

WONDERING ABOUT ECONOMICI AND OTHER SPECIES: A QUIRKY ECONOMIST'S DECONSTRUCTION OF BEHAVIOR, CHOICE, AND WELL-BEING

We exhibit behaviors: working, texting, kissing, and buying M&Ms. Most of us are mostly aware of our actions and perceive most as chosen. But! Do you wonder why you do what you do rather than something else—how much of it you chose — and whether what you do is best for you? I thread these questions around how economists, like me, model behavior: we imagine behaviors are chosen behaviors and that people behave in their own interests. But economic “choice” is inconsistent with the street definition of choice. And what does best mean? And are you truly always experiencing your best available path? I deconstruct the behavioral assumptions of economic choice theory to investigate what we explicitly and implicitly presume about behavior and what we don't. How economists model choice is contrasted with how behavior is modeled by psychologists, neuroscientists, evolutionary biologists, philosophers, and animal behaviorists.____ Common behavioral quirks (e.g., endowment effects, duration biases, and empathy gaps) are discussed. Psychologists believe anyone with them can't be an economicus, but it's not that straightforward.____ Many who model behavior also want to judge it. Most such economists adopt the ethic, welfare economics: increasing the well-being of society's members is right, and decreasing it is wrong. Regular people and moral philosophers reject this criterion, and welfare economists can't fathom why. I consider well-being, or its absence, in other ethics, including Mill's liberalism, Adam Smith's virtue ethics, duty-based ethics, God's rules, Buddhist ethics, and how people really judge right from wrong.

*Behavior, Choice,
Well-being, and
Ethics*

Preface

Preface.....	1
Introduction.....	10
Part I: An ordering of paths, well-being, and economics: how to make you, a chimp, or even a cactus, better off?	22
Chapter 1: Economicus: assumptions of a neoclassical theory of behavior and their implications— my take.....	22
The goal of a theory of behavior is to explain the behavior of individual entities: the sequence of <i>paths</i> each will experience.....	22
A Neoclassical behavior theory (NBT).....	25
Implications of NBT (Assumptions 1-9)	50
What else might be presupposed (implicitly assumed)? What, if any, additional assumptions might you want to add to NBT?.....	54
Thoughts about NBT Plus.....	58
So, what to say about an economicus, a person or thing that follows the rules?,.....	65
Chapter 2: Welfare (well-faring) economics as a theory of what is right and wrong—my take..	73
Welfare economics is a method to test right from wrong.....	73
Welfare economists are consequentialists, specifically welfare consequentialists.....	75
So much for what welfare economists agree on	82
Welfare consequentialism (WC), including welfare economics, scares many moral philosophers because anything goes if it increases WB	84
Are welfare economists utilitarians?.....	87
Besides not being utilitarians (Benthamite, rule, or preference), welfare economists differ from other modern welfare consequentialists.....	94
The moral judgments of regular people: they don't ascribe to the ethics of welfare economics	95
Conclusions and qualifications:	101

Chapter 3: Could economicus be the product of evolution rather than the creation of God?	
Unlikely.....	106
Evolution is driven only by the survival of genes/traits; its objective is not to produce gene containers that are happy or error-free	106
Evolution, since it's typically incremental, gravitates towards a local best rather than the bestest best (the global best)	107
NCT emerged and developed in 19th and early 20th century Britain, a Christian creationist world where man was unique and in God's image	108
Chapter 4: Wanting versus liking, and happiness: an electric14al/chemical stew	111
A brief history of happiness and WB.....	111
Modern happiness/emotional WB: a chemical state of mind	117
So, what is a brain on happiness? Your brain in a nutshell	119
A more macro view: pleasures, emotions, temperaments, and happiness	139
Most economists talk about preferences, not happiness nor WB	149
Chapter 5: Does consuming more increase WB? Is getting rich the path to happiness?.....	154
The evidence on the relationship between income and happiness: the big picture	155
Data and studies on the relationship between income and WB	156
A hypothesis: your WB is determined more by your relative position than by your absolute level of consumption.....	165
Differences in happiness, by income, in the same society at a point in time could, in part, be the result of expectations and aspirations:	170
If emotional or life-satisfaction WB is determined by relative status, what would increase aggregate WB?.....	171
Other data and theories suggest income and wealth are not critical determinants of happiness nor WB.....	174
Chapter 6: Common quirks, incorrect beliefs, and flawed choosing.....	190

Distinguish between NBT violations, common quirks, incorrect beliefs, flawed choosing, and a wrong choice	190
Three + common quirks (duration bias, our emotional empathy-gap, salience effects, and maybe a 4 th , lack of empathy for one's future self). What does each mean for NBT?.....	193
Chapter 7: The endowment effect.....	217
Does the endowment effect cause flawed choosing?.....	220
Explanations for the endowment effect	224
The endowment effect: summing up.....	231
A few thoughts on modeling the common quirks:.....	232
Does NBT survive the common quirks?.....	234
The common quirks, ethics, and welfare economics:	235
The ethical implications of quirks depend on whether the ethical goal is maximizing WB or fulfilling wants. Welfare (well-faring) economics is built on the former. What is right (morally preferred) for the individual is choosing their highest-ranked available WB-path. For society, it is maximizing societal WB in an aggregate sense.	235
Knowledge, beliefs, incorrect beliefs, subjective probabilities, and best estimates	236
Conscious and unconscious mental processes: whose ordering of paths is it?.....	239
Chapter 8: Returning to whether all the different kinds of WB are commensurable and all the different bearers-of-WB are comparable	243
More on why complete WB-incommensurability is important	244
WB-commensurability requires that all the kinds of WB generated by a bearer can be separated from the bearer.....	250
A flawed argument for complete WB-comparability: flawed because it is circular.....	252
Two kinds of WB are WB-incommensurable if an ability to compare them is incompatible with experiencing them.....	254
The kinds of WB produced as you acquire knowledge about a process are WB-incommensurable with the kinds of WB produced by a change in its level.	256

You cannot compare certain kinds of WB because it is morally unacceptable to even imagine such comparisons.....	258
An environmental aside: ecological economists accept as gospel that complete WB-commensurability does not exist.....	260
If you have no control over how much of a bearer you experience, you don't compare it with other bearers.....	262
Inadequate processing skills	264
Saying "Paths i and j are WB-incomparable" does not make it so.....	264
Path incomparability is sometimes mandated by law	264
A few other qualms about WB-commensurability:	265
Ok, so maybe some people cannot WB-rank all conceivable paths, but who cares if most people can WB-rank their current available paths.....	266
Neurological evidence in support of comparability and WB-commensurability	266
Summing up and looking forward	268
Part II: What is a choice? Choice or the illusion of choice?.....	275
Chapter 9: How would you define a <i>choice</i> ? And the difference between a choice and the experience/sense of making a choice.....	275
Necessary conditions for a situation to be a choice?	276
But! Does choice require a choosing experience? Have you made a choice only if you feel you made a choice?.....	276
Contrary to what economists believe and teach, the assumptions of NBT seem to make choice impossible?	276
So, what is <i>free will</i> , and how does it relate to a capacity to make choices?.....	278
Causal determinism.....	282
So, is choice consistent with causal determinism? Compatibilists and Incompatibilists	284
Why do we all feel and believe that we have free will and make choices?.....	287
Chapter 10: The evidence on conscious choice	291

The neurological evidence:	291
Imagine if one never had a choosing experiences	298
Other evidence on choosing experiences and subsequent actions	299
A: believing we had a choosing experience when we didn't.....	300
Summarizing:	306
Part III: Most moral philosophers are not welfare consequentialists, so what are they?.....	310
Chapter 11: A primer on welfare economics vs. other ethics	311
Processists vs. consequentialists	312
Consider the role of WB in different ethics	314
Whose welfare counts (who gets moral standing)?	316
Is Buddhist ethics welfare economics for an economicus who suffers from the endowment effect and other common psychological quirks?	318
John Stuart Mill: a backward-looking utilitarian and a forward-looking liberalist	323
Neither enjoyment, suffering, nor freedom is primary (deontological/duty ethics)	329
Virtue ethics	338
Evolutionary morality theory	339
Contract morality (<i>contractualism</i>).....	340
Ethics based on group rights, things having rights, and inherent value.....	341
All this said about ethics:	345
Chapter 12: A bit more on welfare economics, external effects, and interpersonal comparisons of WB.	349
The ethical goal of welfare economics is maximizing societal WB, not achieving efficiency	349
But, doesn't welfare economics imply that unregulated competitive markets are the right mechanism for allocating resources and distributing goods and services?.....	351

What if we add common quirks, incorrect beliefs, violations of NBT assumptions, and flawed choosing?	355
The practical value of welfare economics is limited because it refuses to compare my WB with yours or even accept that they are comparable	355
Interpersonal WB comparisons are making a comeback amongst economic ethicists: Rawls, Sen, and Dasgupta.....	358
Epilogue: Do you make choices, and if so, has your choosing made you happy?	365

As with all choices, you can now err in two ways: you want to read this book so spend the money and time. But you end up disappointed, so worse off. Or you choose to not read it but would have liked it if you had, so you miss a chance to be better off. You did but shouldn't have, or you should have but didn't.

I've spent 40 years modeling behavior: teaching, researching, and estimating environmental damages. My economic niche includes applied choice theory, welfare economics, and valuing things we like that are not directly bought or sold. The adventure included environmental litigation: estimating \$ damages from environmental injuries caused by large corporations. This required an ability to defend, in litigation, my behavioral assumptions. Because I am anxious and hyper-vigilant, I allocated thousands of hours—a mistake?—to keep that wolf at bay: agonizing about defending my assumptions against assaults from opposing economists and lawyers. The ordeal led me—I was hesitant—on a second adventure: investigating research on behavior and choice in philosophy, psychology, neuroscience, evolutionary biology, and animal behavior. This book is what I have learned and concluded about behavior and how economists model it. While this book is aimed at those with economic interests, it is philosophical—digging down by questioning—hopefully not to fantasy fiction. Don't worry, there is no math(s), but as I tell my students, "It will require critical thinking."

The pandemic has upset our daily routines by adding new constraints and eliminating old ones. For better or worse, family members have been forced to spend more time together and less time with everyone else. We all are at increased risk of sickness and death. Such changes have changed how we now live—a natural experiment for scientists like me who study behavioral changes. As a result, there have been opportunities to reflect on what we do and why. Think about how your expenditures, relationships, and time allocation have changed and why. More or

less exercise? Divorced or newly married? Less restaurant time, more Netflix? New pleasures, new worries? Did you make all the right choices?

The pandemic coincided with my retirement from the University of Colorado. I stumbled into pursuits. Two include more biking and nordic skiing: a goal is to suffer more—a remnant of my Lutheran youth. Besides, anaerobic exhaustion is calming. I've allocated more time to this book. New pursuits include learning to do stuff with my left hand with my eyes shut, understanding first-order predicate logic, and improving my vertical jump. [My upcoming NBA tryout? I started jumping up on rocks while walking the dog. I ended up unable to walk and was scheduled for back surgery. Edward's follies (behavioral errors)?] Have I maximized my well-being? My examples are reallocations of my time: I spent my career thinking more about your demand for different activities than whether you purchase oranges or apples. Needless to say, what I am buying has also changed, but at the end of the day, it all comes down to how you spent your time—you are the limited time you have on earth.

The foundation of applied economics is an individual's behavior can be predicted and explained by assuming the individual is always choosing their highest-ranked available path forward—a behavioral model called *neoclassical choice theory*, NCT. But! most people, including most behavioral scientists, don't believe it. I critique the details—my conclusions surprised me. The distinction between wanting and liking something is paramount.

I start by developing a souped-up NCT, a version that includes everything one might care about, not just market goods and services. I have a life-long interest in valuing things that are not directly bought and sold, like environmental amenities. I allow for uncertainty and risks—*expected-utility theory* is a restrictive example. I deconstruct NCT's standard axioms into their components and consider what it implicitly, but not explicitly, assumes—a lot is implicit.

Entities whose behavior is consistent with NCT are the species *economici* (sing. *economicus*). Economists assume *economici* and *homo-sapiens* are the same species, so *homo-economici*. [NCT is a 19th Century relic of British protestant creationism: humans, but only humans, were created by God in his image.] I distinguish between the two species, asking where in the animal kingdom if anywhere, one might find an *economicus*.

I'm disinclined to believe that human behavior is explained by assumptions fundamentally different from those that explain animal behavior so consider the behavior and

WB of non-human animals. Doing so provides a broader perspective on NCT. All animals sense and perceive: they see, hear, feel, interpret, and respond. All behavior results from the same evolutionary forces. Humans and slugs are 70% identical in DNA; it's 99% with chimps. And, like us naked apes, other animals experience emotions, have relationships, learn, and pass the learning on. We are not that different from a hungry bear searching for breakfast, or Giacomo, my dog, chasing a ball.

Animals, including humans, have beliefs—some consistent with the best science—and beliefs affect behavior. For example, consider the role of beliefs in the recent U.S. Presidential election. Does NCT assume your beliefs are correct?

Emotions—fear, anger, lust, embarrassment, boredom—affect behavior. But sensations, perceptions, and emotions don't play an explicit role in NCT—the economicus brain is a Skinnerian black box. If an economicus has emotions, they run deep; on the surface, he is drab, like faded yellowish/brown wallpaper. What we do depends on what we want, and what we want depends on our emotional state. And humans don't understand that what they want will be different when they are in other emotional states. So, while economics books don't talk about sensations, emotions, and beliefs, I will.

Scientists who model the behavior of animals, plants, planets, and photons deem their models “behavioral”. But economists replace that adjective with “choice”, imagining humans choose how to behave, given their options. But not all behavior is chosen: few humans believe animals choose, and no physicist says the moon chooses its path. So, what is a choice? There is no agreement, and NCT is inconsistent with “choice” in the street sense.

NCT has “choice” in its title because economists believe there is something right, better, or good if what is experienced was chosen by the individual experiencing it. And because economists believe humans choose how to behave. Everyone believes this. The belief that choice is good has its roots in the enlightenment, secular humanism, and the liberalism of John Stuart Mill—it's a common belief. Economists, like everyone else, bring their implicit beliefs to the table.

Welfare (well-faring) *economics* arose in parallel with NCT. Many economists believe it's the way to determine whether a behavior or policy is right or wrong: the ethical goal is to increase social well-being, assuming it's only a function of the WB of each member of society. Others reject this—welfare economists don't understand why. I explain; this requires I contrast welfare economics with WB's role in other ethics such as virtue and duty-based. While one can believe in NCT and reject welfare economics or vice-versa, the two are often conflated, further confusing our understanding of both.

After teaching environmental economics for many years, I developed a critical-thinking course, “Economics, Ethics, and the Environment”—critiquing welfare economics and environmental ethics. Each student had to research a topic that we were both interested in. Because many topics pushed beyond environmental issues, I generalized the course to “Ethics, Happiness, and Choice”.¹ Over five years, this book served as the text. Without my students' thoughtful insights and criticisms, this version would not exist.

ERM June 2022: Boulder CO and Rossland B.C.

¹ I punctuated “...Choice”. rather than “...Choice.”—the right “ to the left of the period. This is called *logical punctuation* (see [Yagoda 2011](#) and [Pinker 2014](#)). A period (or comma) will not appear between the left and right quotation marks unless it is part of the quotation. I am violating the American style rule (it's not a British rule)) that the ending quotation mark comes after the period. The American rule is illogical: logic dictates that what is between quotation marks is what is quoted. This rule exists because a long-dead typesetter decided that sequences like “choice”. disturbs the eye.

Introductionⁱ

08142022

This is a draft: a work in progress

We do things: go to work, watch TV, kiss our partner, and buy M&Ms: we behave in specific ways. We are mostly aware of our actions and perceive them as actions we chose.

This book is for those who wonder why they do what they do and wonder how much they consciously control. Being an economist, I thread the conversation around how economists have historically modeled behavior, but a fondness for economists is not required. This book is for economics students, including graduate and undergraduate economics majors). It is for economic scholars, including the doubters and zealots of economic choice theory. It is also for psychologists, philosophers, and neuroscientists who want to know more about *economicus* behavior—*economici* are the creatures whose behaviors adhere to the neoclassical choice-theory assumptions. Prior knowledge of the theory is not required, so the book is also an invitation to lay readers who want (or would like) a new hobby, rigorously studying behavior and choice.

The distinction between behaviors and choices (a choice being a chosen behavior) pervades my perspective. So does the distinction between what is desired and what will be liked.

It is critical to distinguish between modeling and judging behavior—deciding which behaviors are right and wrong. We all like to judge others and pronounce actions and policies right or wrong. Sometimes our behavior and thoughts are ethical: we are virtuous or do the right thing for ourselves and others. Sometimes we are unethical: we have wrong thoughts and do wrong to ourselves or others. Of course, what you find unethical, I might find ethical. There are numerous ethics (methods for parsing right from wrong), and, not surprisingly, they have different and often conflicting ways of parsing right from wrong. I review prominent ethics and contrast them with *welfare economics*, the ethic many economists use to judge behavior.

In my introductory economics course, I say that economists assume people choose what they perceive best for them, subject to their circumstances (constraints). And—what they perceive as best for them is, in fact, best for them. At every point, you choose your best path forward. Everyone else does the same—no one makes well-being-reducing choices. If I appear to

act contrary to my well-being (*WB*),² it is because I have unusual tastes or constraints. This is, simply put, the conventional economic choice theory: assumptions meant to predict and explain what you will do and why. It is called a “choice” theory because economists believe behaviors are chosen behaviors, in contrast to believing you have little control over what you do or little conscious control over what you do.

Chapter 1 lays out the explicit and implicit details of a/the conventional economic theory of behavior. I will call it *Neoclassical behavior theory*, *NBT* for short. I occasionally refer to it as “my NBT” because it is my modern take on neoclassical choice theory. It emerged in the 19th Century’s last half as the economic theory of human behavior. It remains the one first taught, and it is the foundation for almost all empirical research in economics. Many of its assumptions are implicit, and some seem incorrect.

To behave according to this theory has implications for how brains work. Since NBT is my conception of modern neoclassical theory, I will use “NBT” to refer to Assumptions 1-9 of Chapter 1 and neoclassical choice theory in general, unless the context is the theory’s historical development. Then I will use “neoclassical choice theory”, *NCT*.

If you behave according to NBT, a researcher only needs to identify your constraints (which paths are available) and your ordering of them to predict your behavior. She would then predict that you will move to Maine with Mary and become a stonemason if Maine, Mary, and masoning are part of your highest-ranked available path.

If a path is not available, one cannot experience it. My wife desires to dine with Springsteen, but paths that include Bruce accepting her invitation are not available. For me, Bruce is also out, along with drinking great wines and the apartment overlooking Central Park. In contrast, for Bruce, who suffers bouts of depression, depression-free paths and paths of anonymity are not available. While economists are explicit about which paths you can’t afford money-wise, many other factors constrain you (e.g., your health, gender, age, fitness, culture, morals, and obligations to friends and family). But economists typically do not model these as explicit constraints.³ It is argued that they are not constraints. This is partly because they can be

² Going forward, I will abbreviate “well-being” with “WB”.

³ For now, the pandemic has imposed many constraints, including where we can’t go, and our increased risk of sickness and death.

challenging to model, so it is easier to say that such things affect your ordering of paths. It is also because economists interpret actions as choices rather than constrained behavior. [“The new mom chose to feed her new baby”, rather than “The new mom had to feed her baby.”] I attribute more behavior to constraints than many economists, but I am not alone.⁴

Determining someone else’s ordering is difficult. A researcher could simply ask whether you rank higher Path q or Path g , but many economists are hesitant to believe that if you answer “ q ”, you rank q over g . So, economists simply assume that if an individual chooses q over g , she must rank q over g —whatever one does must be best for them, given their constraints. So, if I drink motor oil and beat the kids, it must be a component of my best path, given my constraints.

While NBT assumes an ordering of paths and selecting the highest-ranked available path (HRAP),⁵ both assumptions are questionable. Maybe you don’t know your ordering or don’t even have an ordering. Moreover, while many believe they act in their own best interests (choose their HRAP), they are convinced that others don’t. It is difficult to look at a specific choice and determine whether it is inconsistent with the theory—NBT allows for wild and crazy behaviors.

Research by psychologists, behavioral economists, and neuroscientists demonstrates that humans have common quirks that can lead to what I will call *flawed choosing*, suggesting that we won’t always experience our HRAP.⁶

An important question is whether these common quirks violate the NBT Assumptions. Two common quirks are *duration bias* and *empathy gaps*. Duration bias is we fail to appreciate

⁴ My research has focused on constraints such as skills (e.g., skiing, fishing, and biking), personality (e.g., extroversion, and competitiveness) life constraints (kids, BMI, fitness, and health), and mental health. These are all constraints that are exogenous to the individual when a path must be selected. They are different from how behavior is affected by what is deemed unacceptable.

Adam Smith, in his *Theory of Moral Sentiments* (1759), argues that many behaviors are avoided simply because they violate “moral sentiments” (one’s moral sense): selecting Path j is simply wrong so is not selected. More recently, Sen (1989) formally distinguishes between which paths are available given the external constraints imposed on the individual, the set S , and the subset of those available paths that are personally admissible, the subset $K(S)$: the individual selects a path from $K(S)$.

⁵ Abbreviating is always a questionable practice. I do it in this case because the expression appears many times and often multiple times in the same paragraph. If it has been a while, I will repeat the full expression.

⁶ Flawed choosing needs to be contrasted with psychological terms such as *abnormal behaviors*, *mental disorders*, *mental illnesses*, and *pathologies* of mind, mood, or behavior. Abnormal and pathological behaviors do not necessarily violate the assumptions of NBT, and flawed choosing does not imply the individual has a mental disorder.

that we will adjust to both good and bad things. What makes us happy today (e.g., yesterday's pay raise) will not make us happy forever. And, even though we would be distraught immediately after a diagnosis of a permanent disability, we will adjust more than we can imagine, eventually returning to a state of happiness (or unhappiness) that can approach what it was before the diagnosis. Duration bias is we incorrectly imagine we won't adjust and make choices based on our false premise. For example, when I am anxious, I am inclined to believe it will persist unless I take a pill or have a drink; mistakenly, I believe the anxiety won't abate by itself, so I drink and pill too much, or maybe not.

Empathy gaps are a lack of empathy for one's other selves. One can lack empathy for their future self or oneself in a different emotional state. Empathy gaps start with the hypothesis that your ordering of paths varies with your emotional state. The gap is that individuals choose as if they will always have their current ordering. For example, when you are angry, you order paths differently and choose based on your angry ordering, so often choose a path that is not highest-ranked after the anger abates. Research on duration bias, the *emotional empathy-gap*, the *future-self empathy-gap*, and *incentive salience effects* are presented in Chapter 6; Chapter 7 covers the *endowment effect*.

While economists assume the ordering is a preference ordering, economists are vague on why a path is "preferred". Economists presume if you rank Path w higher than Path e , you would be better off with w , **or** you desire w more than e , **or** both. The difference between w is better and w is more desired is critical, but economists do not distinguish.

But what does "better-off" mean? Does it mean more pleasurable sensations (more orgasms and less pain)? Does it mean experiencing more positive thoughts and fewer negative thoughts—more "I'm successful." Less, "I'm a bad person."? Does it mean more *emotional WB* (experiencing more positive emotions and fewer negative ones, e.g., more happiness, less sadness)? Does it mean more life satisfaction (*life-satisfaction WB*)? It depends on who you ask, and "better off" often goes undefined. Going forward, I define better off as more WB, and WB as a combination of pleasurable sensations, positive thoughts, emotional WB, and life-satisfaction WB. [Not all species are capable of experiencing them all.] Confusingly, "happiness" is often used as a catchall. Researchers and authors are typically not specific.

What it means to be happy has drastically changed over the ages (Chapter 4). The current notion is that happiness is a chemical/electrical brain state—many of my students hate that happiness is only a choreographed dance of brain synapses, choreographed by God-knows-who or what.

Turning from explaining to judging behavior, a researcher can build a theory to explain and predict what you will do without judging whether your behavior is ethical. But welfare economists want to judge behaviors and policies right or wrong—they want to make ethical judgments. Welfare economics (defined in Chapter 2) assumes that a society's ethical objective should be maximizing its citizens' welfare (well-faring). It is a type of *welfare consequentialism*: how an act affects the WBs of society's members should be the only determinant of whether it's right/ethical/moral. This raises the question of whose WB counts: who is a member? For example, in Canadian society, do Italians count? And, if so, as much as Canadians? How about beavers and moose? Economists presumptuously exclude the WB of beavers and moose, but there isn't anything in NBT that specifies their omission. Consider the implications of expanding the notion of who counts to other species even if you believe their WB should not count. When it comes to humans, welfare economists punt on the issue of which humans should count.

Few moral philosophers are welfare consequentialists.⁷ The few that are reject the WC of welfare economics.

People reject WC: believing that while how an act affects everyone's WB should be a criterion in judging whether it is right or wrong, it should not be the only criterion.⁸ Unfortunately, welfare economists don't deal with the fact that others reject WC as the right ethic to parse right from wrong.

⁷ I define *moral philosophy* and *moral philosophers* as follows: Moral philosophy has two separate meanings: it can refer to the study of how one parses right from wrong behaviors and inclinations, and it can also refer to a set of specific assumptions that can be used to parse right from wrong. To avoid confusion, I will refer to the later as an *ethic*. Moral philosophers are individuals who study ethics. It is important to distinguish between moral philosophers and individuals who adhere to a specific ethic. One can adhere to a specific ethic (e.g., right and wrong are determined by God's words) without being a moral philosopher. One can be a moral philosopher that adheres to no specific ethic, or one can be a moral philosopher that defends a specific ethic such as virtue ethics or utilitarianism. At one extreme is the Oxford philosopher Bernard Williams (1929-2003) who suggested that philosophy might have nothing to contribute to ethics (Williams 1985 and 2006).

⁸ "Act" or "action"? Much of the time they are synonyms, so interchangeable, but not always. I use both, going with the one that sounds, to me, more natural, given the context.

Using WC (including welfare economics) to judge right from wrong is difficult—other ethics share this difficulty. Consider first a society of only one individual—George. Since WC does not have to consider how his behavior will affect others, the best ethical outcome, from the perspective of WC, is the outcome best for George. If George always chooses his best path, the moral imperative is to constrain him as little as possible—a world of no laws or restrictions but vast resources—the Tea-Party dream. If, alternatively, George, left alone, would sometimes not choose his best path, WC suggests sometimes forcing him, for his own good, to do things he would not choose to do. For example, George would be better off if a foreign agent stopped him from drinking motor oil.

Things are much more complicated if society has more than one member, and their behaviors directly affect each other. For example, consider a society of two people: George and Fred. George paints his house purple and fills his yard with plastic statues of Snow White and the Dwarfs: the visual effect brings joy to his heart. But the display makes Fred, the neighbor, sick. On the other hand, Fred loves to get drunk and belt out—badly, and over and over—the swan song from the tragic opera “La Traviata”. In this case, WC requires that George’s welfare be weighed against Fred’s welfare, but welfare economics has few rules for doing this.

Welfare economists can’t imagine ethics where whether an act is right or wrong is independent of WB. But such ethics are common. I was brought up with an ethic that right vs. wrong was all about serving God, honoring God, and following his rules. The objective was not to experience pleasurable sensations: most were sins. Rejecting WC isn’t limited to religious ethics: Kantian ethics, for example, emphasizes doing one’s duty—happiness is ok but not the objective. Virtue ethics is about being virtuous, not about being happy.

Returning to only explaining choice and my second sentence, “Most of us are mostly aware of our actions and perceive them as our own, actions we choose.” We believe we make choices, but what is a choice, and do we really make choices? What does it mean to choose Path *e* over *c*? Part II considers what it means to choose and whether we make choices. Let’s agree that if Path *c* is the only available path, there isn’t anything to choose—choice requires two or more available paths. For example, prison inmates don’t choose to stay in on Saturday night—

going out is not available, and people with no money and no credit can't choose to buy a new Mercedes or choose not to. In contrast, people with money who are not incarcerated can afford—money-wise—many different paths of goods and services, so they seem to have more choices.

It is tempting to conclude that the goods you buy and consume are the goods you chose to buy, and this is the conclusion everyone jumps to, including economists. But is it true? If you have \$100 to spend on groceries, there are thousands of different bags of groceries that would cost \$100 or less. If you walk out of Safeway with 3 pears, chocolate cookies, and a steak, a bag costing \$58, maybe your brain, body, and upbringing constrained you to buy that bundle.

Economists typically think of constraints as external to the individual, but you are also constrained by who you are. Heavy people cannot win bike races with mountain stages, and I can't go downhill fast on my mountain bike because it would make me too anxious—anxiety is a chemical brain state I was born with. Besides—my wife won't let me ride a motorcycle, but she will let me go heli-skiing.⁹ What is a choice cannot be assessed independently of a constraint? Consider how current choices (marriage and donuts) lead to future constraints. The ultimate constraint is reality, and the ultimate reality is your weeks are numbered—four-thousand, give or take ([Oliver Burkeman](#) 2021).

As noted above, many people, including behavioral psychologists, reject the economic assumption that people always behave in their best interests, arguing and demonstrating that people make bad choices. But a bad choice is still a choice. As will be discussed (Chapters 9 and 10), numerous prominent physicists and neuroscientists don't believe we make choices in the street sense of the word. For example, if your unconscious determines you will propose to Wilma, and before you buy the ring, your conscious brain provides you with the experience of choosing to propose, did you choose to propose to Wilma?¹⁰ A substantial body of neurological evidence suggests that the final determinant of whether you do x or y is unconscious. But, before the behavior is initiated, the conscious part of our brain often provides us with the perception of

⁹ Heli-skiing makes me anxious but includes moments of euphoric pleasure.

¹⁰ A few readers of earlier drafts have insisted that unconscious means “passed out”, and that I should use “subconscious”. In my defense, quoting the [Michael Craig Miller, M.D.](#), Senior Editor, Mental Health Publishing, Harvard Health Publications, “As a general rule, then, in most of the professional literature where mental functioning is concerned (including not just psychoanalysis, but also psychiatry, psychology, and neuroscience, among others), writers—like Freud—tend to use the word ‘unconscious’ rather than ‘subconscious’. Although the word ‘subconscious’ continues to appear in the lay literature, it is rarely defined carefully and may or may not be synonymous with ‘unconscious’.”

choosing, an *illusion of choice*. I review this research; you can “choose” how much of it you choose to believe.

There is also the physics of *causal determinism*: what you do this moment, and the next, is entirely determined by the world’s configuration of atoms and the laws of physics, including the atomic configuration of your brain. Both causal determinism and behavior determined by the unconscious brain raise the issue of whether you should be held responsible for your behavior, be it right or wrong.

Economists implicitly adopt a weak notion of *free will*. Free will is a common expression in philosophy; economists would define it as the ability to freely make choices if they thought about it. Philosophers differ regarding what it is and whether it is consistent with choice. If we want to explain behavior, economists need to ask how much of behavior is chosen, starting by understanding the neurology of choice and the philosophical disagreements about choice.

All the above will be investigated referencing recent scientific research. The goal is to introduce you to the philosophy and scientific research that has helped me evaluate economics as a social science and welfare economics as an ethic. Of course, many students of economics are already versed in parts of this literature. My mental reveries scratch surfaces in philosophy, ethics, psychology, neuroscience, evolutionary biology, and animal behavior. Wandering is fun and has led me to think about behavior and choice in new and weird ways, and I hope to convey a bit of that weirdness. Looking ahead, zombies, the philosophical kind, will be discussed, asking whether a zombie can be an economicus.¹¹

A few warnings are appropriate: what is new here is how the different literatures are brought together, packaged, and interpreted, along with my deconstruction of NCT: I am presenting research by others by analyzing and questioning. I believe I bring new insights and contributions to the discussions. Since there is ignorance and disparate views on how brains work, research findings and speculations are sometimes in conflict— so question the research

¹¹ There are movie zombies, voodoo zombies, and philosophical zombies. [Philosophical zombies](#) appear in academic discussions of consciousness. Philosophical zombies and humans are indistinguishable, but, in the case of zombies, no one is home.

and pursue the conflicts, starting with the references and links. The goal is to introduce and critique the literature, not revere it.

There will be environmental examples because I have spent my life studying environmental economics. Thinking about the environment, including its living non-human inhabitants, highlights specifics of the economic approach to ethics and how it is distinct from other ethics. For example, if individual animals have moral standing and people make them suffer, there are drastic implications for what is and is not economically efficient. These implications are worth considering, even if you believe non-human animals shouldn't have standing. A related belief, rejected by economics, is an ecosystem, a species, or a place (e.g., the Grand Canyon) can have value separate from the value you and others ascribe to it. I survey non-economic notions of value.

My examples will push the boundaries of propriety and taste—this is intended. There will be religious examples: religious doctrine is how many distinguish right from wrong (sin). Sexual desire, drugs, and perverse behaviors make excellent examples; they push buttons. Sex is conspicuous in ethics—for instance, is sex for fun moral or a sin? Does its morality depend on which body parts are involved or who it is with, and if so, why? The implications of our ethical assumptions are often only appreciated when they take us to a place we do not want to go.

Final warning: would learning what you now believe is incorrect make you happier? Consider what [Dr. Victor Frankenstein](#) had to say, lamenting the creature he created,¹²

You seek for knowledge and wisdom, as I once did; and I ardently hope that the gratification of your wishes may not be serpent to sting you, as mine has been... Learn from me ...how dangerous is the acquirement of knowledge and how much happier that man is who believes his native town to be the world, than he who aspires to become greater than his nature will allow—[Mary Shelley](#)

Do you know what will make you happy, and, if so, will you take that path? Dr. Frankenstein did not. Could he have chosen a different path? Do any of us choose our path?

While I carefully chose the order to present the topics, and while they collectively form a whole, nothing terrible will happen if you read the chapters in the wrong order or skip some: most stand alone or at least teeter alone. Jumping around within a chapter isn't advised, but skimming over a section is allowed. The book is divided into three parts. The first introduces and

¹² There will be a link to a person when their name first appears in the text. If a name only appears in a footnote, the link will be in the first footnote it appears. Links appear in blue.

critiques an economic theory of choice, discussing many research experiments in psychology designed to learn how choices are made and whether individuals typically act to pursue an overall goal such as well-being or the fulfillment of desires. Part I also introduces and summarizes welfare economics. Part II digs more deeply: asking, “What is a choice?” Do humans and other animals make choices in the street sense? Part II draws heavily from the neurological research on the experience of conscious choosing, the physics of choice, and the philosophical literature on free will and choice. Part III adds to our conversation about welfare economics and contrasts it and WC with other ethics such as virtue, Kantian, and Buddhist. Finally, the end of Part III introduces the ethics of prominent ethicists who are economists but neither welfare economists nor even welfare consequentialists.

You can determine a lot about a book by knowing how often a word, phrase, or name appears. *Utility* and *preferences* don’t much: both are suitcase words—packed with various meanings.¹³ In contrast, *WB* (for well-being) often appears, often with the adjective *emotional* or *life-satisfaction*. *Behavior* will be ubiquitous: it is critical to distinguish between it and a choice (a chosen behavior), same for *wants/desires* vs. *likes*. You will encounter *neurons*, *synapses*, *readiness potential*, and the *mesolimbic dopamine system*. You will encounter *anxiety*, *happiness*, and *emotional-dependent orderings*; philosophical terms include *will*, *free will*, *duty*, and *virtue ethics*. Other vital words and phrases include *beliefs* and *incorrect beliefs*, *bias*, *flawed choosing*, and *violations of choice theory*. Non-economist names you will encounter include Mark Bear, Jeremy Bentham, Kent Berridge, Daniel Dennett, Rene Descartes, Leon Festinger, Michael Gazziniga, Dan Gilbert, Jonathan Haidt, David Hume, Henry and William James, Daniel Kahneman, Benjamin Libet, J.S. Mill, Shaun Nichols, Richard Nisbett, Martha Nussbaum, Derik Parfit, Frank Ramsey, Jean-Jacques Rousseau, Bertrand Russell, Peter Singer, Cass Sunstein, Timothy Wilson, and Daniel Wenger.

¹³ Ken Binmore (2021) argues that this is the word’s charm: in his assessment, maximizing utility simply means one behaves consistently, and it really does not matter what the individual is being consistent with respect to. So, there is no need to equate utility with happiness or well-being. He does note that people often do not behave consistently. Binmore notes the word traces “its origins to the Latin *utilitas*.” and its use by “Daniel Bernoulli [1700-82] in discussing the [St. Petersburg Paradox](#).”

I have attempted to write as if we're having an informal conversation between friends. I will direct your attention to exciting and insightful theories and studies of behavior. The only difference between you and me is that I have spent more time studying this literature. I avoid words and expressions you don't need to know; I will not repeat everything three times; I won't hedge or overly qualify. [e.g., I mostly omit the qualifier "often"; "most" is challenging to get rid of.] I won't include a lot of *signposting* ("Now, I will tell you about... Then I will talk about..."). I will try to be specific rather than abstract. Since it is a conversation, there will be many "I", "me", and "you", and you and I will appear in examples.

Since you are part of a friendly conversation, you have obligations; one is giving me a break with "qualifiers" like "mostly" and "might". When anyone discusses a theory or research study, it is always possible to immediately and legitimately object: descriptions and interpretations often need to be further qualified to make them bombproof. (Lawyers try to make legal documents bombproof, resulting in a thickness of prose that overwhelms nonlawyers).

If, with effort, you can't comprehend what I am trying to say, I am at fault, but you will have to think hard in places. So, please, give me enough rope and stick to garrote myself. I am an academic, so writing clearly is an alien endeavor: success is not guaranteed.

But, if we are just talking, why all the footnotes—conversations don't have footnotes? Because footnotes are critical to the task at hand. References appear as endnotes at the end of each chapter, whereas the footnotes are details and asides about things on that page—conversations do include asides. Footnotes are either hated or loved. Authors often compromise and put the footnotes at the end, making them endnotes—readers who like footnotes hate this. Don't conclude the footnotes mean the book is a tome—hopefully, it isn't. And feel free to ignore them. But if you want to dig deeper, they and the references are the places to start. View them as links embedded in the transcript of our conversation: the book is a guide for exploring further, so it requires clear directions (footnotes and references). Have fun, and if you have questions or comments, email me at Edward.Morey@colorado.edu. The latest draft can be found at EdwardMorey.org. Thanks.

ⁱ Sources:

1. Ken Binmore (2021) *Early Utilitarians: lives and Ideals*. London, Springer.
2. Michael Miller (August 12, 2016) Unconscious or subconscious, *Harvard Health Blog, Harvard Medical School*. At <http://www.health.harvard.edu/blog/unconscious-or-subconscious-20100801255>
3. Oliver Burkeman (2021) *Four thousand weeks: time management for mortals*, NY, Farrar, Straus, and Giroux.
4. Edward Morey 1981. "The demand for site-specific recreational activities: a characteristics approach". *J. of Environ. Econ. And Manag.* Vol. (4): 245-271.
5. Edward Morey Jennifer Thacher, and Bill Breffle 2006. "Using angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model". *Environmental and Resource Economics*. Vol. 34(1): 91-115.
6. Edward Morey, Terry Buchanan, and Donald Waldman 2002. "Estimating the benefits to mountain bikes of changes in trail characteristics, access fees, and site closures: choice experiments and benefits transfer". *Journal of Environmental Management*. 64(4): 411-22.
7. Edward Morey and Mara Thiene 2012. "A parsimonious, latent-class methodology for predicting behavioral heterogeneity in terms of life-constraint heterogeneity". *Ecological Economics* 74: 130-44.
8. Edward Morey and Mara Thiene 2017. "Can your personality explain where and with whom you recreate? A latent-class site-choice model informed by estimates from a mixed-mode LC cluster model with latent personality traits". *Ecological Economics* 138: 223-37.
9. Martha Nussbaum 2012. "Who is a happy warrior? Philosophy, happiness research, and public policy". *International Rev. of Econ.* 59: 335-361.
10. Stephen Pinker (2014) *The sense of style: The thinking persons guide to writing in the 21st Century*, NY, Viking, published by the Penguin Group.
11. Mary Shelley 1818. *Frankenstein; or, the Modern Prometheus*. At <http://books.google.com/books/about/Frankenstein.html?id=5twBAAAAQAAJ>
12. Jennifer Thacher, Edward Morey, and Edward Craighead 2005. "Using patient characteristics and attitudinal data to identify depression treatment preference groups". *Depression and Anxiety*. Vol. 21(2): 47-54.
13. Bernard Williams (1985, 2006) *Ethics and the Limits of Philosophy*. Oxford: Fontana Press 1985, Routhledge 2006.
14. Ben Yagoda 2011, "The rise of 'logical punctuation': The period outside the quotation marks isn't a copy error". *Slate* May 12th. <https://slate.com/human-interest/2011/05/logical-punctuation-should-we-start-placing-commas-outside-quotation-marks.html>

Part I: An ordering of paths, well-being, and economics: how to make you, a chimp, or even a cactus, better off?

Chapter 1: Economicus: assumptions of a neoclassical theory of behavior and their implications— my take

01172022

“My take”: My account of neoclassical choice theory is likely different from yours and the take of others. This must be so if this chapter is worth reading.

The goal of a theory of behavior is to explain the behavior of individual entities: the sequence of *paths* each will experience¹

There is more than one theory of behavior and choice in economics, the same in psychology and neuroscience.¹ Here I present a theory of behavior fundamentally neoclassical, an extension of what is taught in principles-of-economics courses. Its first nine assumptions are my take on neoclassical choice theory. While I teach it to undergraduates, I don’t necessarily take these nine assumptions as truth or the best theory of behavior—in my defense, I am not the only economist in this boat.

The *neoclassical behavior theory* presented here, hereafter, *NBT*, allows for risks and ignorance. It extends the notion of consumption, and consumption paths, to include the environment, public goods and bads, aspects important to economists in labor and other applied fields such as urban/regional and demography. It extends the notion of a path to include all aspects of what might make any economicus better or worse.

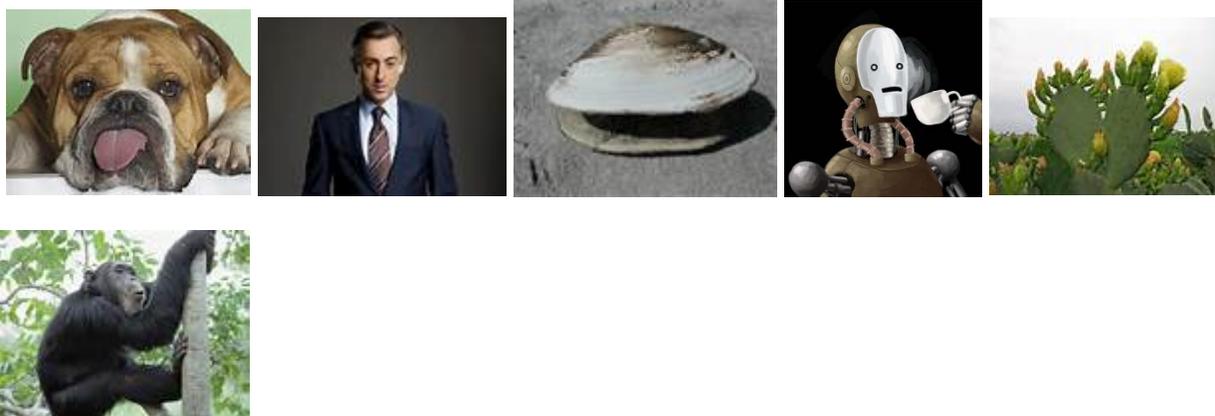
It is “an NBT” rather than “the NBT”. I will, in the assumptions, avoid the word *choice*: economists tend to use the word *choice* rather than *behavior* because they assume human behaviors are chosen. Think of choice theories as a subset of theories of behavior, keeping in mind that many don’t assume behaviors are consciously chosen. Think of NBT as a framework for discussing behavior and asking questions about behavior and choice. It is presented as a theory of how entities behave, not how they should behave. And, at this early point, the issue is

¹ While economists, emphasize the word “choice” other fields use more neutral terms such as *decision theory*, *judgement*, and *behavior*. As an alternative to the neoclassical perspective, there is the newish field of *behavioral economics*.

understanding the theory, not whether it does or does not explain behavior. Later, research will be presented on whether you and others behave in a manner consistent with the assumptions of NBT. Be clear that I am offering my take on modernized neoclassical choice theory, not embracing it.

A dividing line between theories of behavior is whether their intent is only to predict behavior without judging it. Or whether it also predicts what you do is best for you, given your constraints. If the behavior predicted is deemed ethically right, ethics has crept in: there is something right about the behavior. Contrast a choice theory with, for example, a theory to explain the moon's position relative to the earth (Newtonian physics). It would never cross a physicist's mind to suggest their theory implied there was something right or wrong about the moon's path. In contrast, economists believe that choosing your highest-ranked available path, HRAP, is, in many situations, the right thing to do.

Simply put, NBT assumes an economicus does what is best for it, given what is available.² This chapter provides details. An economicus is any entity whose behavior is consistent with my first 9 Assumptions. They might be plants, animals, machines (Rob the Robot, or a toaster), or even humans—whether you believe economici exist is up to you. While NCT was postulated to explain the behavior of humans, many humans are convinced it does not apply to other humans. The following entities might be economici; it depends on whether they follow the rules of NBT.



² How one conceives of and defines “best” will be a big deal.

Dog, Alan Cumming, clam, robot, cactus, and chimpanzee

At this point, I will not restrict NBT to humans: doing so could mean there are no real economici—maybe economici are mythical, like angels and dragons. The behaviors of clams, chimps, and humans all have the same evolutionary root, and all are motivated to propagate their genes. If a theory of behavior explains the behaviors of complex organisms, like humans, wouldn't we expect it to also explain the behaviors of simpler organisms, like clams and cacti, so why exclude them? I worry that if a theory can't explain the behavior of a bivalve, how could it possibly explain the behavior of more complex organisms? Considering clams makes us think more deeply about behavior and choice. Psychologists, neuroscientists, and behavioral economists believe that economici don't exist. If so, and we want to find a real one, we should examine the behaviors and motivations of dogs, clams, and toasters. I would argue that my NBT, Assumptions 1-9, applies to clams and toasters—yes, toasters. I'm not so sure about dogs and cats.

We also need a theory of animal behavior if we want to consider the welfare of non-human animals, so why not a theory that covers all animal behavior? [Don't worry! I'm not going to sweat whether toasters are happy.]

As with all scientific theories, a behavior theory starts with a set of definitions and assumptions that contain no contradictions. Logical deduction determines what they together imply/predict. What goes in [the assumptions] determines what comes out [the predictions]. Another name for a prediction is a hypothesis: *if such and such conditions hold, a certain outcome will result*. A theory might predict that if the price of gasoline rises, you will buy less, and it might predict that if one wears a red dress, the probability of a Saturday night date increases. Keep in mind the distinction between what the assumptions imply and whether you, and others, act in a manner consistent with them. Later chapters present the evidence on whether humans and other animals do.

While you and I would like to believe we are logical, humans, as a species, didn't evolve to be logical or to limit our thinking to the assumptions on the page: we unconsciously add our

beliefs, making it difficult for us to avoid faulty logic.³ So, when we consider what the assumptions of NBT imply and don't imply, we need to make sure we don't add in our beliefs as implicit, additional assumptions. Since many of you, and I, have familiarity with NCT and what it predicts, we will be tempted to suppose the standard predictions before I have specified all the necessary assumptions. To keep us on track, there will be sentences, "Assumptions 1-X neither preclude nor imply..."

A Neoclassical behavior theory (NBT)ⁱⁱ

NBT is simply a set of definitions and assumptions. Since what a theory predicts depends precisely on what is and is not assumed, the definitions and assumptions must be precise. NBT consists of two definitions and nine assumptions: the standard neoclassical assumptions are deconstructed into their components.

NBT is limited to predicting the behavior of *economici*.

An economicus is an entity whose behavior and objective are consistent with Assumptions 1-9.

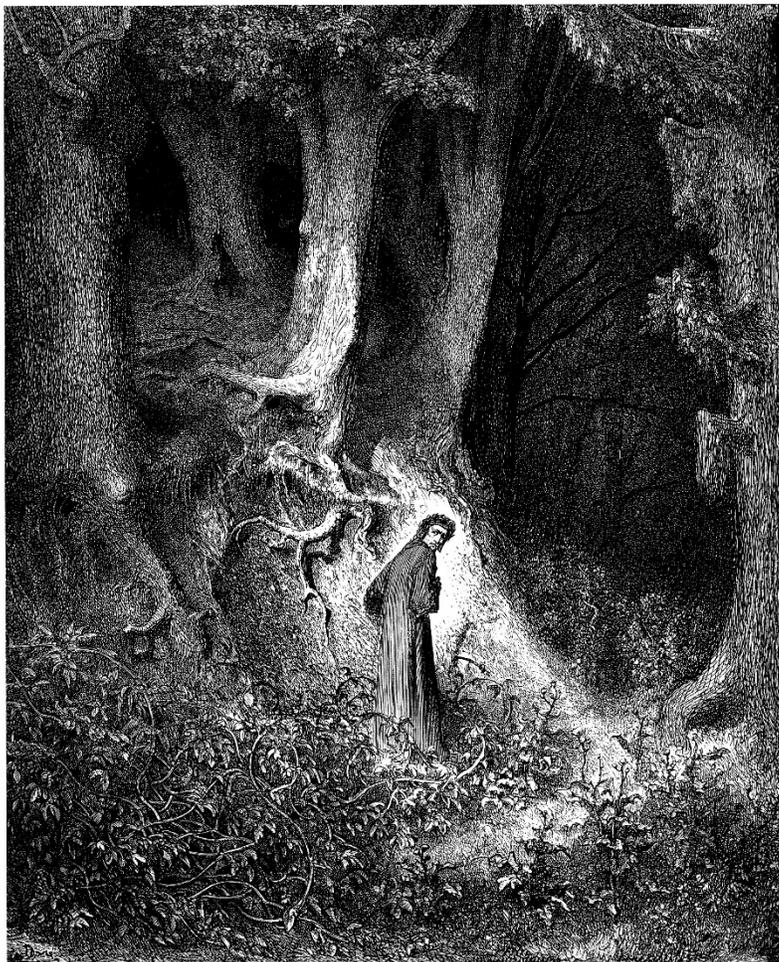
The definition does not imply that the entity is a human: that would be a *homo economicus*, a particular type of economicus. The definition does not even imply that the entity is alive: it could be a computer program, a robot, or even a fridge, as long as its actions are consistent with Assumptions 1-9. These assumptions don't imply human economici, but it is not precluded.

If you look at the descriptions of NCT in texts, they never say the theory applies to humans or only humans. The only suggestion of humanness is the word "*individual*," which my wife thinks only applies to people. You can, if you insist, restrict, by assumption, the theory to homo economici, but keep in mind the possibility that there are no homo economici.

Before the details, a curt description: the assumption, an economicus; the method, logic; the conclusion, behaves like an economicus.

A path:

³ For example, many bright students accept the non sequitur: "Living things need water; fish need water, so fish are living things." They only see that their reasoning is illogical when I replace "fish" with "washing machines". Why the faulty logic? Humans know that fish are alive, so "fish are living things" is, by itself, a correct statement, and so miss the "so". They mistake water being *necessary* for many living things with water being a *sufficient* condition for life.



Midway upon the journey of our life / I found myself within a forest dark, /
For the straightforward pathway had been lost.

Inf. I, lines 1-3

Figure 1, Chapter 1: Dante, the Inferno, Gustave Dore (illus.)

A path for an economicus is an imaginable life as it would unfold through time, starting now. It includes the goods, services, and inputs the entity would consume/use through time on that path (its consumption path). It includes its job, what it would know at different points in time (its knowledge path), what other entities would consume (other's consumption paths), what other entities would know, plus the entity's relationships (social, sexual, work), the relationships of other entities, and the environment (who is now and will be the future Prime Minister of England, air qualities, crime rates, etc. Uncertainty and ignorance are part of every path, and so are beliefs ⁴

⁴ Another term for path would be *state-of-the-world*. I picked *path* because I will discuss different types of states (emotional states, brain states, states of nature, etc.). I am taking the concept of a *consumption bundle* and broadening it to include everything that might affect an economicus. In an earlier draft, I used *bundle*, but found that a bundle is interpreted as something experienced now, rather than a path through time (a sequence of present

For economici, there are typically many paths forward, but all end. A life (an existence) is a sequence of paths taken. An economicus' paths are not quite like those in Yosemite: they can't be loops, and one can't turn back, and the entity must continue forward at the speed of time: deviations would be time travel to the past or future.

A goal of NBT is to predict the sequence of paths an economicus takes. [Keep in mind that an economicus's behavior is not influenced by all path components.]

If you were taking an introductory microeconomics course, a path would be defined by only the amounts of each market good you would consume on that path, and it would be called a *consumption bundle*. But the behavior of humans is influenced by much more than market goods.

There is a path, call it *Path k*, in which I would be married to Shirley for eight years, have a poodle named Fred, smoke Camels, no one would be starving, and Trump is re-elected in 2024. On another path, maybe *Path f*, I would be married to George, not know about global warming, have a cat, drink only zinfandel rose, many people would be starving, and Michael Bloomberg would be the U.S. President. Each sequence of paths taken is a different life. Since economici don't have perfect information, paths include uncertainties.

Note the word *imagine* in the definition, and remember that humans can imagine things that are not possible, and *imaginable* does not mean *available*.⁵ So by my definition, a path can't have non-imaginable components. But what is imaginable can change over time and varies across individuals: it depends on what they have seen and experienced. For example, N. American plains Indians in the 1700s likely had sets of paths vastly different from their contemporary, Adam Smith. Consider animals; humans have bigger imaginations than chimps, and chimps' are bigger than dogs, so my set of imaginable paths is bigger than my dog's: Giacomo has never imagined a trip to Bangkok or the taste of 1961 Chateau Rothchild. At this point, you are asking, "What's with the imaginable restriction? I have never heard an economist use the word imagination." I will explain but note that it limits the number of paths.

bundles). This expansion beyond the consumption bundle isn't new; John Harsanyi (1982), [the 1994 Nobel laureate in Economics](#), included "health, social status, job situation family situation, etc." I also considered the word *prospect* where a prospect is set of uncertain outcomes, each associated with a probability of occurrence. The words *prospect* and *gamble* are often used when considering choice when there are uncertainties.

⁵ If an entity has experienced something or is programmed to do something, it is imaginable. But imaginable paths are not limited to experienced paths.

The definition of a path ended with, “*Uncertainty and ignorance are part of every path, so are beliefs.*” “Uncertainty”, as I am using the term, neither implies nor precludes *probabilities of occurrence* associated with uncertain events. While every path includes uncertainties, the definition of a path does not require that these uncertainties be expressible as probabilities, where the probability of an outcome is a number between zero and one, inclusive, and the probability of something happening is one.⁶

Your beliefs consist of things that you think are correct: they represent that which is the case. For example, Bob’s belief that “the world is flat” means that somewhere in his neurons is stored a representation of this subject and predicate. And it bubbles up whenever the earth’s shape becomes relevant to behavior. [I don’t walk around constantly thinking about how 2 plus 2 is 4 but become conscious that it is when I need to do sums in my head.] Beliefs are often described as “affective attitudes”, “affective”, in this case, meaning they affect behavior. Yours are yours, and mine are mine. They include how you think the world works, including what you think will happen if Event A happens; will it cause Event B or C? Most beliefs are mundane, e.g., my belief that squirrels have tails and I have feet. Beliefs vary in strength: how confident you are that your belief is correct: your confidence in it being the way you think it is.⁷ People express the strength of a belief with terms like “probable”, “unlikely,” and “for sure”, but they can also be expressed as probability numbers; I will call these *subjective probabilities* to

⁶ For example, if the only two possible outcomes are rain tomorrow or not, then probability of it raining, $Prob(r)$, or not, $Prob(nr)$, is one. Probability is often formally defined as follows: for all possible outcomes ($k=1,2,\dots,K$): $Prob(k)\geq 0$ for all k , $Prob(\text{some } k)=1$, and $Prob(j \text{ or } k)$ for all j and k equals $Prob(j)+Prob(k)$ if outcomes j and k are mutually exclusive. These three conditions are called *Kolmogorov’s Probability Calculus*.

For many economists, *uncertain* outcomes fall into two categories: those where it is possible, because of one’s beliefs/knowledge, to associate probabilities (e.g., the probability it will rain tomorrow) and uncertain events where, because of a lack of priors, you are unable to associate probabilities with the different possible outcomes. The former are *risky* events. The later are uncertain situations where you can identify the possible outcomes, but you are unable to associate probabilities with them. E.g., whether it will rain tomorrow is associated with an estimated probability based on weather models, weather data, and painful joints, but if there were neither weather models nor data, and nothing hurt, there would be little basis for a probability estimate, and we would be clueless as to whether it will rain tomorrow.

⁷ A probability of .5 (50%) would indicate that both it and not it are equally likely. Subjective probabilities are often described as how much you would wager on a proposition. For example, consider the proposition that Biden will win reelection. If you would wager, at most, 66 cents for a bet that will pay nothing if Biden is not reelected and one dollar if he is, your subjective probability is .66 (you believe the odds are 2 to one that he will be reelected). Not everyone agrees that beliefs can have degrees. An alternative view is that you either believe something or you don’t. This alternative view would argue you either believe in God, or you don’t: “if you doubt his existence, you are not a believer.” More basically, philosophers don’t all agree on what is a belief ([Eric Schwitzgebel](#) 2019).

distinguish them from other uses of “probability” such as “frequency of occurrence”.⁸ For example, until recently, I believed, with about 80% confidence, that I had had the Covid virus. I still believe I had it, but now my confidence hovers closer to 60%. Your confidence in your beliefs—your subjective probabilities—might or might not align with probability estimates based on the best available data and best scientific practices. Looking ahead, I will define beliefs inconsistent with the best available data and scientific practices as *incorrect beliefs*.

Semantic/declarative knowledge (e.g., $2+2=4$ and Justin Trudeau is the Prime Minister of Canada) is a type of belief, a *correct belief*. In contrast, *procedural knowledge* (e.g., how to snowboard) is not a type of belief. Procedural knowledge is “knowledge-how”; declarative knowledge is “knowledge-that”.

Economists don’t often use the word “beliefs” or discuss how they are formed or updated, which is unfortunate.

Assumption 1: At every point in time, an economicus is on one, and only one, path

Economicus can’t be on two paths at the same time. N.B. (*noto bene*) this assumption does not imply that it is always on the same path. The definition of a path makes this impossible: paths include uncertainties, and as time passes, many become certainties, ending the paths with those uncertainties.

*Assumption 2: At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, An economicus has an ordering if for all paths j and k , either Path j is ranked higher than Path k , Path k is ranked higher than Path j , or paths j and k have the same rank.*⁹

⁸ What you believe is *subjective* in that it is internal to you (you are the *subject* of that belief). Consider the following beliefs: I believe it will rain tomorrow and Trump won the election; I believe in our Lord and Savior, Jesus Christ, I believe the stock market will fall, and I believe a real Burger-King Whopper must taste better than their new non-meat Whopper. These beliefs of mine, and my confidence in their truths, are subjective. The word “probability” has a number of different, but overlapping, interpretations, including *classical*, *logical*, *evidential*, *frequency*, *subjective*, and *propensity* ([Alan Hajek](#) 2019). What I am calling subjective probabilities are more formally known as *subjective Bayesian probabilities* their origin traces back to [Frank Ramsey](#) (1903-30) and [Bruno De Finetti](#) (1906-85). Their contributions to probability theory laid the foundation for expected-utility theory (discussed soon) and game theory. While Ramsey made significant contributions to economics (I had to study his economics in graduate school), economics was a sideline encouraged by John Maynard Keynes. Ramsey was an early 20th Century philosopher, logician, and pure mathematicians—not bad for someone who died at age 26. There is a recent biography (2020) by [Cheryl Misak](#), *Frank Ramsey: a sheer excess of power*.

⁹ This complete ordering is, by definition, transitive. *Transitivity* is defined as, if Path j is ranked at least as high as Path c and Path c is ranked at least as high as Path a , then Path j is ranked at least as high as Path a . Sometimes, transitivity is casually, but incorrectly, defined as if Path j is ranked strictly higher than Path c and Path c is ranked strictly higher than Path a , then Path j is ranked strictly higher than Path a . Formally, this is not transitivity but

Assumption 2 only says that if an economicus is presented with any two paths, it knows, at least operationally, which path is ranked higher. Assumption 2 does not require conscious awareness of the ordering. It does not preclude nor imply anything about the criterion to use to order paths, so it does not imply the ordering is based on WB or on the fulfillment of wants/desires, and it does not imply that it is not. Assumption 2 does not preclude orderings with lexicographic properties.¹⁰ And it neither implies nor precludes the existence of paths with the same rank.

Something caused economicus to have its current ordering, but Assumption 2 is silent on the cause. Assumption 2 excludes entities with multiple simultaneous orderings battling for dominance. It excludes orderings that depend on how the paths are *framed*.¹¹

Assumption 2 limits the orderings to paths that are imaginable.

Since humans have big imaginations, a homo economicus's set of paths is humongous, making a list of how they are ordered longer than all Google searches put end to end. This makes me wonder why an economicus would have a complete ordering of all paths given that 99.9999% aren't available to it—and never will be. A neurophysiologist might wonder where and how this humongous amount of useless information is stored. Why would evolution allocate the space? While my dog imagines fewer paths than I imagine, he still imagines many (sleep on a rug, sleep on his bed, get a drink, run straight, run left, bark once, bark twice). And since his brain is smaller than mine, I wonder where a *canis economicus* stores his ordering.

There is also the question of whether it is possible to rank only minutely different paths. E.g., *paths i* and *j* only differ due to variations in the pattern of your mother-in-law's paisley drapes and the color of her cat's tail hairs. Assumption 2 precludes economicus from saying, "I

quasi-transitivity. Transitivity implies quasi transitivity, but quasi-transitivity does not imply transitivity: quasi-transitivity is a weaker relationship between the paths. Looking ahead, relaxations of transitivity will be relevant when incomplete orderings are discussed.

Distinguish between the word *ordered* and the word *rank*: paths are *ordered*; the *rank* of a path is its "position number" in the ordering. For example, the path with only two paths ordered about it has the rank "3rd".

¹⁰ Commodity *x* lexicographically dominates *y*, if, all else the same, the path with more *x* is always ranked higher independent of the amount of *y*. For example, chocolate is lexicographically preferred to scotch, if for two paths that are identical in all other respects, the one with more chocolate is always ranked higher.

¹¹ Framing is how the path is presented; framing does not affect the components of a path. For example, patients and their loved ones are more likely to adopt a directive that says "allow natural death" than one that says "do not resuscitate" even though they mean the same thing (Venneman et al. (2006)). Orderings influenced by framing violate Assumption 2.

don't know how to order them.”—complete means complete. Assumption two is consistent with you giving them the same rank.

Assumption 2 implies nothing about how quickly an ordering can change; it only specifies that, at every point in time, economicus has only one. If you are a homo economicus, it is consistent with your ordering depending on your emotional state (mad, sad, sexually aroused), and it is consistent with your emotions fluctuating wildly. It is consistent with your ordering changing from minute to minute.

Assumption 3: An economicus's ordering of any two paths, j and k , is independent of what other paths are, or are not, imaginable.

I can imagine driving, walking, or biking to Montreal but can't imagine getting there by transmutation. If I'm an economicus, Assumption 3 says that my ordering of paths that vary in terms of bike vs. walk can't change if I come to imagine a transmutation option. Consider marriage: currently, I can imagine a path where I am married to Alex and another where I am married to Spencer, but, at the moment, I can't imagine a path where I am married to Scout. I rank the Spencer path higher than the Alex path. If I start to imagine a path where I am married to Scout, my imagining can't influence how I rank the Spencer path vs. the Alex path.

Assumption 3 might appear innocuous; it isn't. For example, imagine I have abandonment issues that cause me to rank Spencer higher than Alex because Alex is a flight risk, and I am certain Spencer will not leave me. Spencer does not dance, but Alex does, and I love to dance. Then something makes marriage to Scout imaginable, and married, we would dance, dance, dance—Scout is a great dancer. Just the image of a married life dancing with Scout causes me to hate the idea of a life of no more dancing, driving me to now rank Alex higher than Spencer. Assumption 3 precludes this.

Assumption 4: An economicus's ordering does not change in the time-span behavior is modeled.

An economicus's ordering is stable, at least for a while. Typically, the length is only specified in empirical work. For example, an economist estimates the demand functions for different fruits, because she wants to use those estimates to predict future weekly demand for bananas and strawberries. This estimation exercise requires that everyone's ordering of paths is constant during the data collection period and remains so over the period she wants to predict. If the orderings change, the estimated model predictions would be based on the wrong orderings.

Assumption 4 is rarely made explicit.¹²

Assumption 5: While there are many paths, most paths are not available, and an economicus can only experience an available path.

In explanation, many paths are blocked from the get-go, and all paths run into roadblocks. I don't have the physical talents to play in the NHL or dance with the NYC Ballet, and neither does my dog. I'm not smart enough to understand quantum mechanics. While imaginable, paths that include me dating [Carla Bruni](#) are not available. And paths that are certain about tomorrow's weather are available to none of us, but paths with weather forecasts are.

Which paths are available varies a lot, e.g., Warren Buffet can afford paths where splendid wines are drunk every night, but I cannot. This does not mean Warren will take such a path; it only means he could. On the other hand, paths that include being married to Taylor Swift or playing in the NHL aren't available to Warren, even though he has tons of money. Which paths are available to you depends on your resources and what is happening in the world. Your resources include intelligence, abilities, age, appearance, propensity to work, education, friends, and inheritance. What is happening includes prices, wage rates, who is the President of France, pollution, public goods, and whether Taylor exists and, if so, what kind of guys or women she likes.)

The passage of time reveals what has happened to now; this restricts which paths are available going forward. If I have recently gained thirty pounds and learned about global warming, every available future path must start with a heavier me who knows about global

¹² An exception is *De Gustibus Non Est Disputandum* (In matters of taste, there can be no disputes) by [George Stigler](#) and [Gary Becker](#) (1977).

The establishment of the proportion that one may usefully treat tastes as stable over times and similar among people is the central task of this essay.

If I understand, they are not arguing that the everyone has the same, and stable, ordering, but rather a lot of changes in an individual's behavior over time can be explained by changes in which paths are available, and across individuals by differences in which paths are available. In a bit of a dig,

the problem of [changing tastes] is abandoned at this point to whoever studies and explains tastes (psychologists? anthropologists? phrenologists? sociobiologists?).

I am sympathetic to the notion that constraints can explain a lot, particularly when you use the term broadly, but they can't explain a lot of why you and I experience different paths. 1977 was long ago.

warming. Even though I would like to begin the future lighter and less knowledgeable. The passage of time turns probabilities into certainties: Today, there is a 70% chance it will not rain tomorrow, so paths without rain tomorrow are still possible. But the arrival of a rainy tomorrow will make my choice today to suntan tomorrow no longer available. See Figure 2: everything is going smoothly for Scout, then it rains; it wasn't supposed to.

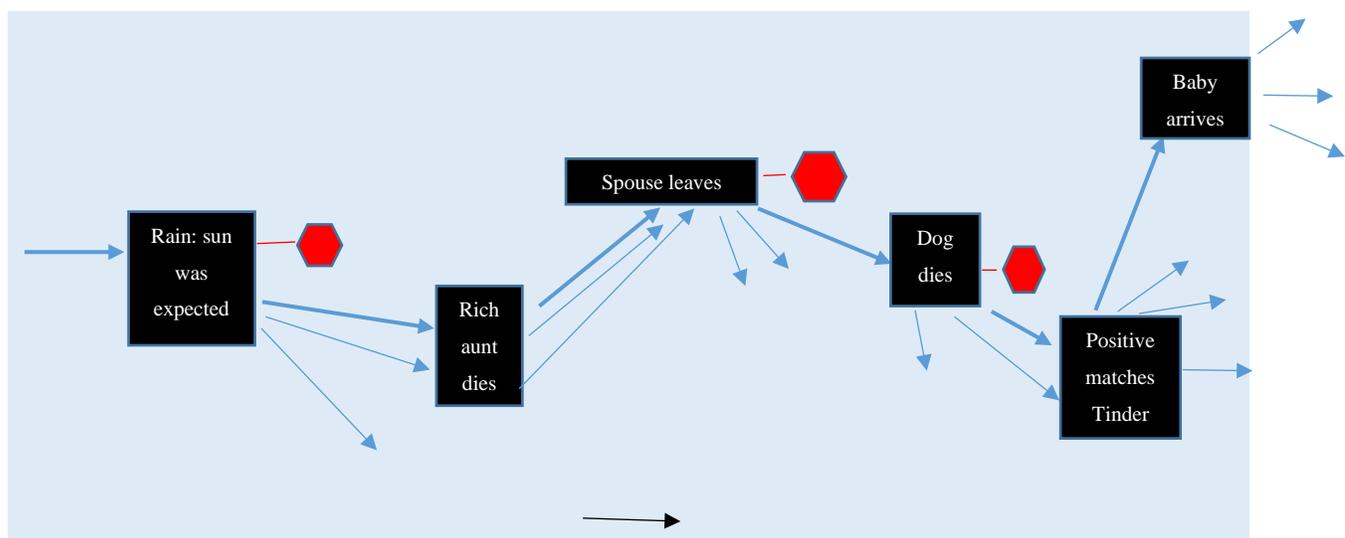


Figure 2, Chapter 1: Scout's paths over time

The black boxes are points in time where which paths are available changes. The blue arrows are the paths available at each point. The thick blue arrow is the path taken. A stop sign indicates that the current path is no longer available. Scout had planned a sunny day at the beach, but unexpected rain blocked that path, so Scout was forced to take a different path. But then, given her chosen path, her rich aunt dies, and, suddenly, she can afford more goods and services.¹³

Unplugging my toaster would drastically limit the paths available to it—limit it to one path, just sitting there.

The objective of a theory of behavior is to explain and predict how behaviors will change when what is available changes (a price or income increases, partner leaves). The objective is to explain and predict the economicus's sequence of paths as the set of available paths changes.

¹³ I drew the figure so after it rains there are three paths going forward, and her aunt dies of a broken neck in two of them. If she had chosen the lowest path, the aunt would still be alive: maybe because on that path Scout would have called an ambulance. Or maybe the other two included Scott breaking her aunt's neck. As drawn, the spouse leaves no matter which path Scott chooses.

I have defined feasibility more restrictively than do many descriptions of NCT. In those, only one constraint is explicit, the budget constraint (economicus can't purchase what it can't afford). Assumption 5 excludes paths that are not affordable in terms of income and wealth, but it also excludes paths that are not available because of ability, intelligence, cultural restraints, laws, and moral commitments.¹⁴ Other agents' orderings reduce, for you, which paths are available. E.g., paths with Bob as a friend are not available if he dislikes you. More rights for you, political and otherwise, increase your set of available paths, but they reduce those available to me: your right to do something constrains me from stopping you.

Assumption 6: An economicus's ordering of paths can't depend on which are currently available (6a), and which are currently available can't depend on their current ordering (6b).

So, eliminating a path from the set of available paths, or making additional paths available, can't affect how the other paths are ordered. Since the path currently experienced must be available, 6a implies that the ordering can't be a function of it. Assumption 6a is referred to as the *independence-of-irrelevant-alternatives* (IIA) assumption: the order of Path j relative to Path k can't depend on the availability of Path h .¹⁵ 6b says that what is available is independent of the ordering.¹⁶

The largest set in Figure 3 is all possible orderings of all imaginable paths: if an unimaginable path becomes imaginable, this set enlarges. Assum:3 is the subset unaffected by adding or subtracting imaginable paths. Assum:6a is the subset unaffected by adding or

¹⁴ Neoclassical choice theory does not explicitly mention duties (self-imposed constraints), commitments, or responsibilities. Hausman (2012), an exception, views moral commitments as important determinants of behavior, but treats them as a determinant of the ordering rather than as a constraint.

¹⁵ A different name for Assumption 6a is *choice coherence* (Kreps 2013). The *weak axiom of revealed preference*, WARP, accomplishes much the same thing as IIA but is a restriction directly on behavior, whereas IIA is a restriction on the ordering of paths (Mas-Colell, Whinston, and Green (1995 or 2012). If 6a is violated the individual's choices can be, in Kreps' judgement, "incoherent". His example: you order coffee and pie, you are given a choice between apple and cherry, you choose apple. The waiter then says, "Oh, we also have peach," and this added alternative makes your change your order to coffee and cherry pie. That is, how you order two paths changes when another path becomes available. Assumption 6a takes center stage in Chapter 7 where I discuss the comment quirk called the *endowment effect*.

¹⁶ Assumptions 2, 3, 6, plus the assumption that the set of paths is finite (an assumption I haven't made) imply that the ordering of paths can be represented by a *utility function* that is unique up to increasing monotonic transformations. A utility function is a mathematical function that associates a number with each path as function of its components such that a higher-ranked path is associated with larger number. I will rarely mention utility functions or the number the function associates with each path (called *utility*).

subtracting available paths. [Changing what is imaginable does not necessarily change what is available (e.g., newly imagining transmutation to Montreal does not make it available).

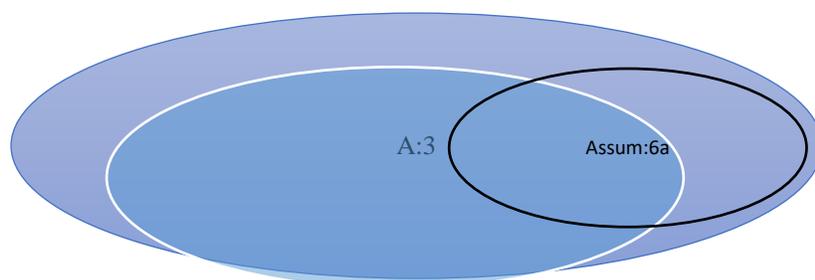


Figure 3, Chapter 1: orderings, at a pt. in time consistent with Assum. 3 and 6a

And changing what is available does not necessarily change what is imaginable (e.g., a train trip to Montreal becoming affordable does not imply it wasn't previously imaginable).] The intersection of Assum:6a and Assum:3 represents the set of orderings that are not influenced by changes in imaginable or available paths.¹⁷ NBT restricts orderings to the intersection of the A:3 and Assum:6a sets.

*Assumption 7: At every point in time, economicus takes one of their highest-ranked available paths (HRAPs).*¹⁸

Assumption 7 is an *optimizing axiom*. Optimization mistakes are not allowed.

Even if the ordering of paths is constant, what is available changes because time continuously updates our knowledge about the world and its future. This, by itself, will cause the HRAP to change.

My definition of a path along with these seven assumptions is a theory of behavior. If Assumptions 1-7 hold for Lady Gaga, and you know her ordering of paths and which paths are available, you can immediately identify which path she will now take if her HRAP is unique.

¹⁷ That part of Assum:6a that isn't part of Assum:3 is the set of orderings that are unaffected by what is available but affected by what is imaginable. That part of Assum:3 that isn't part of Assum:6a is the set of orderings that are unaffected by what is imaginable but affected by what is available.

¹⁸ Implicit is the restriction that there exists a HRAP (or paths). Assumption 7 is typically stated as a single highest-ranked available path, but there could be a number of available paths that are all tied for the highest rank.

The taken path will vary across economici because what is imaginable differs, orderings differ, and what is available differs. Assumptions 1-7 is a behavioral theory, but there isn't anything in these seven assumptions that implies that an economicus makes conscious choices, consciously debates the alternatives, or that the debating influences what they experience.

Assumption 5 (economicus can only be on an available path), along with 1-4 and 6-7, leads to the central prediction of NBT: demand curves usually slope down (price increases cause demand to decrease).

A weakness, or strength, of Assumptions 1-7 as a theory of behavior is it assumes nothing about the ordering criterion, making Assumption 1-7 consistent with every possible ordering criterion. They are, for example, consistent with a randomly-generated ordering or an ordering based on how many times the letter *q* appears in the description of each path.¹⁹ With only Assumptions 1-7, the brain of an economicus is a **B.F. Skinner** black box, so there is no need to consider psychology or cognitive neurology.²⁰ The downside, and upside, is that the ordering "is what it is". Assumptions 1-7 don't preclude nor imply that your chosen path is, in any sense, *best* (better for you than any other available paths). So, economists add a few additional assumptions.

While I have not seen the word *senses* or *sensations* in a description of NCT, I'm pretty sure economists presume economici have senses, so let's make that explicit.

Assumption 8a: Economici have senses: they are sentient. And the sensations experienced and their intensities depend on which path is taken.

Humans have six senses: sight, hearing, touch, taste, smell, and proprioception.²¹ Bats have echolocation; humans don't. A sensation is the awareness of a stimulus through one or more

¹⁹ While Assumption 1-7 are consistent with one's ordering being randomly generated, it precludes higgledy-piggledy behavior: Assumption 4 restricts behavior to what is feasible. you might find Assumption 1-7 with each entity's ordering randomly generated at each choice occasion, preferable to NCT. In explanation, everyone agrees that behaviors are limited to feasible behaviors, and what is feasible at a choice occasion is largely determined by prior behavior.

²⁰ In the 20th Century's first half, economists worked hard to purge choice theory of psychology. At the same time, psychology was dominated by behaviorism (led by Skinner); it considered the brain a box not worth opening—only behavior matters. While psychology has left behaviorism behind, many economists still have a principled blindness when it comes to the brain. Neoclassical choice theory predates modern neuroscience and behavioral economics.

²¹ Animals have nerve receptors that fire only when a specific stimuli is encountered. For example, you have taste receptors on your tongue that only fire in the presence of sweets, and others that only fire in the presence of sour,

senses; they enable an economicus to observe its environment (outside stimuli), so the ability to change its behavior in response to changes in its environment (respond to environmental stimuli). For example, head for the shade when the sun gets too hot and run out the back door when you hear a wolf howling at the front door. My refrigerator senses its internal temperature and set temperature but is sightless; in contrast, fridges in development have cameras that can sense whether it contains milk. My toaster senses only which button I pushed (*bread, bagel, or defrost*) and whether its lever was pushed down. It sends electricity to its heating coils when it senses that its lever has been pushed down.

Eating an ice cream cone produces a different melody of sensations than getting a massage. Animals experience pain, a sensation, and try to avoid the behavior that caused it.²² Economicus is motivated to repeat the behavior that caused a positive sensation. If the theory is Assumptions 1-9a, animals are not excluded.

But what about plants? [Natalie Angier](#) (2009) describes how plants sense their surroundings and react in extraordinary ways. Charles Darwin and his son wrote a book (Darwin 1880) about how plants respond to stimuli. Everyone agrees that plants are sentient—they gather information about their environment and change their behavior based on what they sense. For example, they move towards sources that enhance their WB (light, nutrients, water) and put up chemical defenses when their WB is attacked by pests. They even share resources with other plants depending on their relative needs. The neuro-botanist Stefano Mancuso (Bremmer et al. 2006, and Mancuso and Viola 2015) further claims that, while plants do not have a central brain, they do process information and their different parts communicate (signal), sharing information in ways that are analogous to how animal brains work: they have cells that behave like neurons. He does not claim that plants experience pleasure or pain, but it makes one wonder. Many

and you have receptors on your retina that only fire if they experience the color red. Such firings are transmitted to you brain. These are sensations. Physiology and experiments indicate that there is more similarity in a sensation within a species than across species; for example, the sensation of seeing is more similar across humans than between humans and frogs.

²² An animal does not even need a backbone to experience pain. After using pain-inducing chemicals on spiders, [Thomas Eisner](#) (2003, p253), the recently deceased “father of chemical ecology”, concluded, “I have no doubt they did [experience pain]...we came to the conclusion that invertebrates perceive pain, and that their sensory basis for doing so may not be much different from our own.”

botanists reject Mancuso’s claim that plants are intelligent; it comes down to how you define intelligence.²³ He defines it as the ability to solve problems.

What about robots? My Roomba vacuum cleaner senses walls and drop-offs and changes its behavior (directions) when it encounters walls and drop-offs. It can also identify objects using its camera and artificial intelligence. The ability to sense things—including pain and pleasure—requires neither self-awareness nor a mental map of what is happening. Assumption 8a neither implies nor precludes economicus from having thoughts and emotions.

Assumption 8a implies an entity without sensors isn’t an economicus. But all plants and animals have sensors, so Assumption 8 does not exclude them. It also does not exclude toasters, refrigerators, or computer programs. It does exclude inanimate objects like rocks and steel bars.

I believe that advocates of NCT believe/suppose economici are sentient, but Assumption 8a is not listed as an assumption.

Assumptions 1-8a imply that a sequence of sensations is associated with each path.

Assumption 8b: An economicus’s ordering of paths is determined, at least in part, by the sequence of sensations that would be produced by each path.

That is, whether Path j is ranked higher than Path k is determined, at least in part, by the sensations that would be produced by each.

Assumption 9a: An economicus’s ordering of paths is based on its WB (well-being), WB, accounting for the uncertainties. The more WB economicus associates with a path, the higher its rank. And WB is determined, in part, by sensations.

The phrase “accounting for uncertainty” is to be interpreted broadly, allowing different economici to account for uncertainty differently.²⁴ Assumptions 1-9a don’t imply an economicus will always take the available path that would have produced the most WB. Going forward, for brevity, I will omit the qualifier “accounting for the uncertainties”. Note that 9a neither precludes

²³ Mancuso was recently physically attacked by vegetarian extremists (Scorranese 2021): their reasoning was that saying plants have feelings makes it more OK to eat meat. He rejects their reasoning, but there is a logic to their conclusion.

²⁴ I initially wrote, “in a statistical sense.” But for many econometricians “statistics” implies the use of probability measures. Assumption 9a neither requires nor precludes probability measures.

nor implies that the economicus orders paths by their expected level of WB—*expected* meaning *statistical expected value*—details later.

Or

Assumption 9b: An economicus's ordering of paths is based on its wants and desires. Paths that are more wanted/desired/are ranked higher. And wants and desires are determined, in part, by sensations.

Or

Both 9a and b: the two orderings are identical.

NBT (neoclassical behavior theory) is Assumptions 1-9, and any entity whose behaviors and motivations are consistent with Assumption 1-9 is an economicus. The choice theory taught in introductory microeconomics courses is a restrictive case of NBT.²⁵ What is known as *Expected-utility theory* is as well—details soon. After discussing Assumptions 9a and b and then the implications of Assumptions 1-9, I will consider additional assumptions you might want to add.

The distinction between 9a and 9b is critical: it is the distinction between an economicus getting what it wants and getting what will increase its WB. Imagine three possible paths, j , k , and l , such that your WB-ranking is k , l , and j , and your wanting-ranking is k , j , and l . The two orderings differ, but Path k is ranked higher than l in both. Picture the degree of overlap between two paths as a continuum of correlations. If 9a and b both hold, the correlation is 1.0; the other end is -1.0; the midpoint, 0, being completely uncorrelated. My guess is that many believe that the WB-ranking and the want/desire-ranking are either identical or highly positively correlated. In contrast, Buddhism believes the default correlation is negative and defines enlightened as a state of being where 9a and b merge. Wow! economicus is enlightened.²⁶ The amount of actual overlap is an empirical question. A later chapter assesses the research.

²⁵Typically, in introductory microeconomics courses, a path consists of only the goods and services you would personally consume on that path—your *consumption bundle*: other important path components are ignored. Which consumption bundles are available is only restricted by prices and income—the *budget constraint*. There is no future (it's a one-period world); there are no risks, and you have full knowledge—there is no ignorance.

²⁶ Buddhists talk about “craving” rather than wanting and desiring, a synonym with a negative edge, “craving” chosen to draw their critical distinction between want fulfillment and WB (see, for example, [Stephen Batchelor 2015](#)).

At this point, WB²⁷ and *wants* and *desires* are not explicitly defined, but they are determined by sensations, at least in part.²⁸ Being more explicit will require specifying economicus's experiential capabilities. Dogs, humans, and refrigerators have different experiential capabilities; for example, while they all have temperature receptors, dogs and humans, but not refrigerators, have smell receptors, and dogs' receptors are more sensitive. Humans can think abstractly and can self-reflect; dogs and refrigerators can't. So, the specifics of WB are likely different for dogs, humans, and refrigerators.

Assumption 9 allows for two broad but distinct ordering categories: a WB criterion and a want/desire fulfillment criterion. So, without further assumptions the ordering criteria is still unclear.

A question. Does making the ordering criterion either WB or wants and desires (Assumption 9) further restrict what entities can be economici? Or, said another way, can entities that are neither animals nor plants have WB or have wants and desires? Sure! For example, when broken, an entity exhibits less WB (being well). WB for a machine or computer program is increasing in how well and long it performs its dictated tasks and at what cost. So, performance reflects the user's WB. [I wish my word processor, WORD, exhibited more WB, particularly the part that "corrects" my grammar.] I defined WB as a combination of emotional WB, life-satisfaction WB, and pleasurable sensations, but not all entities are capable of experiencing all three. Machines, for example, are incapable of emotional WB and pleasurable sensations. And most species can't experience life satisfaction. Experiential capabilities vary with DNA by species and across individuals within a species. Keep in mind that animals, like machines, do not get to pick their capabilities.

Assumption 9a implies one concept of *value*, *economic value*, in that higher-ordered paths have more value. Since orderings depend only on economicus's WB, objects don't have

²⁷ What *WB* means is often ignored. One view is that it does not matter what it means: it is "in the eyes of the beholder", like *porn*. I reject this view. Another view is that WB means more *utility*, but this simply defines WB as an ill-defined word, *utility*, a lousy synonym for WB. Many of my students think utility means more than it does, and while some surveyors ask people about their WB, no one asks them about their utility. See **Footnote 12 in Chapter x.**

²⁸ Looking ahead, *wants* and *desires* are synonyms. For some economici, *drives* and *instincts* might be more appropriate words.

objective/inherent value, only extrinsic value. Without 9, there would be no such thing as economic value.²⁹

Assumption 9a also implies that expanding the set of available paths can't decrease WB (having additional options never hurts—even a million colors of house paint).

With either 9a or 9b, the ordering need not depend on every path component. While who is the Prince of Denmark is a component of every path, it does not have to affect an economicus's ordering: if three paths are identical except for whether it's Hans, Bob, or Hamlet, and the economicus is indifferent to who is the Prince, the three paths would all have the same rank. Your ordering might depend on how much food people in China consume in each path, while mine depends only on what I eat: I'm indifferent to what the wife, kid, and dog eat.

There are many reasons why an animal's ordering could depend on what other animals are doing. These include: (1) what another animal is doing negatively affects the animal (it is smoking, driving drunk, or trying to kill and eat you), (3) what another animal is doing positively affects the animal (it's protecting, feeding, helping), (4) animals care about their position (status, influence) relative to other animals, particularly their peers, and (4) animals of the same species sexually desire one another.

For humans, their ordering depends on each path's present and future components. And many humans put less weight on path outcomes the farther they occur in the future—they discount the future. But there is nothing in Assumptions 1-8 that implies economici have the cognitive ability to consider the future—and there is also nothing that precludes it.

Knowledge, or its absence (ignorance), could be a factor in how an economicus orders paths. All else constant, I might rank higher paths where I don't know the black spot on my toe is toe cancer, even if I can find out for free. Knowledge and ignorance are commodities, just like peanut butter.³⁰ My ordering could also depend on freedom or rights (Free speech? Free Love?) I would have in each path. And it could depend on whether you have those same rights.

²⁹ So, for you, apples, the smell of roses, and the existence of a pristine ecosystem only have value if including them in a path would increase that path's rank. Values depend on the orderings of paths, so commodities only have value if they affect the ordering; value, or lack of, isn't an objective characteristic of a commodity.

³⁰ In a recent study, [Ganguly and Tasoff \(2014\)](#) found that 16% of the university students sampled were willing to forego a payment of \$10 to avoid learning whether they had genital herpes. [David Eli](#) and [Justin Rao \(2011\)](#) have

At this point, you might wonder why an entity could have an ordering not based on either WB or want fulfillment (Assumption 9a or b). A straightforward answer is that entities neither select their ordering nor their ordering criterion. Rather, your ordering is imposed by genetics, God, parents' parenting skills, and other past experiences, which you can now do nothing about. In evolutionary genetics, the objective of plants and animals is to increase the probability they will pass along their DNA, which is different from making them better off. What ordering a God would impose depends on which god, but the one in my church was not into WB or reducing suffering while people are alive. He was into maximizing ill-being for most dead people—those in hell.

NBT assumes economicus's ordering is given and assumes nothing about its creation. Economists don't seem to care where it came from—but should. Quoting [Franz Dietrich](#) and [hi List](#), and then [Samuel Bowles](#):

Preferences are exogenously given in that the standard theory can't explain how they are formed. An agent's preferences are simply taken to be an essential but inexplicable feature of the agent's personal identity.

The axiom of exogenous preferences is as old as liberal political philosophy itself... Reflecting this canon, most economists have not asked how we come to want and value the things we do.

If there is no theory of how a WB-ordering is created, any is possible, and my WB could mean sticking pins in my eyes and drinking antifreeze. If I am an economicus, eyeball pin-sticker, it is either because I have a bizarre ordering or terrible options. If economists could explain the ordering's creation and evolution, they would be better able to predict behavior. Your ordering is likely determined by genetics and learning, but both words are absent from economic theories of choice.³¹

found that people will pay to not learn about their IQ or how attractive others find them. See also [Sweeny](#) et al. (2010)

³¹ There is a small literature on the formation of your ordering, not much by economists. In one strain (going back to [Kelvin Lancaster](#) (1966)), an ordering of goods and services is derivative of a more basic ordering in terms of the characteristics of those goods and services, but it does not ask where the ordering in terms of characteristics came from or how and why it might change over time. In the ordering-formation literature a distinction is made between worlds of complete and incomplete information. In a world of incomplete information, experiencing different commodities and activities provides information (experiencing what Vegemite tastes like) that might cause you to modify your ordering of paths. Also, belief formation, and the evolution of beliefs affects your orderings. Dietrich and List (2012) model your ordering in terms of the motivational salient properties of each path (the properties you pay attention to). In which case, your ordering will change over time as properties gain or lose salience. "For example, after having suffered from gallstones, an agent may form his or her preferences over various kinds of food on the basis of whether they are healthy and may no longer be interested in whether their taste is rich (something he

*WB-comparable and WB-commensurable*ⁱⁱⁱ

Assumptions 2 and 9a require that all paths and their components are *WB-comparable*, which requires that all the kinds of WB are *WB-commensurable*. If there isn't complete WB-comparability and commensurability, the ordering is incomplete, contradicting Assumption 2.³²

Explaining: distinguish between kinds of WB and *bearers-of-WB*. Feeling safe is a kind of WB, whereas door locks are a bearer (producer) of this kind of WB, and so are security services.³³ A bearer is a producer of one or more kinds of WB. A path is a bearer, and so are each of its components. Since, for me, the taste sensation sweet-'n-salty is one kind of WB, bacon covered in maple syrup, and caramels with sea salt are bearers of this kind of WB; all else constant, I order paths with more sweet-'n-salty higher.

An ice-cream cone is a bearer of multiple kinds of WB, including pleasurable taste, gustatory and visual sensations, positive thoughts, and positive emotions (*happiness*). But, at the same time, its consumption can decrease other kinds of WB: my later thought: *I'm depressed because my diet is ruined*.

Distinguish between comparing kinds of WB (a sense of calm vs. sexual pleasure) and comparing bearers (a meditation class vs. a willing sexual partner).

WB-comparability is a property that two paths either have or don't have. If two bearers aren't WB-comparable, they are *WB-incomparable* (Chang 1997). Paths *h* and *k* are WB-

or she cared about before the illness)." They do not model why salience differs across individuals at a point in time, or over time for a given individual.

In the 1998 *J. of Economic Literature*, Samuel Bowles reviews how economic institutions (markets, families) influence your ordering. I have modeled (Morey and Thiene 2017) one's ordering as a function of one's personality. See also the references to the determinants of the ordering in Dietrich and List (2012), including the articles authored by Fenrong Liu and the book, *Preference Change: Approaches from Philosophy, Economics and Psychology*. For another take, see Hausman (2012). He views the determination of one's ordering as a challenging process of weighing WB, desires, moral obligations, and beliefs, so subject to change. Philosophers discuss the logic (or illogic) of changing one's ordering.

³² It is a sufficient, but not necessary, to make the ordering incomplete. Sen (1997), for example lists "(1) *process significance* (preferences may be sensitive to the choice process... and (2) *decisional inescapability* (choices have to be made whether or not the judgemental process has been completed)." Sen (2018) lists *epistemic inadequacy*: "we quite often do not know all of the consequences that will follow from taking one action rather than another."—another term for lack of full information.

³³ For some, guns are a bearer of feeling safe, for others, they are bearers of fear. Political philosophers have used *freedom*, *equality*, and *pleasure* as examples of kinds of WB and the Bill of Rights as a bearer of freedom and equality. Philosophers concerned with the motivators of choice typically use the word *value* rather than WB. [Ruth Chang](#) (1997) is responsible for the adjective *bearers* to keep clear the distinction between kinds of WB and the bearers that produce the different kinds of WB.

incomparable if **none** of the following statements are correct: h is ranked higher than k , k is ranked higher than h , or h and k have the same rank. They are WB-comparable if one, and only one, of these statements, is correct. For example, if Path k is identical to Path h , except Path k has x more pandas in the wild but y has fewer teachers in your daughter's school, and if you can determine which path you rank higher, these two paths are WB-comparable. [WB-(in)comparability is a pair-wise property:³⁴]

WB-comparability is a property that two path components have or don't have. Components x and y are pair-wise WB-incompatible if increasing the amount of one while decreasing the amount of the other makes the new path WB-incomparable with the initial path. For example, skiing, s , and eating chocolate, c , are not WB-comparable if increasing s or c while decreasing the other makes this new path WB-incomparable with the initial one. In contrast, a few examples of comparability: I would give up my friendship with George for two apples, my friendship with Phil for five apples, and for no finite number of apples would I give up my friendship with Wanda. I would give up my right to free speech for the freedom to do whatever else I please, but I would not give up free speech for 12 ice cream cones. A complete ordering of paths based on WB means you know whether and at what rate you would trade any path component for any other path component, be it friends, The Bill of Rights, or cheeseburgers. Keep in mind that the rate can be zero or infinite—extreme examples of WB-comparable.

Figure 4 shows what paths are ranked or not ranked when there are WB-incompatibilities

³⁴ If Paths j and k are incomparable, and Paths k and l are incomparable, that does not imply j and l are incomparable.

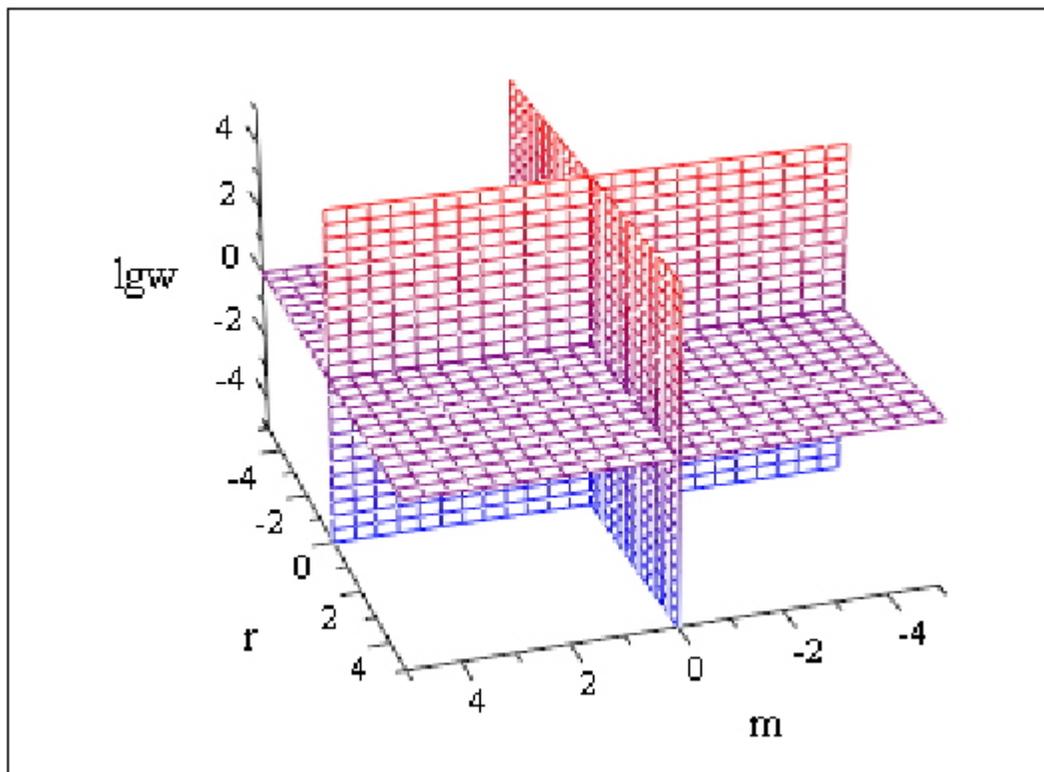


Figure 4, Chapter 1: m - r - lgw paths

Since it's a figure, paths are limited to three components. Imagine one is the rate of global warming, gw , the other two the individual's consumption of market goods, m , and her positive personal relationships, r (familial, sexual, and platonic (friendships)). In Figure 4: m increases as one moves left, r increases as one moves forward, and the rate of gw declines as one moves up, lgw for less gw . The intersection of the three planes can represent any chosen path. The axes are expressed as deviations from this path. Paths with more of at least one of the three but none with less are in the upper-left front quadrant, Quadrant A. Paths with less of at least one, but not more of any, are in the lower-back right quadrant, Quadrant G—label the quadrants alphabetically, A–H, going counter-clockwise, starting with A top left. Complete WB-comparability means the individual can rank every conceivable pair of paths.

But what if some of the path components are WB-incomparable? Paths in Quadrant A are still ranked higher than the origin path and paths in G lower. Which other rankings remain depends on what is and isn't comparable.

If nothing is comparable, only the Quadrants A and G paths are ranked relative to the origin path.

What if m and r are WB-comparable, but gw is WB-incomparable with both? Then, paths on the horizontal plan through the origin (gw constant) are also ranked related to the origin path. Paths in Quadrants B, C, and D interiors are unranked relative to the origin path. These paths have less gw and less of either m or r than the origin path. Paths in Quadrants E, F, and H interiors are also unranked relative to the origin path. These paths have more gw and more m or r than the origin path.

Alternatively, what if gw and m are WB-comparable, but r is incomparable with both? Then, all paths on the vertical plan through the origin (r constant) are ranked related to the origin path. Paths in Quadrants B, E, and F interiors are unranked relative to the origin path. These paths have more r and either less m or more gw than the origin path. Paths in Quadrants C, D, and H interiors are also unranked relative to the origin path. These paths have less r and more m or less gw than the origin path.

To summarize, if a path component is not WB comparable with other components, paths with different levels of that component often can't be ranked.

WB-incommensurability: WB of kinds A and B are WB-incommensurable if you cannot compare them in terms of WB ([Wiggins](#) 1997 and [Hsieh](#) 2016). For example, it requires you to be able to compare music's sensual pleasure with the relief from finding out pandas won't be going extinct. Complete commensurability means you have a complete ordering over all conceivable *packets* of different kinds of WB. Incomplete commensurability means you can't order all the WB-packets.³⁵ In philosophy, the practice is to use the words "comparable" and "incomparable"

³⁵ You will still have a partial ordering of packets. For example, if all the kinds of WB generated by eating different types and quantities of chocolate are commensurable, there will an WB ordering over packets that vary only in the types and quantities of chocolate eaten. There can also be a partial ordering in that Packet c generates more WB than Packet m , while there are packets that generate more WB than c but less than m but can't be ranked relative to each other.

when comparing goods, activities, and paths, and to use the words “commensurable” and “incommensurable” when comparing kinds of WB (Chang 1997).³⁶

Feeling safe and being in love are WB-incommensurable if you cannot order all packets that vary only in feeling safe and experiencing romantic love. Likewise, if you can’t compare the grief from the loss of a loved one with global-warming existential-angst, they are WB-incommensurable.³⁷ [Right now, the pandemic is causing both death and a reduction in CO₂ emissions.]

Complete WB-commensurability implies an economicus can order all packets in terms of **utility**, with packets differing by kinds and magnitudes of WB. E.g., picture a three-dimensional graph with *satisfaction*, *calm*, and *pleasurable sensations* on the three axes. Each point in Figure 5 represents a different packet. The black WB package (dot) has positive amounts of satisfaction and calm but no pleasure, whereas the green has pleasure but no satisfaction nor calm. If an individual can’t order every packet, she suffers from WB-incommensurability. Alternatively, if I could always order paths that vary only in the amounts of contentment and excitement they produce, contentment and excitement are, for me, WB-commensurable.

³⁶ I add the prefix “WB-“ to both “comparable” and “commensurable” because one can compare based on many criteria besides WB, so comparability does not always mean *WB-comparability*. And because a few economists have used “commensurable” to mean the ordering of paths is complete and has no lexicographic properties, which is different from how I define *WB-commensurable*. See [Clive Spash](#) and [Nick Hanley](#) (1995), [David Pearce](#) 2000, [Mika Rekola](#) 2003, and [Aldred 2002?](#), 06 and 13).

³⁷ Even if Packets *h* and *k* are WB-incomparable, they still can be comparable in terms of specific kinds of WB (e.g., Packet *k* generates more happiness than *h*, but less life satisfaction). And Packets *h* and *k* being comparable does not imply that either is comparable with any other packets.

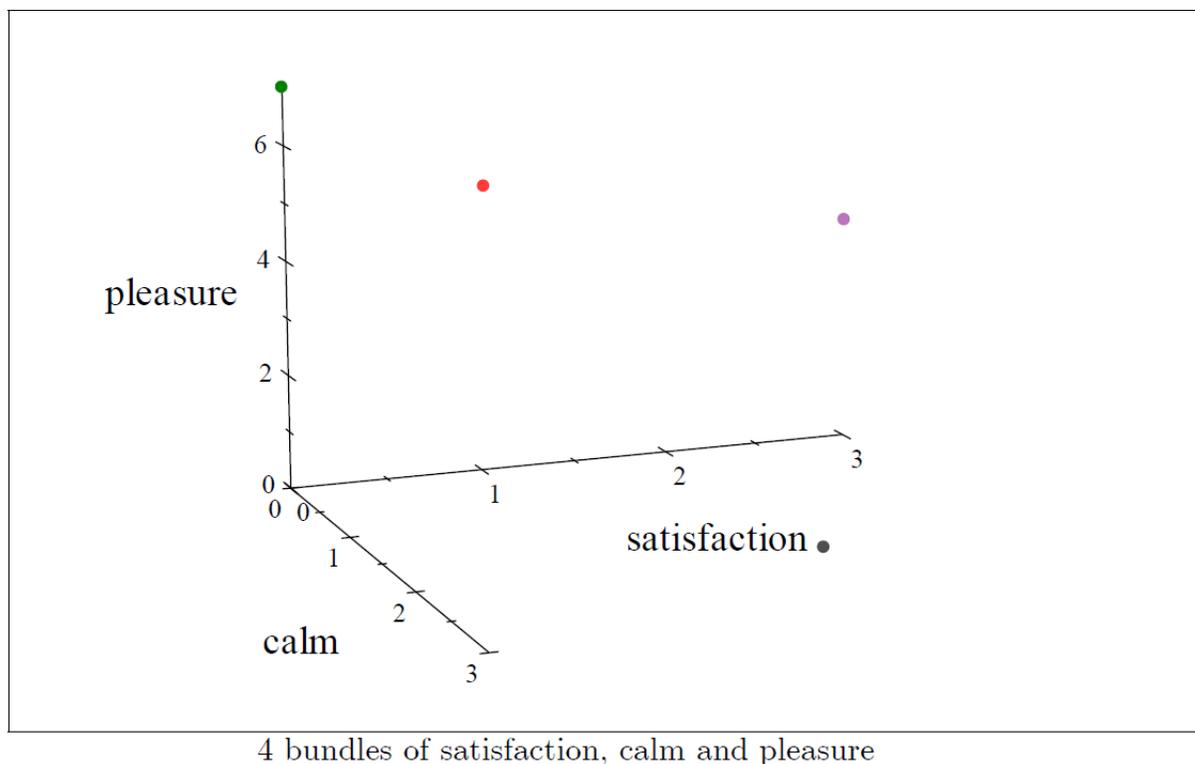


Figure 5, Chapter 1: four combinations of satisfaction, calm, and pleasure.

WB-commensurability gives a second and equivalent way of defining pair-wise WB-incompatibility for two components. Commodities s and c are WB-incomparable if the kinds of WB produced by s are incommensurable with the kinds produced by c .³⁸

Complete WB-commensurability is necessary for complete WB-comparability.

You can guarantee that WB-incommensurabilities will not occur if you are willing to assume there is only one type of WB: if so, whether two kinds of WB are commensurable is a non-issue. But, many people, including economists and philosophers, are *WB pluralists* (a philosopher would say *value pluralists*): we believe there is more than one kind of WB, more than one way to be better off. [Daniel Kahneman](#) and [Angus Deaton](#), two recent Nobel laureates in Economics, are WB pluralists. The Harvard legal scholar [Cass Sunstein](#) (1994) makes a

³⁸ Incomparability of a bearer can be discontinuous: at some levels, you can WB-compare and at other levels you can't. For example, while I'm not sure I could compare all possible income reductions with all possible reductions in the rate-of-global-warming, I know I could compare a few of them: for example, ten dollars less income with a 10% less gw is ranked higher than my current path. All that is needed for WB-incomparability is that there are income reductions where I wouldn't be able to order the current path vs. a 10% reduction in gw combined with any of those income reductions.

compelling case for WB pluralism. The utilitarian Jeremy Bentham was not a WB pluralist: he was a WB monist; the only WB is pleasure (a unidimensional pleasure/pain continuum), a sensation that varies only by magnitude and duration.³⁹ The philosopher and legal scholar [Donald Regan](#) is, I believe, a rare example of a modern WB monist.⁴⁰

Aristotle, a WB pluralist, argued that many kinds of WB are WB-incommensurable; many modern philosophers agree ([Martha Nussbaum](#) 2012)—many regular people agree.⁴¹ If not all kinds of WB are commensurable, Assumptions 2 and 9a are violated: economicus does not have a complete ordering of paths based on WB. Chapter 8 provides more detail and summarizes the arguments for and against bearer-of-WB comparability and WB-commensurability.

In contrast, if one adopts Assumptions 9b instead of 9a, there will not be a complete ordering unless there is complete comparability in terms of bearer-of-want/desire fulfillment. And this requires complete commensurability of all the different wants and desires.

Cardinal WB?

Assumptions 2 and 9a don't imply that WB has cardinal properties; they also don't preclude it. If an individual's WB is *weakly cardinal*, she can order paths in terms of WB; and she can, in addition, order differences between paths in terms of WB. She can comprehend whether a shift from Path *j* to Path *k* would increase her WB more or less than a shift from Path *c* to Path *f*. If her WB is *strongly cardinal*, she can also comprehend whether the ratio of the WB's associated with

³⁹ He was preceded in that belief by John Locke (Locke, Book II: Chapter XX; Patricia Sheridan 2016, 19). Whether the utilitarian J.S. Mill was a WB pluralist is a matter of debate. The philosopher Martha Nussbaum identifies him as one because he counts “music, virtue, and health as major pleasures” (Nussbaum 2012, 338). In contrast, Mark Schroeder in his survey of Value Theory (2012) notes that “whether Mill properly counts as a pluralist about value depends on whether his view was that there is only one value — happiness — but two different kinds of pleasure which contribute to it, one more effectively than the other, or whether his view was that each kind of pleasure is a distinctive value,”

⁴⁰ Quoting him (1997, 129): “In this volume I am the ‘designated eccentric’, appointed to take a position no one else would touch with a barge pole,...Specifically, I believe the following two propositions: (1) There is one and only one sort of value that matters to practical reason in the final analysis. This unique final value is G.E. Moore’s ‘good’ [‘good’ as in approved of]. (2) Given any two items (objects, experiences, states of affairs, whatever) sufficiently well specified so that it is apposite to inquire into their (intrinsic) value in the Moorean sense, then either one is better than the other, or the two are precisely equal in value.”

⁴¹ Philosophers accept the idea that some kinds of WB are commensurable (maybe the pleasure of a cold drink and the pleasure of an ice-cream cone), but they reject that all kinds of WB can be compared. For example, the WB loss generated by environmental injuries and animal extinctions might not be comparable with the WB gain generated by a cold soda.

Paths j and k is greater than the ratio of the WB's associated with Paths c and f . For example, an individual whose WB-ordering is strongly cardinal could comprehend that a shift from j to k would increase her WB three times more than a shift from c to f .⁴²

Complete WB-comparability and commensurability don't imply that WB has cardinal properties.

Implications of NBT (Assumptions 1-9)

Bivalves and plants order paths consistent with Assumptions 1-9. I am not sure about dogs.

Consider a computer program that orders words alphabetically and selects the highest-ranked available word. But words that start with the letters a , c , f , or ba are unavailable. It would select bb , the "size of a shot fired by an air rifle". The program selects its most wanted path, the bb path. Consider my refrigerator: the relevant components of every path are internal temperature, set temperature (set by me), and whether the refrigeration is on or off. The paths are ordered by how much it is too hot or cold. Specifically, the smaller the absolute difference between the internal and set temperatures, the higher the path is ranked. If there are two paths where the temperature is 5 degrees too warm, the one with refrigeration on is ranked higher than the one with it off. If the internal temperature is either at the set level or too cold, it experiences/chooses refrigeration off; if it is too hot, it experiences/chooses refrigeration on. It is doing the best it can; it is being well. Life and consciousness aren't required. If Assumptions 1-9 apply to humans, they also apply to philosophical zombies, creatures that, by definition, have no conscious experiences.

Assumptions 1-9 do not imply that *economici* have thoughts or emotions. These assumptions do not require a memory except for knowing the ordering. And they don't imply that *economici* are self-aware. They don't even imply that *economici* are conscious. Assumptions 1-9 do not require that an *economicus* is a living plant, animal, or human. Assumptions 1-9 only

⁴² Consider numbers. Numbers are ordered in that 10 is larger than 5 and 5 is larger than 3. In addition, numbers are weakly cardinal in that the differences between numbers can be ordered. E.g., the difference between 10 and 5 is greater than the difference between 6 and 3. An ordering where the differences can be ordered is, at least, *weakly cardinal*. Numbers also have the property that the ratios of numbers can be ordered. For example, $10/5$ is larger than $5/3$. If an ordering has this ratio property it is *strong cardinality*. Strong cardinality implies weak cardinality. In contrast, consider the ordering properties of letters (a , b , c , ...). Letters are ordered in that "c" comes before "m" in the alphabet but a difference between any two letters is meaningless, so differences between letters can't be ordered, neither can their ratios. So, letters have an ordering, but that ordering has no cardinal properties: no one claims the letter q divided by the letter m is greater than the letter d divided by the letter w .

imply rationality in the weakest sense. And Assumptions 1-9 don't imply that WB has cardinal properties.

Assumptions 1-9 don't imply that the path experienced was chosen (consciously or unconsciously), which makes me wonder why NCT includes the adjective *choice*. Here I am defining choice in the street sense: you have a choice if you could have behaved differently. Assumptions 1-9 apply to worms, but many humans believe that worms are *incapable of choice*: worms have limited options, and evolution has programmed their instincts to crawl, dig and eat to maximize gene transmission.⁴³ Does incapable of choice apply to you as well? Maybe? This is discussed in Part II.

Assumptions 1-9 imply that there is no randomness in behavior. By this, I mean that if two economici at a point in time have the same ordering of paths and the same available paths, both will take the same path.

Assumptions 1-9 are consistent with taking a path with future doom and gloom written all over it. For example, an economicus might drink a bottle of gin every day—it tastes great and, at the time, makes it feel great—even though it knows it will soon lead to a massive drop in future WB. Such drinking could be part of economicus's best available path if the weight it puts today on future WB is less than it puts on immediate WB. As noted earlier, Assumption 9a is consistent with any type of, and degree of, what economists call *temporal discounting* (discounting the future).

If you subscribe to an ethic that assumes more WB is the goal, then whether Assumption 9a holds (the ordering is based on WB) is critical. Welfare economics is such an ethic: anything that increases the WB of a member of society is ethically right, as long as it does not decrease the WB of other members. If more WB is the ethical goal, Assumptions 1-9a support the argument that Robin Crusoe⁴⁴ should have unlimited freedom: he can do no wrong, so restricting Robin's freedom is wrong/unethical. If more WB is the ethical goal, and Assumption 9b correct, but not 9a, Robin will not even do what is ethically right for himself, eliminating the ethical justification for unlimited freedom. Assumption 9a, combined with the ethic that more WB is ethically

⁴³“Worms exhibit dietary choice, they hunt for high quality food and leave hard-to-eat bacteria...Using laser ablations and mutant analysis, we show that AIY neurons serve to extent roaming periods and are essential for food seeking” (Boris Shtonda and Leon Avery 2006).

⁴⁴ The classic example of a society where no one else is affected by the one-member's behavior.

preferred, implies something right about an economicus's behavior, at least from its perspective—this is the ethical creep I mentioned earlier.

Alternatively, if you subscribe to an ethic that assumes the goal is to fulfill more wants and desires, then if for Robin, Assumption 9b holds, but not 9a, Robin can do no wrong.

An Important question is whether humans violate the assumptions of NBT (Assumptions 1-9). Much of Part I is devoted to this question.

Expected utility theory, EUT, is a restrictive case of NBT (Assumptions 1-9)^{iv}

Expected-utility theory (EUT)—developed in the 20th Century—restrictively specifies how one chooses in an uncertain world. To start, it assumes WB is cardinally meaningful.⁴⁵ Specifically, economici can order WB differences (WB is at least weakly cardinal).⁴⁶ It also requires that a probability of occurrence is associated with every uncertain event. It is a restrictive case of NBT and predicts economicus will choose the available path with the highest expected WB.

An example of expected WB: imagine that you and your partner prefer having a girl over no child, and no child over having a boy. You can, with certainty, have a baby, and the probability it is female is .5. Your cardinal WB from not having a baby is 0. If you have a child, your expected WB is the WB associated with a girl multiplied by .5, plus your WB with a boy multiplied by .5. If WB with a girl is 1 and WB with a boy is -.5, the expected WB with a child is .25, which is greater than zero. So having a child would maximize expected WB. The ordering of alternatives would be the same if the WB from a girl was .5 and from a boy -1, but in this different cardinal situation, EUT says you would choose to not have a baby because the expected WB from having a baby is less than expected WB without one.

⁴⁵ EUT is a prominent form of *Decision theory* which is an umbrella term for theories designed to model choices in an uncertain world. It is an interdisciplinary endeavor with contributions from statisticians, philosophers, theorists, and economists. It also includes, or not, game theory and social-choice theory. For introductions to Decision Theory and EUT, see [Peterson \(2017\)](#), [Steele and Stefánsson \(2016\)](#), and [Briggs \(2017\)](#).

⁴⁶ Imagine that Path *j* generates more WB than Path *k*, Path *k* more than Path *s*, and Path *s* at least more than Path *t*. EUT assumes the individual knows whether going from *k* to *j* increases WB more, less, or the by the same amount, as going from *s* to *t*. If WB has this property, the ordering of paths is, at least, *weakly cardinal*. See the earlier discussion of weak and strong cardinality.

[John von Neumann](#) (1903-1957) and [Oscar Morgenstern](#) (1902-1977) famously demonstrated that an economicus will choose the available path that maximizes their expected WB if one adds to Assumptions 1-8 and 9a four additional assumptions.

- (1) Differences in WB can be ordered, and
- (2) A probability can be associated with each set of uncertain alternatives in each path.
- (3) A continuity assumption, and
- (4) An independence assumption.⁴⁷

Subjects violate the independence assumption in many choice situations, so EUT is not ideal for explaining choices in an uncertain world.⁴⁸

EUT has two forms: a possible explanation of how one should choose in a world where outcomes are probabilistic (the normative form) and a description of how people choose (the descriptive form). Here, EUT is a descriptive theory, a restrictive case of Assumption 1-9a.⁴⁹ Unlike many choice modelers, I never understood why a human would or should always choose to maximize expected WB. I implicitly questioned the continuity and independence assumptions, but without knowing it.

⁴⁷ Consider first the continuity assumption expressed as an example. The upcoming U.S. Presidential election is part of every path. Imagine you rank Donald higher than Bernie, and Bernie higher than Elizabeth, all else the same. The continuity assumption says that there must have been a specific probability, p_d , such that you would have been indifferent between a path where Bernie wins for sure, and a path where Donald wins with probability p_d and Elizabeth wins with probability $(1-p_d)$. This assumption is restrictive but not drastically so. More restrictively, the independence assumption says, all else constant, that if you ranked Donald higher than Bernie, you ranked—for every p —a race between Donald with probability p and Elizabeth with probability $(1-p)$, higher than a race between Bernie with probability p and Elizabeth with probability $(1-p)$. Put simply, the independence assumption says adding Elizabeth to a race that was initially between Donald and Bernie cannot influence your cardinal ordering between Donald and Bernie.

⁴⁸ See Levin (2006) and Kahneman and Tversky (1979). A famous example, the *Allais Paradox*, is from 1953 by the French economist, and Noble Laurate, [Maurice Allais](#). Put simply subjects are presented with two different gambles and asked to indicate which alternative they prefer in each gamble. Many respondents' two choices together imply that they are violating the independence axiom. Johan Lehrer (2010) has a nice, and simple, vacation-choice example. Kahneman and Tversky suggest that the violation occurs because people value sure things (an outcome with certainty) more than is consistent with the independence assumption.

⁴⁹ Note again the phrase “WB, accounting for the uncertainties” in Assumption 9a: EUT restrictively assumes that the ordering criterion is expected WB. Uncertainty can be accounted for in many ways besides a probability-weighted expected value. For example, if one is extremely risk adverse, one might order paths by the least amount of WB that could be realized in each path, a maxi-min strategy—a strategy consistent with NBT but inconsistent with maximizing expected WB.

What else might be presupposed (implicitly assumed)? What, if any, additional assumptions might you want to add to NBT?^v

Presuppositions are things you take for granted (pre-suppose), imagining others do as well. Since tacit, economists rarely think about them, so there is no way of knowing what is presupposed unless one asks. Which, if any, of the following additional assumptions (10-12) have you already implicitly included and now want to make explicit? What have you not already implicitly assumed but now want to? Adding could serve two purposes. (a) Adding 10-12 in sequence further restricts the set of entities whose behavior the theory explains. And (b) Adding each gives the entity more ways to experience a path, and this will allow us to be more specific about what it might mean for different sorts of *economici* to increase their WB. For example, adding Assumption 10 means the theory no longer applies to robots and toasters.

Going forward, I will adopt the following terminology: continue to label any entity whose behavior is consistent with Assumptions 1-9 as an *economicus*. Entities whose behavior is consistent with Assumptions 1-11 are *super-economici*. Those whose behavior is consistent with Assumptions 1-9 plus Assumption 13 are *homo-economici*. If Assumptions 1-9 hold, but not 10 nor 11, the entity is a *primitive-economicus*. What increasing WB means is different for super and primitive *economici*. Deem Assumptions 1-11 *NBT Plus*. If Assumptions 1-9 plus 12 hold, the entity is a *choice economicus*, and if Assumptions 1-12 hold, the entity is a *super choice economicus*.

Assumption 10a: Economici experience thoughts and emotions. Both are provoked by past and present sensations, perceptions, memories, knowledge, cognitive abilities, plus past thoughts and emotions.

A perception is a type of thought, how one interprets a sensation. A perception can depend on what caused the sensation: pain caused by a fall, knee surgery, or a mobster hitting you with a bat. If touched, you sense the touch, but how you perceive it depends on who is doing the touching. An elk sees and smells a wolf (sensations), and these sensations cause a perception. A sensation is necessary but only part of what produces the perception.

A perception is a thought, but not all thoughts are perceptions. While *thoughts* are challenging to characterize, I hope we all agree that thoughts are the product of thinking, and thinking is a cognitive process involving neurons firing. And we agree that a thought is a discrete mental state (they come and go) and differ from sensations. Thoughts can be conscious or

unconscious: unconscious thoughts are ones you are unaware of. Many thoughts are questions: Why is the dog hungry, and Where is the nearest restaurant? Saying “around the corner” is a spoken thought.

A belief is a type of thought (a stored thought), something the individual who holds it regards to be true. *It's a 1964 Malibu. Assumption X implies Assumption Y. Trump made America great again, and if not for a Democratic conspiracy, he would still be President.* Subjective probabilities are beliefs. *It is likely to rain tomorrow. That bus is going to hit that car.*

Humans typically believe that only humans and a few other animals are capable of thought and that only humans have thoughts about anything other than the immediate past and future. Independent of what people believe, not much is known about animal thinking. Whether a machine can have thoughts is debated

There is no agreed-upon definition or theory of *emotions*. There isn't even agreement on how many there are.⁵⁰ Emotions are viewed from three perspectives: “as experiences, as evaluations, and as motivations” ([Andrea Scarantino](#) and [Ronald de Sousa](#) 2018). For example, anger and arousal are experienced; they are evaluations of what's going on and motivate action. Emotions differ on numerous dimensions making it challenging to come up with a one-size-fits-all definition:

some emotions are occurrences (e.g., panic), and others are dispositions (e.g., hostility); some are short-lived (e.g., anger) and others are long-lived (e.g., grief); some involve primitive cognitive processing (e.g., fear of a suddenly looming object), and others involve sophisticated cognitive processing (e.g., fear of losing a chess match); some are conscious (e.g., disgust about an insect in the mouth) and others are unconscious (e.g., unconscious fear of failing in life); some have prototypical facial expressions (e.g., surprise) and others lack them (e.g., regret). Some involve strong motivations to act (e.g., rage) and others do not (e.g., sadness). Some are present across species (e.g., fear) and others are exclusively human (e.g., schadenfreude [pleasure from another's misfortune]) (Scarantino and de Sousa 2018).

⁵⁰ Until recently the number was finite: The Confucian text *Liji* (*Book of Rites*) lists seven; Descartes lists six; in the 1970's the psychologist Paul Ekman listed six. Now there are 27 with fuzzy boundaries ([Alan Cowen](#) and [Dacher Keltner](#) 2017). New ones include aesthetic appreciation, awe, awkwardness, empathetic pain, and nostalgia ([Melissa Dahl](#) 2020).

There is constructed-emotion theory ([Lisa Barrett](#) 2017a and b): while there are basic ones, many are constructed based on one's language and culture. Not everyone buys constructed-emotion theory. It suggests that if you can give the feeling a name it is a type of emotion. Tongue in cheek, consider *Force Majeure* [“a vague sense of humiliation that accompanies an instinctual flight response to a perceived danger. named for the Swedish film in which a man runs away from his family to save himself from an avalanche and thereby is diminished by his fear”] And, *Itchy Teeth* [“the restless urge to explode your own life by doing something that can't easily be undone.”], and *Liegasms* [“The sick pleasure of knowing you're lying and getting away with it”] (editors, NY Magazine 2020).

Emotions have physiological components (e.g., hormonal, cardiovascular, and facial expressions) that can be observed. Still, emotions are more. They're simultaneously physiological and mental states: the same physiological state can be associated with either a positive or negative emotion depending on what is being experienced (e.g., pursued by a lion or watching a horror flick) and who is experiencing it.

It seems that one could have thoughts even if one cannot experience sensations, but I'm not prepared to argue the point.

Assumption 10a implies that each path is associated with a sequence of thoughts and emotions. Assumption 10a, like Assumption 8a, limits what can be an economicus: no more toasters or robots, at least for now.

Assumption 10b: An economicus's ordering of paths (either in terms of WB or in terms of wants and desires) is determined, at least in part, by the sequence of thought and emotions that would be produced by each path.

Whether Path j is associated with more WB than Path k is determined, in part, by the thoughts and emotions it would produce. And whether Path j is more wanted/desired than Path k is determined, in part, by the thoughts and emotions that each would produce. Assumption 10a is necessary but not sufficient for Assumption 10b. For humans, WB-enhancing emotions include pleased and relaxed, and disgust reduces WB.

Assumptions 10a and b don't imply the individual is self-aware.

Assumption 11: Economici are self-aware

Self-awareness is the "ability to take oneself as the object of one's own attention and thoughts", recognize yourself as a physical object, separate from all else, and recognize that what you feel belongs to you. In nature, there are degrees of self-awareness: it consists of five distinct cognitive abilities, and only humans seem to have the ability to think of themselves abstractly and symbolically.⁵¹ Many other species have one or two of these abilities, and a few have four of the five.

⁵¹ This is just one definition of self-awareness; feel free to replace it with another. In philosophy, self-awareness is one type of consciousness. The five types are awareness of yourself in relationship to your immediate physical

Neither Assumption 8a nor 10a implies Assumption 11: animals experience sensations, and animals can have thoughts and emotions, but that does not imply they are self-aware. For example, my dog isn't, but he does suffer from anxiety. Whacking a duck with a stick will cause it pain, but the duck can neither reflect on the pain nor abstractly think about his aches and pains—something humans can and do, making it difficult for us to imagine duckness.⁵² Self-recognition is often used to indicate self-awareness, but they are not the same.

Assumption 11 further limits whose behaviors the theory explains. My dog fails the standard test for self-awareness: he has been known to bark at himself in the mirror, even after I draw a blue dot on his forehead while asleep. [Researchers draw a blue dot on the animal's forehead while it is sleeping. They place a mirror in the room and wait behind the one-way glass to see what it will do when it wakes up. If, when seeing itself in the mirror, the animal touches its forehead, it self-recognizes, and they conclude it is self-aware. Professor Gordon Gallup came up with this test when he cut himself shaving, seeing, in the mirror, blood. Chimps tend to touch the blue dot.] Only a few species (the higher primates, elephants, dolphins, magpies, and possibly pigs) have passed the mirror test, but not everyone is convinced it is a self-awareness test. E.g., self-awareness does not require visual self-recognition.⁵³ Assumption 11 excludes robots, cacti, most animals, a couple of humans I know, and new babies.

Assumption 11 introduces the possibility of “I” and “me” thoughts such as *Wanda loves me*: one can't think reflectively about oneself without self-awareness. *I'm happy* (or *content*, *excited*, *awed*, *calm*, *relaxed*, *proud*, *superior*, *free*, *liked*, or *depressed*). So are, *I'm satisfied*; *I'm accomplished*, or *I'm a Fuck-up*. So, *I want to go skiing*. For the religious, so are *I am experiencing God's grace*, and *I'm a sinner*.

environment (knowing where your body ends and the world starts), awareness of interactions with other members of your species, the ability to reflect on yourself over time, the ability to reflect on your feelings, and the ability to think about yourself symbolically and abstractly (Leary and Buttermore 2003). Giacomo has, at least, the first two.

⁵² A few humans feel no physical pain. The disorder is associated with the SCN9A gene. Ashlyn Blocker, a self-aware teenager with this disorder, was recently profiled in the *N.Y. Times* Sunday Magazine. Ashlyn feels pressure, but not physical pain. She does experience emotional pain. Someone in a deep coma neither feels nor is self-aware.

⁵³ For a video on self-recognition by an elephant go to <http://www.pnas.org/content/suppl/2006/10/26/0608062103.DC1>

Thoughts about NBT Plus

Like NBT, NBT Plus does not require much in the way of logic or rationality. Super-economici are no more rational than primitive economici; Assumptions 10 and 11 only add a bit of cogitating.

Assumptions 9-11 raise many questions about how sensations, thoughts, and emotions map into WB and wants and desires. Consider all the different things that affect your WB: sensual pleasures (tastes, sounds, sights), happiness, contentment, excitement, a sense of calm, relaxation, a meaningful/rich life, satisfaction (including life satisfaction), other positive thoughts (including those about the future), learning, pride, loving, being loved, a sense of accomplishment, interactions with others, interactions with animals and nature, being free, being equal or being superior, justice being served, different sorts of personal pain, anxiety, worry, depression, hunger, thirst, embarrassment, fear, sadness, anger, feeling sick, experiencing disability, your reactions to these by others, negative thoughts, loneliness, discriminated against, feeling inferior, and feeling controlled. Assumptions 9-11 force me and you to think about the mapping process and what it might look like. For example, the mapping into WB requires that all these diverse sensations, thoughts, and emotions are commensurable, and each can be separated from the act or situation that produced it. Wow!

But—I have no recollection of seeing the words, *sensations, perceptions, thoughts, or emotions* in the standard economic literature on choice. Economists don't talk about them. This is unfortunate: they are what make us feel alive.

Assumptions 8, 10, and 11 (self-awareness) imply a self—a **me**—who experiences the world with sensations, thoughts, and emotions, all from a personal me perspective. There is a me embodied in every super-economicus, a me who experiences sensations, thoughts, and emotions as his.⁵⁴ In contrast, when my dog experiences pain and perceives danger, he reacts, but he does not consciously think about it as his pain and a danger to him because there is no conscious him

⁵⁴ The late Harvard psychologist Daniel Wegner has an insightful definition of the self: "The sense we have that we are agents who do things and experience things, and who in some regard are the same from one time to another. This sense of identity is inherent in the aspect that William James (1890) called the 'knower,' the self that is the seat of experience, the self that is doing all our thinking and living. Another way of conceptualizing identity is the self as the object one can think about, the aspect that James called 'known'."

Another take on the self, a Buddhist take, is it does not exist, it's empty—*no-self*.

for my dog. Self-awareness complicates the mapping from sensations, thoughts, and emotions to WB. Same for the mapping to wants and desires.

Together with Assumptions 8, 10, and 11, Assumption 9b is [David Hume's](#) (1711-76) view of human nature: emotions, not reason, drive behavior— “Reason is, and ought only to be the slave of the passions and can never pretend to any other office than to serve and obey them.” (Hume 1738).

With Assumptions 8, 10, and 11, both the researcher and policymaker must consider what WB means for super economici: economici with sensations, perceptions, thoughts, and emotions. My view is that WB can depend on them all, but how much weight each gets likely varies across super economici and can vary over time for an individual super-economicus. And maybe for some individuals, some kinds of WB are not comparable (WB-incommensurable).

WB: a trichotomy

For [Daniel Kahneman](#) and [Angus Deaton](#) (2010)—both Nobel Laureates in Economics—kinds of WB fall into two categories: *emotional WB* and *life-satisfaction WB*. I prefer three categories. Their first is “the emotional quality of an individual’s everyday experiences—the frequency and intensity of experiences of joy, stress, sadness, anger, and affection that makes one’s life pleasant or unpleasant.”⁵⁵ This component of WB is only relevant to economici that experience emotions, so only relevant to super economici.

Life-satisfaction WB, in contrast, is economicus’s thoughts about how well their life is going. This component of WB is only relevant to economici that have thoughts and are self-aware, so, again, only relevant to super economici. Life-satisfaction WB is the thought component of WB.

Standard survey questions about emotional WB include: “Did you experience a lot of stress (enjoyment, happiness, anger, sadness, stress, worry) yesterday? Another is “How happy are you these days?” where happiness (unhappiness) is adopted as a catchall for all the components of emotional WB. In contrast, “Rate your life on a ladder scale of 0 to 10 where ‘0 is the worst possible life for you’, and ‘10 is the best possible life for you’” is a life-satisfaction

⁵⁵ Emotional WB is sometimes defined as the ability to manage one’s emotions. It is not how I am defining it here.

question. Another is “How satisfied are you with your life these days?” Few researchers ask emotional WB and life-satisfaction questions.

Until recently, those studying WB have either not thought about the distinction between emotional WB and life-satisfaction WB or viewed the distinction as unimportant because they incorrectly believed the two were influenced by the same factors in the same ways. But this isn’t the case, and the distinction is critical. For example, after a crucial dollar amount, more income will increase your life-satisfaction WB but maybe not your emotional WB.

My third category is *sensations WB*: all else constant, WB increases (decreases) with the duration and intensity of pleasurable (dis-pleasurable) sensations. Orgasm is a classic example of a pleasurable sensation. As noted, for me, the taste of sweet-n-salty is near the top. The sensations associated with burning skin and not being able to breathe are dis-pleasurable, and so is hunger. All else constant, economici repeat behaviors that produce pleasurable sensations and don’t repeat behaviors that produce displeasurable sensations. But “All else constant” is important. Gasping for breath could increase overall WB if it, for example, ends with you winning the Boston Marathon.

I don’t recollect seeing a survey that asks people about the number and intensities of their pleasurable sensations. Doctors and nurses often ask, “On a scale of 1-10, how much does it hurt?”

Do you want to require WB to have cardinal properties or require that it has no cardinal properties?^{vi}

As noted earlier, Assumptions 2 and 9a imply that economicus has, at a minimum, a complete ordering of paths based on WB. But together, Assumptions 2 and 9a neither imply nor preclude WB from having cardinal properties. Because of this, you might want to replace Assumption 2 with the assumption that that ordering must have cardinal properties. Or, you might want to replace 2 with an ordering that does not have cardinal properties.

Assumption 2: At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, An economicus has an ordering if for all paths j and k, either Path j is ranked higher than Path k, Path k is ranked higher than Path j, or paths j and k have the same rank.

Assumption 2(Cardinal): *At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, An economicus has an ordering if for all paths j and k, either Path j is ranked higher than Path k, Path k is ranked higher than Path j, or paths j and k have the same rank. And, economicus can rank the shift from i to k relative to a shift from l to m.*

Assumption 2(Ordinal): *At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, An economicus has an ordering if for all paths j and k, either Path j is ranked higher than Path k, Path k is ranked higher than Path j, or paths j and k have the same rank. And economicus can't rank a shift from i to k relative to the shift from l to m.*

Deem economici who conform to Assumption 2(Cardinal), *cardinal economici*. Deem economici who conform to Assumption 2(Ordinal), *ordinal economici*. An ordinal economicus or other. So, why might you want to replace Assumption 2 with one of these more restrictive assumptions?

Some history: 19th Century neoclassical economists adopted Assumption 2(Cardinal) because they thought cardinal WB was necessary to predict behavior. They were wrong. In the late 19th Century, economists figured out that cardinal WB in a world of certainty adds nothing to the model's predictive power;⁵⁶ the theory with Assumptions 2(Cardinal) and the theory with Assumption 2 predicts the same thing: economicus will experience their highest-ranked available path. He will experience it whether he comprehends it's associated with a lot more, or a little more, WB than the second-ranked available path. And, if he can't comprehend such differences, he still goes with the HRAP. If the two available paths are marry Shirley and have five kids or marry Fred and adopt two, and the Fred path is associated with more WB, you marry Fred—end of story.

Since Assumptions 2 and 9a generate identical predictions to Assumptions 2(Cardinal) and 9a, and since 2(Cardinal) further restricts who the theory applies to, Assumption 2 (Cardinal) was dropped from NCT.⁵⁷ But do you replace it with Assumption 2, an assumption that does not

⁵⁶ Dropping the assumption that WB is cardinal was first proposed by the German economist and mathematician Andreas Voigt in 1893, but he has been largely forgotten with the credit going to Pareto (Schmidt and Weber 2008). The switch from cardinal to ordinal was solidified by R.G.D. Allen and John Hicks in 1934—who said there was no need for cardinal WB in choice theory.

⁵⁷ If a theory's predictions don't change if a restrictive assumption is removed, the assumption should be removed. At least William of Ockham (1287-1347) thought so: if a restrictive assumption is removed, it applies in more situations. Use *Ockham's razor* to shave away the unnecessary restrictions. Replacing an ordering with cardinal

exclude cardinality, or do you replace it with Assumption 2(Ordinal), which, together with 8a, bans cardinal WB? NBT adopts Assumption 2, but for much of my career, I thought 2(Ordinal) was assumed. The distinction is between allowing for cardinal WB in a theory where its existence is superfluous to explaining behavior and banning cardinal WB because the WB of economici and humans don't have cardinal properties.

How to decide? To explain behavior in a world of uncertainty? Cardinal WB is not required to explain behavior in a world of uncertainty. But is required if you explain behavior in a world of uncertainty with a theory that predicts economici maximizes expected-WB (expected utility theory, EUT). Proponents of EUT require Assumption 2(Cardinal).

Another reason to suppose cardinal WB is that you believe that many economici have it. We would like to think that we can comprehend and compare whether a shift from Path k to m would increase WB more or less than a shift from c to l . Humans describe their sensations, perceptions, and emotions in cardinal ways. For example, the perceptions and emotions I experience seem to have intensity ("The taste sensation I get from chocolate ice cream is a little more pleasurable than from vanilla, but the pleasure sensation from tutti-frutti is way more pleasurable than from chocolate."). Wouldn't you expect WB to be cardinal if sensations, perceptions, and emotions have cardinal properties? For example, I would like to think that a starving me obtaining enough to eat would increase my WB more than my current WB would increase if someone gave me their new SUV in exchange for my old SUV.

But maybe I am deluded in thinking that my WB has cardinal properties. Perhaps my WB is ordinal. It is an unanswered neurological research question.

Would you expect evolution to lead to cardinal WB? Since we live with uncertainty, it would depend on how evolutionary advantageous cardinal WB is vs. ordinal WB when there is uncertainty.⁵⁸ I don't know.

If WB depends on sensations, thoughts, and emotions, and if behavior is motivated by WB, understanding your behavior requires understanding how your sensations, thoughts, and

properties (which is an ordering of paths with intensity added) with only an ordering is such a generalization/relaxation.

⁵⁸ Recollect, that cardinal WB does not affect behavior when there is certainty.

emotions map into your WB. While explaining this mapping is beyond my pay grade, the map depends on which experiences have or don't have cardinal properties.

Whether WB has cardinal meaning and to what degree is an empirical question. So are whether sensations, thoughts, and emotions have cardinal properties.

Another implicit assumption

Assumption 12: Before a path is taken, an economicus has the experience of consciously choosing some components of that path. And that perception of choosing caused them to take a path with those components.

Assumption 12 is a choosing axiom; it makes economicus consciously in charge of which path will be experienced—if you are an economicus and 12 holds, you choose, and this determines the path you take—so you are responsible for your behavior. A conscious perception of choosing adds to economicus's sense of self (who they are) and the belief that others are responsible for their behaviors. [We commonly ask if a person should be held morally accountable if their behavior wasn't chosen. For example, Pinocchio should be excused for not listening to his father: Geppetto forget to carve him ears. We find children, the insane, and animals not morally responsible, but there are exceptions: in 1386, the French hanged a pig for murdering a child—it had flopped down, suffocating the kid.]

A reason to adopt Assumption 12 is many humans believe it's true—at least for humans. [Of course, you can't choose every component of your path. For example, that Mexico is south of the U.S. and Trudeau is Canada's P.M. are components of the path I'm experiencing (same for you), but I chose neither Canada's P.M. nor the U.S.'s location. And, while what you consume and do is a component of my path, typically, I don't get to choose what you consume or do.] Assumption 12 is critical to the economic sense of choice and the ethics of economics, and it's crucial in ethics that assume more freedom is better than less.

Assumption 12 embeds two assumptions:^{vii}

Assumption 12a: Before a path is taken, an economicus has the experience of consciously choosing at least some components of that path.

And Assumption 12b: This perception of choosing caused them to take the path they did.

A necessary condition for 12a is consciousness: the ability to experience conscious mental states.⁵⁹ Therefore, assumptions 10a and 11 are required. While 12b requires 12a, 12a does not imply 12b. If you experience choosing (12a), but that experience does not, in fact, influence what you experience (12b), you suffer from the *illusion of choice*. People, including economists, accept Assumption 12b as gospel. However, many eminent neurologists reject it. Their reasons are presented in Part II.

Assumption 12 implies an economicus can consciously retrieve his ordering, and could, if so inclined, tell a researcher which alternative he would choose in any set of alternatives.⁶⁰ Without this conscious knowledge, you wouldn't be able to do this, and researchers wouldn't be able to learn about your ordering simply by asking you (something I have done many times in surveys).

More significantly, Assumption 12 implies that economicus experience the available path with the most WB only because they went through a conscious process—choosing. If Assumption 12 is violated, the predictions of NBT and NBT Plus won't change, but the path taken wouldn't be consciously chosen, making me, and probably you, hesitant to say, "Economicus made a conscious choice".⁶¹ A question concerning Assumption 12 is whether it is consistent with Assumptions 1-9, and, if so, what *conscious choosing* might mean. Without Assumption 12, the path taken could be determined entirely by an unconscious process—a simple process (identify the HRAP). There isn't much to think about, either consciously or unconsciously.

If you insist, you can add Assumption 13,

⁵⁹ This is simply one of many definitions of consciousness (Robert Van Gulick 2018). "On one common reading, a conscious mental state is simply a mental state one is aware of being in ... Conscious states in this sense involve a form of *meta-mentality* or *metaintentionality* in so far as they require mental states that are themselves about mental states. To have a conscious desire for a cup of coffee is to have such a desire and also to be simultaneously and directly aware that one has such a desire. Unconscious thoughts and desires in this sense are simply those we have without being aware of having them" (Van Gulick 2018). Thoughts are conscious if you are aware of them.

⁶⁰ You can be self-aware without being aware of your ordering of paths. If you are consciously aware of your ordering and your constraints, you know whether you are doing the best you can. Alternatively, if you do not consciously know your orderings, you have no conscious awareness that you are doing the best you can. In which case, your incentive to keep doing the same thing, if you have such an incentive, must be unconscious.

⁶¹ If you do not have conscious awareness of your ordering, the path you experience isn't determined by any conscious thought process—even if you are experiences choosing. And, in this case, if you believe you can correctly describe why you do what you do, you are wrong.

*Assumption 13: **Economici are Homo-sapiens.***

With it, the theory limits itself to explaining the behaviors of humans who behave in a manner consistent with Assumptions 1-9. Earlier, I laid out reasons for not limiting my NBT to humans; researchers are more likely to find *economici* amongst the entities who reside in zoos, forests, and gardens than at football games. And, if we exclude non-human animals but want to adopt an ethic that gives more standing to them or asserts humans have a duty to them, we will need a different model for animal behavior and WB.

One reason to include Assumption 13 is that many people want to believe, and need to believe, that humans are superior and separate from the animal kingdom. A way to separate us from the beasts is to assume human behaviors are chosen and animal behaviors are not. If one is in this category, they need their theory of human behavior to be a choice theory that does not apply to other animals.

Assumption 13 implies Assumptions 8a, 9a, 10a, and 11, making them redundant if all living humans have sensations, thoughts, and emotions and are self-aware.

So, what to say about an *economicus*, a person or thing that follows the rules?^{viii}, NBT, like more theories of behavior, assumes one theory of behavior explains the behavior of all humans: the way people think and perceive the world, along with their sense of self, is universal. [Joseph Henrich](#), [Steven Heine](#), and [Ara Norenzayan](#) explain the assumption [not what they believe].

We all have the same hardwiring, so we must all make choices in the same way...The fund analyst in NYC operates with the same axioms as the subsistence farmer in Peru.

My NBT even expands the applicability to other living things. The assumption that all human behavior can be explained by one set of assumptions is based on two things: (a) We all share the same genome, the *human genome* (“23 chromosome pairs with a total of about 3 billion DNA base pairs...24 distinct human chromosomes: 22 autosomal chromosomes, plus the sex-determining X and Y chromosomes.”). And (b), Significant differences in how humans perceive and think and what motivates them have not been observed. The latter, (b), is because, until recently, no one looked. Even today, most behavior research is done on *WEIRD* people (Western, Educated, Industrialized, Rich, and Democratic), often undergraduate psychology or economic majors. Remember that NCT was created by old, white, male protestants, many of whom lived in

19th Century Great Britain—they were not thinking about explaining the behaviors of “savages” and other foreigners.

Maybe one theory of behavior can’t explain the behaviors of all humans. Recent research indicates that when it comes to how we think and perceive the world, WEIRDS are the statistical outliers. Paraphrasing Henrich et al. (2005 and 2010): the universality assumption has been under assault, starting in the 1960s when fundamental differences in visual perception across cultures were observed. It turns out that people's sense of fairness, how they decide, their sense of self, their moral reasoning, and even their brain scans vary significantly by culture, and WEIRD people are typically the statistical outliers. The point is how individual humans behave varies by population and culture.

Homo economicus is a Western sort of guy^{ix}

Economicus can’t have an inconsistent ordering of paths. In addition, his behavior is directed, without flaws, to the HRAP. Every kind of WB is commensurable for him: he’s a Western ideal. Formal logic was first studied in Ancient Greece, infusing how we in the West view thinking, decision-making, and choice. Asian thinking tends to be more intuitive, holistic, and accepting of what Westerners view as contradictions.⁶² In Ancient China, there was nothing comparable to the Ancient Greek system of formal logic; instead, a dialectic (method of reasoning and deciding) was developed that allows for contradictions, and this approach has infused Asian thinking. Of course, we all do both types, and thinking patterns vary widely within a culture, but how people tend to think and value different ways of thinking varies across cultures.⁶³

Research by the psychologists [Emma Buchtel](#) and Ara Norenzayan demonstrates that Asians think more intuitively, judging intuitive reasoning as more reasonable than analytic

⁶²To clarify the distinction, quoting Henrich et al.: *Holistic thought involves an orientation to the context or field as a whole, including attention to relationships between a focal object and the field, and a preference for explaining and predicting events on the basis of such relationships. Analytic thought involves a detachment of objects from contexts, a tendency to focus on objects’ attributes, and a preference for using categorical rules to explain and predict behavior. This distinction between habits of thought rests on a theoretical partition between two reasoning systems. One system is associative, and its computations reflect similarity and contiguity (i.e., whether two stimuli share perceptual resemblances and co-occur in time); the other system relies on abstract, symbolic representational systems, and its computations reflect a rule-based structure.*

⁶³ Learning about logic, marginal analysis, and benefit-cost analysis might cause one to reason more analytically, causing economists to believe man behaves logically.

reasoning. Their article starts with a quote on analytical versus intuitive thinking from a Chinese scholar in 1939. Expressing the Chinese view:

. . . We see an opposition of 'logic' versus common sense, which takes the place of inductive and deductive reasoning in China. Common sense is often saner because the analytic reasoning looks at truth by cutting it up into various aspects, thus throwing them out of their natural bearings, while common sense seizes the situation as a living whole . . . Logic without such common sense is dangerous . . . (Dr. Yutang Lin)

Different ways of thinking aren't just an East/West thing. For example, southern Italians think more holistically than Northern Italians, and working-class Southern Italians think more holistically than middle-class Southern Italians. [Richard Nisbett](#), a social psychologist, has been a pioneer in finding cross-cultural differences in cognition and evidence supporting the hypothesis (the *Social Orientation Hypothesis*) that these differences are due to culture rather than genetic or linguistic differences.

Do any economici exist, or are they mythical creatures? And does it matter?

Much of Part I is devoted to reporting on and discussing the recent psychological, neurological, philosophical, and behavioral economic research on the behavior of humans and other animals. In brief, Assumptions 1 and 4 are on solid footing: you can only live one life at every point in time, and you can't do the impossible. Concerning Assumption 2 (an economicus has a complete ordering of all paths), it's hard to imagine why it would be true. Based on the evidence that will be presented, I doubt whether Assumptions 7 (you experience your HRAP) and 9a (paths with more WB are consistently ranked higher) apply to all, or even some, humans, and doubt whether they apply to chimps, dolphins, and elephants. The realism of other assumptions is also in doubt.

That said, no theory's assumptions realistically describe all the situations where the model is applied—the apple that fell on Newton's head was not falling in a vacuum. Theories are not rejected solely because they assume unrealistic stuff. Milton Friedman's stance was that realism is immaterial, and all that matters is whether the predictions do a good/adequate job predicting behavior. A more nuanced view is a model that adequately predicts behavior because its assumptions, while not realistic, approximate reality. Of course, adequately predicting is a matter of degree: different applications of NCT, including NBT and NBT plus, explain how people react, in aggregate, to changes in their constraints (estimated aggregate-demand curves slope down and shift out when aggregate income increases), but they are not adept at explaining

the behavior of specific individuals.⁶⁴ NCT has been applied to estimate the demands for thousands of goods and services, public goods, environmental quality, and fishing sites—you name it. And the results are used to inform policy by estimating how much demand will change if, for example, a price changes or estimating what an average person will pay for a quality change. Economists have used it to estimate how much the demand for cigarettes by teenagers will drop if the tax on cigarettes increases and estimate willing-to-pay for improvement in environmental quality. I have spent much of my life estimating site-specific recreational demand as a function of environmental conditions, and my estimates have informed damages estimates in numerous NRDA's (Natural Resource Damage Assessments). To summarize, different versions of NCT are widely used in research and to inform policy.

Neoclassical choice theory, including my NBT, is an idealized caricature of how an individual behaves, its creation motivated partly by the mathematization of economics that started in the 19th Century. Caricatures intentionally distort. NCT has morphed into a story about choice based on optimizing, a self-appealing story. It is also a simple story. That said, the current economicus is quite different from economic man as conceived by [Alfred Marshall](#) (1842-1924), and his conception is different from [Adam Smith](#)'s (1723-1790). Economicus continues to evolve; hopefully, he will grow into a more complicated and flawed creature who behaves but does not always choose his behavior—behavioral man—a subject of Behavioral Economics. What is unfortunate is that the current economicus is the only character that many students of economics encounter, but this is starting to change.

Before continuing our discussion of behavior and choice, Chapter 2 jumps to welfare economics: how many economic ethicists parse right from wrong actions. Chapter 3 returns to behavior and choice, absent ethical considerations. Finally, in Chapter 11, I will come back to ethics, briefly surveying other ethics.

⁶⁴ A cynic might argue that NCT predicts, for individuals, just one thing: if you can't afford to buy q units of x , you won't. Picturing an individual's demand function for a particular good (price on the horizontal axis, units purchased on the vertical). At current prices, most people do not purchase most goods and won't start purchasing them if their prices fall. I have never purchased a Bugatti, inline roller blades, or liver, and won't even if their prices fall 50%. And for much of what we do purchase, we also don't necessarily buy more when the price drops—I can't drink more Coke Zero than I already do. individual demand functions for many goods are flat over large ranges in price.

ⁱ Sources:

1. Hugo Mercier and Dan Sperber (2017) *The Enigma of Reason*. Cambridge MA, Harvard University Press

ⁱⁱ Sources:

1. Natalie Angier (2009) Sorry Vegans: Brussels sprouts like to live too. *N.Y. Times*, December 22. <http://www.nytimes.com/2009/12/22/science/22angi.html>
2. Stephen Batchelor (2015) *After Buddhism, rethinking the dharma for a secular age* Yale University Press. <http://yalebooks.yale.edu/book/9780300205183/after-buddhism>
3. Samuel Bowles (1998) Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. *J. of Econ. Literature* Vol. 36 (1): 75-111. <http://www.jstor.org/stable/2564952>
4. Eric Bremner, Rainer Stahlberg, Stefano Mancuso, Jorge Vivanco, Frantisek Baluska, and Elizabeth Von Volkenburg (2006) Plant Neurobiology: An integrated view of plant signaling, *Trends in Plant Science* Vol.11(8):1360-85.
5. James J. Cox, Frank Riemann, Adeline K. Nicholas, Gemma Thornton, Emma Roberts, Kelly Springell, Gulshan Karbani, Hassain Jafri, Jovaria Mannan, Yasmin Raashid, Lihadh Al-Gazali, Henan Hamamy, Enza Maria Valente, Shaun Gorman, Richard Williams, Duncan P. McHale, John N. Wood, Fiona M. Gribble and C. Geoffrey Woods (2006) SCN9A channelopathy cause a congenital inability to experience pain. *Nature* 444, 894-898.
6. Charles and Francis Darwin (1880) The power of movement in plants, <https://gutenberg.org/ebooks/5605>
7. Franz Dietrich and Christian List (2012) Where do preferences come from? *International J. of Game Theory*. ISSN 0020-7276, DOI: [10.1007/s00182-012-0333-y](https://doi.org/10.1007/s00182-012-0333-y)
8. Thomas Eisner (2003) *For love of insects*, Cambridge, Harvard University Press. <https://books.google.com/books?isbn=0674018273>
9. David Eli and Justin Rao (2011) The good-news-bad news effect: asymmetric processing of objective information about yourself. *Amer. Econ. J.: Microeconomics* 3(2):114-118.
10. Robert Elwood (2012) Evidence for pain in decapod crustaceans. *Animal Welfare* Vol. 21(SUPPL. 2): 23-27. DOI: [10.7120/096272812X13353700593365](https://doi.org/10.7120/096272812X13353700593365)
11. Ananda Ganguly and Joshua Tasoff (2014) Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future. A Claremont McKenna College Robert Day School of Economics and Finance research paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2370983
12. T. Grüne-Yanoff and S. O. Hansson, eds. (2009) *Preference Change: Approaches from Philosophy, Economics and Psychology*, Dordrecht, Springer.
13. Alan Hajek (2019) Interpretation of Probability. Stanford Encyclopedia of Philosophy.
14. John Harsanyi (1982) Morality and the theory of rational behavior. In *Utilitarianism and Beyond* (eds. Amartya Sen and Bernard Williams), Cambridge, Cambridge University Press. Originally published in 1977 in Soc. *Thical sRes.* 44(4).
15. Daniel M. Hausman (2012) *Preference, value, choice, and welfare*. New York: Cambridge University Press.
16. Justin Heckert (2012) The hazards of growing up painlessly. *N.Y. Times, Sunday Magazine*, November 15. <http://www.nytimes.com/2012/11/18/magazine/ashlyn-blocker-feels-no-pain.html>
17. David Hume (1739-40) A treatise of human nature <https://people.rit.edu/wlrgsh/HumeTreatise.pdf>
18. Daniel Kahneman and Angus Deaton (2010) High income improves evaluation of life but not emotional well-being. *Proc. of the National Academy of Sci.* 107(38): 16489-93.
19. Daniel Kahneman and Robert Sugden (2005) Experienced utility as a standard for policy. *Envir. and Resource Econ.* 32:161-81.
20. David Kreps (2013) *Microeconomic Foundations I: Choice and Competitive Markets/* Princeton University Press.
21. Kelvin Lancaster (1966) A New Approach to Consumer Theory. *J. of Political Economy* 74(2): 132-157. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.295.4494&rep=rep1&type=pdf>
22. Stefano Mancuso and Alessandra Viola (2018) *Brilliant Green: The Surprising History and Science of Plant Intelligence*, Island Press.
23. Edward Morey and Mara Thiene (2017) Can personality traits explain where and with whom you recreate? A latent-class site-choice model informed by estimates from mixed-mode LC cluster models with latent-personality traits. *Ecological Economics* 138: 223-37.

-
24. Bertrand Russell (1950) *Unpopular Essays*: London: Allen and Unwin (first published in Routledge Classics in 2009).
 25. Eric Schwitzgebel (Fall 2019), Belief, SEP
 26. Roberta Scorrane (Sept. 21, 2021) Mancuso, il botanico attaccato dagli estremisti vegani: Mi accusano di diffendere chi produce carne, *Corriere Della Sera*.
 27. Boris Shtonda and Leon Avery (2006) Dietary choice behavior in *Caenorhabditis elegans*. *J. of experimental biology* 209(1): 89-102. jeb.biologists.org/content/209/1/89.short
 28. George Stigler and Gary Becker (1977) De gustibus non est disputandum, *Amer. Econ. Rev.* 67: 76-90.
 29. Kate Sweeny, Darya Melnyk, Wendi Miller, and James Shepperd, (2010) Information Avoidance: Who, What, When, and Why. *Rev. of Gen. Psych.* 14 (4), 340-353. <http://psycnet.apa.org/journals/gpr/14/4/340/> includes a review of literature on the avoidance of health information.
 30. S.S. Venneman, P. Narnor-Harris, M. Perish and M. Hamilton (2006) ‘Allow natural death’ versus ‘do not resuscitate’; three words that can change a life. *J. of Medical Ethics* 34:2-6.
 31. Karen Weintraub (2018) The Swiss consider the lobster. It feels pain. They decide, *N.Y. Times*, January 12. www.nytimes.com/2018/01/12/science/lobster-pain-swiss.html

iii Sources:

1. Jonathan Aldred 2006. “Incommensurability and monetary valuation”. *Land Economics*. 82(2): 141-161.
2. _____ 2013. “Justifying precautionary policies: incommensurability and uncertainty.” *Ecological Economics* 96: 132-40.
3. Ruth Chang (1997) “Introduction,” in *Incommensurability, Incomparability, and Practical Reason*. Ruth Chang (ed.), Cambridge, Harvard University Press.
4. Nien-hê Hsieh (2007) Is incomparability a problem for anyone? *Economics and Philosophy*, 23: 65-80. doi:10.1017/S0266267107001228
5. Locke, J. 1700. 4th ed. “Book II: Chapter XX: On modes of pleasure and pain,” In *An essay on human understanding*. Oxford: Oxford University Press, 1975.
6. Martha Nussbaum (2012) Who is the happy warrior? Philosophy, happiness research, and public policy. *International Review of Econ* 59:335–361.
7. David Pearce 2000. “Cost-Benefit Analysis and Environmental Policy”. In *Environmental Policy*, ed. D. Helm. Oxford: Oxford University Press.
8. Rekola, M. 2003. “Lexicographic preferences in contingent valuation: a theoretical framework with illustrations”. *Land Economics*. 79(2): 277-91.
9. Amartya Sen 1997. “Maximization and the act of choice”. *Econometrica*. 65(4): 745-799.
10. Patricia Sheridan (Summer 2016) “Locke's Moral Philosophy”. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). <https://plato.stanford.edu/entries/locke-moral/>
11. Clive Spash and Nick Hanley 1995. “Preferences, information and biodiversity preservation”. *Ecological Economics*. 12: 191-208.
12. Cass Sunstein (1994) Incommensurability and valuation in law. *Michigan Law Review* 92(4): 779-861.
13. D. Wiggins 1997, Incommensurability: Four Proposals, In *Incommensurability, Incomparability, and Practical Reason*, Ruth Chang (ed.), Cambridge, Harvard University Press.

iv Sources:

1. Ray Briggs (2017) SEP on Normative Theories of Rational Choice. Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/spr2017/entries/rationality-normative-utility>
2. Daniel Kahneman and Amos Tversky (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47(2): 263-291. At <https://www.jstor.org/stable/1914185>. This article has been cited over 45K times.
3. Jonan Lehrer (10/21/2010) “Allais paradox.” *Wired*.
4. Jonathan Levin (2006) Lecture notes on *Choice under uncertainty* for Economics 202: Graduate Microeconomics (first quarter), Stanford University. At <http://web.stanford.edu/~jdlevin/teaching.html>
5. Martin Peterson (2017) *An introduction to decision theory*, 2nd edition. Oxford: Oxford University Press.
6. Katie Steele and Orri Stefánsson (2016) SEP on Decision Theory. Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/win2016/entries/decision-theory/>

-
- John von Neumann and Oskar Morgenstern (1947) *Theory of games and economic behavior*, 2nd ed., Princeton, Princeton University Press.

^v Sources:

- Simon Blackburn (2001) *Being good: a short introduction to ethics*. Oxford, Oxford University Press—a nice read.**

Lisa Barrett 2017. *How Emotions Are Made: The Secret Life the Brain*. New York, NY: Houghton-Mifflin-Harcourt.

- _____ 2017b. “The theory of constructed emotion: an active inference account of interoception and categorization.” *Social Cognition and Affective Neuroscience*. 12(1): 1-23. doi.org/10.1093
- Stephen Batchelor (2015) *After Buddhism, rethinking the dharma for a secular age*, Yale University Press. <http://yalebooks.yale.edu/book/9780300205183/after-buddhism>
- Alan Cowen and Dacher Keltner. 2017. “Self-report captures 27 distinct categories of emotion bridged by continuous gradients.” *PNAS* 114 (38): E7900-09. <https://doi.org/10.1073/pnas.1702247114>
- Melissa Dahl Feb. 3, 2020. “If you can say it, you can feel it: Some scientists believe you have infinite emotions so long as we can name them.” *New York Magazine*. <https://www.thecut.com/2020/02/new-theory-of-complex-emotions.html>
- Editors, *New York Magazine*. Feb. 3, 2020. “Liegasm, jealousy, and feminamity: introducing 78 new emotions.”
- Mark Leary and Nicole Buttermore (2003) The evolution of the human self: tracing the natural history of self-awareness. *J. for the Theory of Soc. Behavior* 33(4):
- J.M. Plotnik, F.B.M. de Waal and D. Reiss (2006) Self-recognition in an Asian elephant. *Proc.s of the National Academy of Sci. of the United States of America*, published online October 2006. <http://www.pnas.org/content/103/45/17053.full.pdf+html>
- Andrea Scarantino and Ronald de Sousa (Winter 2018) Emotion. *Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/entries/emotion/>
- Robert Van Gulick (Spring 2018) Consciousness. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2018/entries/consciousness>

^{vi} Sources:

- John Hicks and R. G. D. Allen (1934) A Reconsideration of the Theory of Value, Parts 1 and 2. *Economica* 1: 52-76 and 2: 196-219. As a graduate student I spent many hours studying the works of Hicks and Allen.
- Edward Morey (1984) Confuser Surplus, *Amer. Econ. Rev.* 74 (1):163-173.
- Vilfredo Pareto (1909) *Manual of Political Economy [Manuale di economia politica]* English translation by A. S. Schwier, N.Y., Augustus M. Kelley, 1971. Vilfredo, a famous Swiss economist, gave up economics for sociology.
- Torsten Schmidt and Christian Weber (2008) On the Origins of Ordinal Utility: Andreas Heinrich Voigt and the Mathematicians. *History of Political Economy* 40(3): 481-510.

^{vii} Sources:

- Daniel Wenger (2002) *The Illusion of Conscious Will*, Cambridge MA, MIT Press.

^{viii} Sources:

- Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Smith, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q. Patton, David Tracer, and Richard McElreath (2005) Economic man' in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. and Brain Sci.* 28(6): 795-855.
- Joseph Henrich, Steven Heine, and Ara Norenzayan (2010) The weirdest people in the world. *Behav. and Brain Sci.* 33, 61-135.
- Ethan Watters (2013) We Aren't the World. *Pacific Standard* April: 46-53. This is a lay introduction to the previous two articles.

^{ix} Sources:

-
4. Emma Buchtel and Ara Norenzayan (2008) Which should you use, intuition or logic? Cultural differences in injunctive norms about reasoning. *Asian J. of Soc. Psych.* 11: 264-273.
 5. Joseph Henrich, Steven Heine and Ara Norenzayan (2010) The weirdest people in the world? *Behav. and Brain Sci.* 33: 61-135.
 6. Nicola Knight and Richard Nisbett (2007) Culture, class, and cognition: Evidence from Italy. *J. of Cognition and Culture* 7:283–291.
 7. Richard Nisbett, Kaiping Peng, Incheol Choi and Ara Norenzayan (2001) Culture and Systems of Thought: Holistic Versus Analytic Cognition. *Psych. Rev.* 108(2): 291-310.
 8. Michael Varnum, Igor Grossmann, Shinoku Kitayama and Richard Nisbett, R. E. (2010) The origin of cultural differences in cognition: The social orientation hypothesis. *Current Directions in Psy. Sci* 19(1): 9-13.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2838233/pdf/nihms-178724.pdf>

Chapter 2: Welfare (well-faring) economics as a theory of what is right and wrong—my take¹

Draft: 08142022

Introducing *welfare economics* detours from wondering about the determinants of behavior and choice. The topic is here, rather than later, because its foundation is the NBT prediction (with Assumption 9a, not 9b) that everyone maximizes their own WB. For welfare economics, the moral objective is to maximize societal WB—where societal WB is an aggregation of the WB's of society's members. Both NBT and welfare economics are about WB and maximization. In Chapter 11, other ethics will be reviewed and contrasted with welfare economics.

Four questions: (1) What is welfare economics? (2) Is it an ethic? Yes. (3) How do welfare economists differ from one another? And (4) how do they differ from other economic ethicists? Then utilitarianism is discussed. I was taught and have taught that welfare economics is a type of utilitarianism. I was wrong; it isn't. Welfare economics is an atypical form of *welfare consequentialism*: consequentialist in that whether an act or policy is right or wrong is a function of only its consequences—the adjective “welfare” because the only consequences that matter are the welfare (well-faring) consequences. Most welfare consequentialists are not welfare economists, and few moral philosophers are welfare consequentialists. People are not welfare consequentialists. Welfare economists who accept, without critical thought, that welfare economics is the ethic for parsing right from wrong are not moral philosophers. The few economists who study ethics and argue that welfare economics is the preferred ethic are moral philosophers.

Welfare economics is a method to test right from wrong.ⁱ

Whether it is depends on how one defines both welfare economics and right and wrong. This first section starts by defining an ethic. In the words of [John Stuart Mill](#) (1806-93), an ethic is "a method to test right from wrong." Quoting from the *Journal of Economic Literature* article *Taking Ethics Seriously*, [Hausman](#) and [McPherson](#) (1993)

Welfare economists have sometimes thought that their techniques are virtually morally neutral because they rest on innocuous and uncontroversial moral premises such as the Pareto principle. But the standard definition of a social optimum compares social alternatives exclusively in terms of the goodness of their outcomes (rather than the

¹ Earlier parts of this chapter appeared as Morey (2018)

rightness of their procedures) and identifies the goodness of outcomes with satisfaction of individual preferences. These commitments to value only outcomes and to measure outcomes only in terms of individual utilities are neither neutral nor uncontroversial (see Amartya Sen 1979, and Sen and Bernard Williams 1982).

A *Pareto Improvement, P.I.*, is a change that increases the WB of some members of society without decreasing the WB of any other members. That a P.I. is always right, while obvious to welfare economists, is rejected by, and inconsistent with, many other ethical criteria. Ethics that reject it include virtue ethics,² rights theory, Kantian and other duty-based ethics, ethics based on religion (sins are wrong, no matter how they affect WB), and ethics that judge actions right or wrong based entirely, or in part, on process and intent. John Rawls (1971) rejected the P.I. criterion, arguing that increasing the WB of the better-off while holding constant the WB of the worst-off is wrong. To summarize, the Pareto objective is inconsistent with many other ethics, which leaves no doubt that it is an ethic.

“A moral principle can be viewed as a requirement to exclude the use of certain types of information in moral judgments” (Sen and Williams 1982). Welfare economics excludes the cause and the actor’s intent unless they affect the WB produced. It also excludes from moral consideration how a policy or act affects the WB of individuals who are not members of society.

And welfare economics assumes that my behavior and choices are right for me and society if they don’t affect others.

As the economic philosopher Daniel Hausman (1998) put it, referring to economists, “they are [always] ready with advice on how to make life better.”

So, welfare economics is an ethic: put simply, the moral objective is to maximize a particular aggregation of everyone’s WB (welfare), the aggregation over only society’s members.³ But, at this point, skepticism is in order because I have not fully articulated welfare economics. And—keep in mind that some economists advocate ethics inconsistent with welfare economics, but welfare economics is the predominant ethic amongst economic ethicists.

² Adam Smith was a virtue ethicist (McCloskey 2008); he suggested two guides for behavior: rules against abhorrent acts (killing, raping) and virtue for everything else (Smith 1776 and Fleischacker 2013). He thought the morality of an act depended on both the intent and the consequences, but if the intent was virtuous, the behavior is good, unless, of course, the outcome is disastrous. Virtue (“proper gratitude, kindness, courage, patience, and endurance”) results from empathy for others, our sympathy of others—our “moral sentiments”. Smith was not a welfare economist.

³ Some argue, in my view incorrectly, that it isn’t an ethic but rather a valuable tool for moral philosophers because it shines a harsh light on consequences. A reviewer of one of my papers said, ‘So maybe a necessary stage of moral decision making should be the rigid consequentialist phase: just stop and consider this apart from everything else.’ But, if one takes this view, one can’t parse right from wrong policies based on welfare economics.

Ethics judge actions. Actions include individual behaviors (eating peanuts, smoking in a crowded room, or driving drunk). They include government policies such as going to war, taxation to redistribute income, and taxing carbon emissions. Since economists think on the margin, the welfare-economic question is typically not whether the action is right or wrong, but what is the right amount—the right amounts of chocolate and global warming?

My first inclination was to use the words “good” and “bad” acts but will instead use “right” or “moral” and “wrong” or “immoral” as in “morally right” and “morally wrong”. The words “good” and “bad” are used in many ways, so words that are easy to misinterpret.⁴ Sound good? Morally right sounds pretentious but is more precise. Morally wrong means it should not happen for moral reasons. To appear less pretentious, I will typically drop the adjective “morally” and stick with “right” and “wrong”.

Based on its criterion for right vs. wrong, welfare economics can, in theory, judge whether an act is right or wrong but often can’t in practice. This is true of many ethics. Policymakers are lucky when an ethic indicates whether a particular act is right or wrong. Looking ahead, you might disagree with my characterization of welfare economics, citing examples of individuals who self-identify as welfare economists but who would disagree with my definition. This is OK: the objective isn’t for you and me to agree on what welfare economists agree. Instead, the aim is to evaluate ethical assumptions.

It is crucial to know that almost everyone rejects welfare economics as a criterion to parse right from wrong. More on this, but only after we list its commonalities. Only after welfare consequentialism and economics are fully articulated will I address whether welfare economists are best described as utilitarians. This chapter reviews the research on what criteria people actually use to morally judge.

Welfare economists are consequentialists, specifically welfare consequentialistsⁱⁱ

Welfare consequentialism (hereafter “WC”) was defined above; now, it’s detailed. First, I shift back and forth between WC (the general category) and welfare economics (an atypical type).⁵ Then I list major concerns about WC as a criterion for parsing right from wrong. Even many

⁴ Consider: “Phil is a good guy to take along if you want to pick up girls: he’s friendly and nice but not that attractive.” Or “Boy is that pie good.”

⁵ Since welfare economics is a subset of WC, if one rejects a tenant of WC, one rejects welfare economics, but one could reject welfare economics and still be a WC.

modern welfare-consequentialists reject the assumptions of welfare economics (e.g., the Canadian philosopher [L.W. Sumner](#) and the Oxford philosopher [John Broome](#)).

After laying out these distinctions, I return to the relationship between welfare economics and utilitarianism. Benthamite utilitarianism was the original manifestation of WC. But it is inconsistent with welfare economics, and it is neither the face of modern WC nor modern utilitarianism.



Figure 1, Chapter 2:The rape of Lucretia (Titan)

Like economic theories of choice, there is no one set of assumptions that all welfare economists agree to. But they are all *consequentialists*: whether an action is right or wrong depends only on its actual consequences, the actor's intent is immaterial, and so is what was expected.⁶ Consider the behavior of Harvey Weinstein. Consider the [rape of Lucretia](#) by Tarquin at the end of the 6th-Century B.C. (Tarquin was the son of Rome's last king). Lucretia killed herself; Shakespeare wrote a [poem about it](#); the rape and suicide were memorialized by Titan (1571) and other Renaissance artists.

The citizens of Rome found Tarquin's behavior so offensive they overthrew the king and founded the Roman Empire; so, the overall effect was, for Romans, WB increasing—but the result of a rape. Suppose the past actions of Harvey Weinstein lead to less sexual harassment and fewer rapes (we will see). In that case, a welfare consequentialist could/would argue that his behavior increased overall WB (so were WC-right), even though his motives were objectional, and his victims suffered. For welfare consequentialists, the same act can be right in some circumstances and wrong in others: it is determined by the act's consequences. When one thinks of modern welfare consequentialists, additional names include the U. of Mass. philosopher [Fred Feldman](#) and the Yale philosopher [Shelley Kagan](#).⁷

⁶ At the other extreme are *processists* who judge on only the process that caused the act (e.g., a majority vote), not the consequences. For a summary of consequentialism, see [Sinnott-Armstrong \(2015\)](#)

⁷ Peter Singer, while a consequentialist, isn't a welfare consequentialist. Sinnott-Armstrong (2015) identifies Broome, Feldman, and Kagan as welfare consequentialists; they might disagree.

These rape examples indicate that it can be difficult for a welfare consequentialist to determine right from wrong since the long-run consequences are unknown when the act occurs. Even if one waits and sees how the world plays out, there is still the issue of what can be attributed to the act.

People are not consequentialists; while they care about consequences, they also care about what caused them (the process), so they care about intentions, virtues, duties, commitments, and rights. Famous economist ethicists who have articulated theories of right and wrong who are neither welfare economists nor consequentialists include John Rawls (1921-2002)⁸ and Amartya Sen.⁹

Since welfare economists care only about WB, the only consequences that concern them are how the act affects the WB of those individuals who matter—the only consequences that matter are the welfare consequences—so they are welfare consequentialists. Recollect *Assumption 9a: An economicus's ordering of paths is based on its WB (well-being), WB, accounting for the uncertainties. The more WB economicus associates with a path, the higher its rank. And WB is determined, in part, by sensations.*

And the discussion of WB after Assumptions 10 and 11 were added.

Before proceeding, again distinguish between an ordering based on WB and an ordering based on wants/desires. Consider the alternative assumption: *Assumption 9b: Paths that are more desired/wanted are ranked higher.* The two assumptions generate the same ordering only when (1) paths that are more desired are associated with more WB, and (2) paths associated with more WB are more desired. Whether either, or both, approximate reality is an empirical question, but welfare economists assume, often implicitly, that they hold simultaneously (wants/desires ↔ WB). I point out the distinction here because while welfare economics' objective is WB, other ethicists specify fulfilling wants and desires as the ethical objective rather than increasing WB. *Peter Singer* (1997 and 2011) is in this camp. I label this camp *want/desire consequentialists*. They are neither welfare economists nor welfare consequentialists.

Table 1 makes the distinction clearer:

⁸ For introductions to Rawls, see [Wolff](#) (2010) and [Wenar](#) (2017).

⁹ It is common to lump together all economists who judge policies and call them welfare economists, but this is incorrect.

	Welfare consequentialist: nothing by WB effects matter	WB matters, but WB effects are not the sole determinant	WB effects play no role in determining
Wants and desires play no role in determining	I:	II:	III:
Wants and desires matter, but they are not the sole determinant	IV: Empty (if only WB effects matter, wants and desires can't unless they are always in sync with WB)	V:	VI.
Want/desire consequentialist: nothing but wants and desires matter	VII: want/desires ↔WB (the two orderings are identical)	VIII: Empty (if wants and desires matter, WB can't unless it is always in sync with wants and desires)	IX:

Table 1, Chapter 2: Different ethics in terms of whether want/desire fulfillment or WB are determinants of right from wrong

Welfare consequentialists are in I or VII. Want/desire consequentialists in VII or IX. In VII, want/desires↔WB. Welfare economists are a subset of I or VII. III includes Kantian ethics and the will of God. Category III ethics are alien to welfare economists, so difficult for them to get their heads around. Ethicists who care about process and WB are in either II or V; many non-philosophers are in II or V, and religious conservatives are in III.

Two properties shared by all welfare economists, but few other welfare consequentialists, are (a) in the choice period, the individual's ordering of paths is exogenous, and (b) the individual must experience their highest-ranked available path. These standard-NCT assumptions (Assumptions 1-7) take a component of ethics off the table: how to lead an ethical life. Economicus can't choose to behave more ethically.¹⁰ If my highest-ranked available path, HRAP, includes immoral behavior, I am immoral because I'm constrained to be immoral. In this sense, welfare economics is an unusual ethic: making me live a more ethical life can only be

¹⁰ If one defines an ethic solely as personal rules for living more ethically (being virtuous, having empathy, respecting authority, etc.) welfare economics isn't an ethic. Looking ahead, most welfare consequentialists embrace the possibility that you can choose to act differently than you do.

achieved by the State or other exogenous agent forcing me to. They can cause me to act less immorally by constraining or rewarding me to produce fewer negative *external effects*. [I produce an *external effect* when my behavior directly affects another individual. External effects can be positive or negative. Second-hand smoke from my cigarette is an example of a negative external effect; my beautiful garden likely produces positive external effects.]

Welfare economists typically assume (without thinking much) that only humans count but don't have a criterion for determining which humans matter/count. They don't justify or

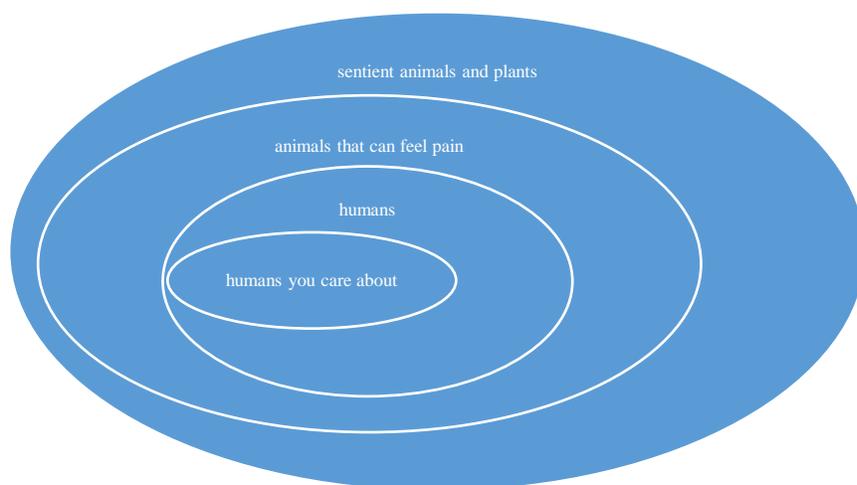


Figure 2, Chapter 2: Individuals--who counts?

explain why only humans matter. This contrasts with other welfare-based ethics, which assert and defend criteria such as *all individuals who can feel pain should count*, or *all sentient individuals should count*. When applying welfare economics, welfare economists typically accept their employer's definition of who matters.¹¹ Not surprisingly, what is right or wrong in WC comes down to who counts. A huge issue is whether future humans count and, if so, how much relative to us. The issue of future humans is further complicated because current humans affect the number of future humans. Within WC, welfare economics is restrictive regarding who counts.¹²

¹¹ For years I consulted for NOAA and the U.S. Justice Department estimating the dollar damages associated with environmental injuries resulting from oil spills, mine tailing, and toxic wastes. This makes me, and the many economists I worked with, prime examples of welfare economists. Damages to non-citizens, and damages to future citizens are not included in the assessment of U.S. damages, nor are damages to other species. If I were valuing the damages from emissions in Paris, I would have accepted a dictate to only estimate damages to French citizens residing in Paris.

¹² One could be a welfare economist (an unusual one) and include everything that can feel pain, or everyone sentient. I lean towards everyone who can feel pain.

Welfare economists agree that the only relevant consequence is how society's WB is affected. And this is only a function of the WB of each member of society. Let's call the WB of society *societal WB*.¹³ Actions that increase societal WB are right, those that decrease it are wrong, and those that increase societal WB more are righter (morally preferred). Welfare economists impose the P.I. criterion on the social-welfare function.^{14,15}

Think what WC implies about a voluntary trade: if the WB of the seller and buyer are both increased by the trade, and if the production and consumption of what is traded decrease no other individuals' WB, the trade is, by assumption, necessarily right. Moreover, this trade could even be right if it makes third parties worse if they are not made too worse.

A problem for welfare economics, and all other ethics where WB matters are that most actions increase some individuals' WB while decreasing others' WB, so are neither P.I.s nor *Pareto Deteriorations*. So, determining how an act affects societal WB requires that the analyst determine how the action affects every individual's WB. And she has a method for comparing WB losses and gains across members.

Above I mentioned rights (as in the "right to free speech" or the right to vote). Many people believe in inalienable rights. And if you have the right to do something, it is wrong (according to *Rights Theory*) for me to interfere with your ability to exercise that right (your right is a constraint on everyone else). Welfare economists do not recognize inalienable rights; they would, for example, talk about the right amount of free speech in the same way they would about the right amount of pollution. Welfare economists who advocate for property rights do so only because they believe that having and enforcing property rights will increase societal WB, not because they believe in an inalienable right to own stuff.

Honest welfare economists agree that increasing efficiency is neither necessary nor sufficient to increase societal WB. An act is efficiency increasing if the gain to the gainers, in money terms, is large enough so that the gainers could compensate any losers for their loss and still be gainers. [The compensation does not have to occur for the act to be efficiency increasing.]

¹³ I chose the adjective "societal" rather than the "social" to make clear I am not referring to the specific type of WB one might get from socializing with others.

¹⁴ Interestingly, any policy can be made a P.I. simply by banning the losers from the roles of society—one can't argue that who counts, and who does not, isn't an ethical judgement.

¹⁵ [Vilfredo Pareto](#) (1848-1923) in his later years, frustrated with how mathematical economics theories failed in the real world, turned to sociology. He came to believe that men act non-logically "but they make believe they are acting logically."—he could have been a 21st Century psychologist.

Said another way, an increase in net real wealth (in \$) does not imply societal WB has increased;¹⁶ money income isn't an accurate measure of WB. This is unfortunate since money is easy to count. An example demonstrates that an efficiency increase can decrease societal WB. Donald, a rich guy, would be willing to pay you \$1000 for a ring you inherited from your mother. You treasure the ring but are poor, so you would sell it for \$200. So, selling it to him for more than \$200 but less than \$1000 would make both of you better off. And, if no one else would have been affected by the transfer, societal WB and efficiency would have increased. However, Donald steals it, saving himself \$200. The stealing is efficiency increasing but does not necessarily lead to an increase in societal WB. For Donald, it is an additional trinket; for you, a treasure was lost (your WB likely went down more than his increased, causing societal WB to decline—if that comparison has meaning). While welfare economists agree on all this, for better or worse—and trying to be relevant—it is common for talking heads (sometimes those same welfare economists) to misleadingly suggest that more efficient means more societal WB.¹⁷¹⁸

A few ethicists argue that increased wealth (increased efficiency), not increased WB, should be the criterion to parse right from wrong—the Federal judge and legal scholar [Richard Posner](#) is in this camp. But these ethicists are not welfare consequentialists; they are *wealth consequentialists*.¹⁹ Few academic normative economists are wealth consequentialists.

¹⁶ Measuring wealth is, itself difficult. To correctly measure a society's aggregate wealth one must include many things besides the present value of its GDP stream. For example, increase in pollution or a decrease in a county's natural resource stock both affect its wealth, but neither is included in GDP.

¹⁷ [Hotelling](#) (1938) argued for using efficiency increasing alone as a criterion for increasing societal WB: if society invoked a sequence of efficiency increasing policies (ignoring who won and who lost at each step), eventually everyone would end up with more WB. The merit of this argument depends on the questionable assumption that the sequence of efficiency increasing policies has the property that each policy is random with regards to who wins and who loses. The argument also requires the assumption that the policies tend to be of equal WB-increasing magnitudes. (Hotelling was clear that one specific efficiency increasing policy would not necessarily increase societal WB.) For details see Hotelling (1938), Hicks (1941), [Little](#) (1950), and [Scitovsky](#) (1951).

¹⁸ That said, a lot of economists run around saying "more efficient is better than less efficient". Quoting Broome (1988): "*They [welfare economists] have slipped from identifying efficient states, which can be done without interpersonal comparisons, to claiming that efficient states are better than others, which generally one cannot. Robbins was right about this.*"

¹⁹ Posner (1983) argues that, while wealth and welfare are not the same thing, they are related, and that the goal of increased wealth, as compared to increased welfare, provides an ethical justification for honesty, commitment, etc., because these characteristics decrease the cost of market transactions. (A welfare consequentialist would advocate dishonesty if it would increase societal WB.). Posner's view mimics the Keynesian view that prosperity is better than depression—just because.

So much for what welfare economists agree onⁱⁱⁱ

Some welfare economists assume, but many don't, that societal WB should be the simple sum of the WB of each member of society. I will refer to this simple sum as *aggregate WB*, so when I refer to aggregate WB, I mean the simple sum, and when I refer to societal WB, I mean the more comprehensive concept.²⁰ Aggregate WB is invariant to how WB is distributed across society's members; it is impartial to which member experienced it.²¹ If maximizing aggregate WB is the moral objective, individuals, as individuals, are unimportant—they are simply containers that generate WB. And whether society consists of many containers or only one big one is immaterial.²² Many welfare economists are unwilling to go this far, wanting, for example, to allow the possibility that the WB of some members should count more, for example, citizens more than foreigners or those at the bottom of the WB ladder. Other welfare economists believe that economists have no expertise in determining a fair distribution of WB,²³ and, as discussed next, many believe that WB can't be compared across individuals.

Welfare economists don't all agree on what WB means or how it should be measured.²⁴ Some assume WB is simply emotional WB/happiness, some that it's life-satisfaction WB, some a combination. Some, like me, add pleasurable sensations.

Welfare economists differ on whether WB is cardinally meaningful (e.g., WB=4 is twice as much as WB=2) or only ordinally meaningful (4 is more than 2, but degrees of more are not comprehended).

²⁰ An aside: many welfare consequentialists argue that the least cost way to increase societal WB is to decrease suffering, a different, and appealing, perspective.

²¹ Consider whether you would choose the maximization of aggregate WB as the ethical criterion under two different scenarios: (1) you are an *impartial spectator* choosing whether a group should adopt this objective, or (2) you know you will be a member, so are not an impartial spectator, but do not know your specific position in the group. The second is choosing behind a *veil of ignorance*. Hume and Adam Smith considered the impartial spectator; Rawls and Harsanyi considered the veil. Harsanyi (1977 and 82) argues that an individual whose objective is to maximize his expected WB would, behind the veil, choose maximizing average WB (maximizing aggregate WB). **XX** does not buy his reasoning. Rawls (1971), making different assumptions, reaches a different conclusion. That an impartial spectator might choose maximizing aggregate WB.

²² Increasing aggregate WB does not necessarily increase societal WB unless everyone's WB has equal weight. In which case, societal WB is aggregate WB.

²³ Non-expertise is commonly stated in introductory microeconomics. For example, "Exactly how far policy makers should go in promoting equity over efficiency is a difficult question that goes to the heart of the political process. As such, it not a question economists can answer" (Krugman and Wells 2014).

²⁴ Chapter 4 extends the discussion and review of the different ways WB can be defined.

They also disagree about whether the overall levels of WB are comparable across individuals (measurable on the same scale).²⁵ Incomparable across individuals is a common assumption. Consider four camps of welfare consequentialists: ordinal/comparable, ordinal/incomparable, cardinal/comparable, and cardinal/incomparable.

	WB is comparable across individuals	WB isn't comparable across individuals.
For an individual, different levels of WB are cardinally meaningful	Cardinal/comparable	Cardinal/incomparable
are only ordinal	Ordinal/comparable	Ordinal/incomparable

Table 2, Chapter 2: Cardinal, ordinal, comparable, incomparable

It was challenging to get my head around ordinal/comparable, easiest for cardinal/comparable. If everyone's WB is a number on the same cardinal scale (cardinal/comparable), aggregate WB has cardinal meaning.

As noted in Chapter 1, whether WB has cardinal meaning is an unanswered empirical question. People think they comprehend and rank changes in WB, but that does not make it so.

Whether it is possible to compare WB across individuals is also an empirical question. People assert such comparisons. "I am not as happy as George but am happier than Glen, who is always depressed." Or "Donald stealing my mother's ring increased his happiness less than it decreased mine." While, at the same time, admitting that we can't be sure about such comparisons. There are two issues: whether it is meaningful to compare your WB to mine and, if so, how to do it.

If WB is ordinal/non-comparable, adding WB numbers across individuals makes no sense. However, a welfare consequentialist can still ask whether a rule could exist that would take everyone's ordering of paths and spit out a societal WB ordering of sets of paths. Ideally, a rule that worked for all possible individual orderings. They would want to impose ethical restrictions on the rule. For example, if every member of society ranks a *Set X* higher than *Set Y*, the rule must rank *Set X* higher in terms of societal WB.²⁶ [Kenneth Arrow](#) (1961/63) proved that if this and a few other reasonable ones are imposed, no rule can exist that would work for all possible individual orderings (this is Arrow's famous impossibility theorem). Put simply, if each

²⁵ Whether WB is comparable across individuals is different from whether bearers-of-WB are comparable.

²⁶ *Set X* is the set of paths taken by each member of society in scenario X.

of us only has an ordering of paths, and if these orderings are WB-incomparable across individuals, for many individual orderings, it will be impossible to order sets of paths in terms of societal WB.²⁷

I was surprised when I learned that rules exist for ordinal/comparable. Consider, for example, the following rule from Rawls (1971): *Set X* is associated with more societal WB than *Set Y* if, and only if, the individual with the least amount of WB in *Set X* has more WB than the individual with the least amount of WB in *Set Y*. (They don't have to be the same individual.²⁸) *Set X* is morally preferred to *Set Y*, only if the worst off is better off with set X. Attractively, this rule is inconsistent with only one individual's ordering being the sole determinant of the societal ordering. It also has the desirable property that how it ranks *X* vs. *Y* does not depend on what other paths are available: it is consistent with the *independence of irrelevant alternatives restriction*. Despite these desirable properties, The Donald, and many others, would not like this rule.

Welfare consequentialism (WC), including welfare economics, scares many moral philosophers because anything goes if it increases WB^{iv}

In WC, no act nor behavior is inherently right or wrong. Lying isn't inherently wrong; it is morally right if it increases societal WB, same for breaking promises and dereliction of duties. There is no inherent right to free speech nor equal treatment under the law. There's no proscription on racial, cultural, or sexual discrimination. There is no right to liberty, so there is no limit on the amount the government can interfere in your life. Since speaking freely increases

²⁷ Ordinal/incomparable, a main theoretical mantra of much of 20th Century choice theory, leaves welfare economists nothing to conclude other than Pareto Improvements (Deteriorations) are right (wrong). Robbins (1938) famously pointed this out. "He maintained that if economics was to have the objectivity of a science, economists may not make interpersonal comparisons and may not, in their capacity as economists, argue for or against any policy or change of policy that would make some people better and others worse off" (Scitovsky 1951). Bucking this ordinal/incompatible trend, numerous Cambridge (e.g. Pigou 1932 and Kahn 1935) and some American economists (e.g. Irving Fisher and Frank Knight 1944) continued to advocate, sometimes implicitly, for comparability, arguing that we all have a similar capacity for WB, such that similar amounts of money generate similar amounts of comparable WB. See Scitovsky for additional details on the varying views on welfare economics in the 20th Century's first half.

²⁸ Getting technical, with ordinal/comparable WB, WB has no cardinal meaning from the individual perspective. One can achieve comparability across individuals without imposing cardinality by assuming WB functions can only vary across individuals up to the same increasing monotonic transformation. [WB_i^* is an increasing monotonic transformation of WB_i , if $WB_i^* = \phi_i(WB_i)$ such that the derivative of $\phi_i(WB_i)$ is positive. Comparability is imposed by requiring that $\phi_i = \phi$ (that it's the same function for everyone).]

the WB of those speaking, but it often decreases the WB of those who must hear it, WC requires restrictions on speech.

For welfare economists, the sources of an individual's WB are irrelevant for determining right from wrong. Consider the implications. The emotional WB the neighbor kid gets from his mother's love is worth no more, nor less, than the emotional WB he gets from torturing cats. My decrease in WB from knowing that people are starving counts for no more than the decrease in WB I experience from having neighbors who are ethnically different from me. Abhorrent to many is that the WB I get from causing others to suffer, or their misfortune, counts as much as the WB I get from helping a friend or loved one.

If the sources of an individual's WB are irrelevant for determining right from wrong—the WB counts no matter how despicable or disgusting its source. Looking ahead to research on disgust, the pleasure George gets from having sex with a dead chicken he bought at the supermarket, then roasted and enjoyed for dinner, counts. [Of course, WC must consider that the WB of others could decrease if they were to know what George was up to.²⁹]

Imagine the sources of an individual's WB are irrelevant, and the moral objective is to maximize aggregate WB. Then, any act that increases aggregate WB is right (morally preferred) no matter how the WB was generated. Start with a simple example from [Amartya Sen](#) (1970), an economic ethicist and Nobel laureate, but not a welfare consequentialist. Consider a society of two individuals, Lewd and Prude, where Prude gets WB by denying Lewd pleasure, and Lewd gets WB by making Prude do something he does not want to do.

Society has to rank three states: (A) no one reads Lady Chatterley's Lover, (B) only Prude reads it, or (C) only Lewd reads it. Prude, being a prude, does not want to read it but prefers he read it rather than Lewd: Prude hates the idea of Lewd enjoying the book. Best for Prude is no one reads the book. Lewd, on the other hand, would enjoy the book, but loves the idea of Prude being forced to read it. No one reading the book is Lewd's worst outcome.

To the discern of many, Prude being forced to read the book could increase aggregate WB: both Prude and Lewd are better off if Prude, rather than Lewd, reads it, and while Prude would have

²⁹ Mill did not specifically discuss chickens, but rejected counting external effects of this sort, providing further evidence that he was neither a utilitarian nor a welfare consequentialist. Reacting to the argument that society can limit drinking simply because it decreases security and weakens and demoralizes society. Mill angrily responded that if this is harm, then we would all, wrongly, have an interest in everyone's "moral, intellectual and physical perfection" ([Reeves](#) 2007).

more WB if no one read it, this may be more than offset by the great pleasure Lewd gets from Prude being forced to read it. Objectors, including Sen, would argue that forcing Prude to read a book hates can't be the moral high ground. Also, objectors would object to having to include both the pleasure Lewd would get from Prude being disgusted and the pleasure Prude would get from Lewd not having the pleasure of reading the book. Shifting examples, Bob's pleasure from living in a neighborhood with only straight, white people could mean that keeping others out is WC morally preferred to an integrated neighborhood.

If the sources of an individual's WB are irrelevant and the moral objective is to maximize aggregate WB, WB generated by ignorance counts as much as WB generated by learning. In WC, there is nothing inherently wrong with fake news: it increases the WB of some. There is also nothing inherently right about real news: it decreases the WB of some. The pleasure I get from believing global warming isn't real counts as much as the WB I would get from a great ski day or pulling a drowning kid from a ditch.

If one takes WC to the extreme, one can create examples that would make almost everyone scream "NO", no to awful acts that could, in theory, be justified by WC. These include torture, racial discrimination, sadistic acts, dictatorial rule, and killing humans for fun and sport.³⁰ Anything can be morally justified by WC simply by arguing that doing it would increase societal WB. One could reject the assumption that the sources of WB are irrelevant, but one would not be a welfare economist. Welfare economists include the negative WB impacts of my actions on others, but they would never deem specific sources of my WB as immoral per se.

In contrast, [John Harsanyi](#) (1920-2000), the 1994 winner of the Nobel in Economics, did. He argued that society should, on ethical grounds, exclude from WB pleasures and pains caused by lack of information, misinformation, and morally objectionable orderings (those based on racism, sexism, sadism, maliciousness)—orderings that lead people to trample on the dignity and equality of others (Harsanyi 1977 and 82).³¹ He was not a welfare economist. As a Hungarian

³⁰ Welfare consequentialists tend to argue that while despicable acts could in theory increase WB, this is unlikely in practice when one takes account of those who suffer from despicable acts. Of course, their suffering is not taken into account if the sufferers are not members of society.

³¹ An important question is who determines what orderings are immoral. And if society decides an ordering based on hate is immoral, there is the further issue of whether an ordering is, or isn't, hate-based. And, if hate is deemed immoral, how does one determine what a hate-filled person's ordering would be if he were hatred free?

Jew in 1944, he worked in a forced-labor unit, escaping from the train at the Budapest station as his unit was sent to an Austrian concentration camp (Harsanyi 1995).

Consider the implications of adding to WC the welfare economic assumption that everyone maximizes their WB subject to their constraints (there is no flawed choosing). This makes, for many, welfare economics even more objectionable. It is more objectionable because the assumption implies it is always wrong to interfere in actions that affect no one other than the actor. You and the government have a moral imperative not to intervene when I beat my head against the wall, drink 12 shots of tequila, or try to commit suicide—unless someone else would be negatively affected—a lot. In contrast, a welfare consequentialist who recognizes a choice would reduce WB would be morally compelled to intervene, even if it would not decrease anyone else's WB. We believe other people make mistakes, and we are happy to intervene.

And finally, there is the complaint that welfare economists don't require that everyone counts and everyone's WB counts equally.

Are welfare economists utilitarians? ^v

The Dictionary of Bad Definitions

Utilitarian, *n.*

Etymology < post-classical Latin *utilitarius* >

A person who studies, worships, or is otherwise obsessed with the word *utility*.

Economists must be utilitarians because they are always talking about utility. (Shakespeare's brother Bob, 1604, in *My letters to Will*)

Whether welfare economists are utilitarians needs clarification. Why? (1) Many economists are led to believe (I was) that welfare economics is grounded in the ethic called utilitarianism—this is incorrect. Welfare economists need to be clear on what they assume and believe—they need to be precise. We should not be misleading ourselves or our students. (2) Utilitarianism is objectionable to many—a dirty word—so why unnecessarily call yourself a utilitarian? [Even without the title, many would still find the ethics of welfare economics objectionable.] (3) And it is of interest to understand how the label “utilitarian” came to be associated with welfare economics: it is because Benthamite utilitarianism was the first welfare-consequentialist ethic out of the gate, so, for a long time, it was reasonable, and initially not misleading, to refer to all welfare consequentialists as utilitarians. But now, there are different sorts of welfare consequentialists, so it is misleading to label them all utilitarians when most are not.

So, what is utilitarianism, and how does it relate to welfare economics? The first difficulty with answering is few modern moral philosophers are welfare consequentialists, and the few who are, do not self-identify as utilitarians—it has a bad name in large spheres of moral philosophy.

So, I start with whether welfare economists adhere to the classic utilitarianism of [Jeremy Bentham](#) (1748-1832), the founder of utilitarianism?^{32 33} Most do not. Benthamite utilitarianism is WC if WB is only pleasure minus pain. Welfare consequentialists, including welfare economists, take a broader view of WB. Pain might cover all the things that would cause a spider's WB to suffer, but many things that decrease the WB of my dog and I are not physical pain (the firings of specific types of muscular nerve cells), and pain does not always cause a decrease in our WB. Chewing on sticks to dislodge one's baby teeth hurts, but the puppy persists. On the way to the dentist, my tongue plays with the loose tooth trying to find the amount of pushing that hurts best. After a competitive bike ride, the leg pain tells me I worked hard, my goal, and I know the pain will be gone in the morning. But pain you do not control won't go away and is a symptom of illness causes a decrease in WB. The word pleasure is also too narrow for humans. Like pain, pleasure is a sensation. Many pleasures (the taste of good chocolate, orgasm) are specific to certain body parts. A pleasure can cause enjoyment, but it can also cause suffering—I am upset with myself for eating those tasty onion rings. Or, it can go the other way: consider poor Fyodor Dostoevsky going on about “the enjoyment, of course, of despair; ... in despair there are the most intense enjoyments” ([Dostoyevsky](#) 1864:1996). Whether

³² Utilitarianism was developed and popularized by Bentham's *Introduction to the Principles of Morals and Legislation* (1789). Put simply, humans and other animals experience only pleasure and pain which can be netted and then added across individuals (Bentham's *felicity calculus*). An action is moral if it increases aggregate net happiness. The first sentence of Bentham's book says it all,

Nature has placed mankind under the governance of two sovereign masters: pain and pleasure. It is for them alone to determine what we ought to do, as well as to determine what we shall do.

³³ Initially Bentham did not call it utilitarianism. John Stuart Mill adopted the term after seeing it used in a novel, unaware that Bentham had earlier used the word in a letter (Reeves 2007). Mill suggests that Plato was the first utilitarian (Mill (1879)). "To each his due" suggests actions should be assessed on the basis of how they affect welfare. The expression "the greatest happiness of the greatest number", while popularized by Bentham, can be traced back to the philosopher [Francis Hutcheson](#), a founding father of the Scottish Enlightenment (Reeves 2007). One root of utilitarianism is [Epicurus](#). Bentham claimed inspiration from Hume. When Bentham came across Hume's discussion of morality he felt as though “scales had fallen from his eyes” (quoted by [Gottlieb](#) 2016) “It appears there was never any quality recommend as a virtue, or moral excellence, but on account of it being *useful* or *agreeable* to a man *himself*, or to *others*,” (Hume 1758). Hume differed from Bentham in that Hume was noting how people judge, not on how they should judge, and he did not think WB was quantifiable (Gottlieb 2016)

a pleasurable or painful sensation increases WB depends on the circumstances in which it occurs. Bentham was a pain and pleasure guy; welfare economists are not.

Also, in Benthamite utilitarianism, net pleasure has cardinal meaning, and everyone's pleasure is measured on the same cardinal scale (cardinal, comparable). Also, the moral objective of Benthamite utilitarianism is to increase aggregate net pleasure, defined as the sum of aggregate pleasures minus aggregate pains. Bentham advocated maximizing aggregate WB, limiting WB to one pain/pleasure continuum.

And, for Bentham, this includes the pleasure and pains of all animals that can experience either, independent of species. For Bentham, the pleasures and pains of ducks count as much as yours and mine, and all count equally, including spider pain.

The day may come when the rest of the animal creation may acquire those rights which never could have been withholden from them but by the hand of tyranny. The French have already discovered that the blackness of the skin is no reason a human being should be abandoned without redress to the caprice of a tormentor. It may one day come to be recognized that the number of the legs, the villosity of the skin, or the termination of the os sacrum [whether one has a tail] are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason or perhaps the faculty of discourse? But a full-grown horse or dog is beyond comparison a more rational, as well as a more conversable animal, than an infant of a day or a week or even a month, old. But suppose the case were otherwise, what would it avail? The question isn't, can they reason? nor can they talk? But, can they suffer? (Bentham 1789)

Holding all else constant, imagine two worlds: in the first, I get 10 units of net pleasure, and Wilbur, a pig, gets -5, and in the other, the numbers are switched. The Benthamite utilitarian is indifferent/impartial: who gains and loses is immaterial if the net gain is unchanged. This *impartiality* is a critical component of Benthamite utilitarianism (Driver 2014), but not welfare economics. Welfare economists typically do not require that members' WB get equal weight. The objective of Benthamite utilitarianism is to maximize aggregate WB, but many welfare economists don't assume societal WB is aggregate emotional WB, a critical difference.

Benthamite utilitarianism would judge a behavior as moral or immoral by whether it increases or decreases aggregate WB in his pleasure/pain sense, and a larger increase in aggregate WB is preferred to a lesser increase. Bentham would judge the singer [Lana del Rey](#) smoking another cigarette wrong if her pleasure is less than the displeasure produced by her second-hand smoke. Welfare economists, in contrast, would ask whether the act increases societal WB, defining WB more broadly, not necessarily counting Lana's, and everyone else's WB, equally. Neither welfare economics nor Benthamite utilitarianism implies your act is moral simply because it increases your WB. As noted earlier, since economicus is constrained to

maximize his own WB, economicus is constrained to act Bentham- immorally if an act increases his personal WB but decreases aggregate WB. Economici are commonly referred to as *psychological egoists*, or *egoists*—which is true if an egoist is defined as someone who cannot act contrary to their self-interest. Whether Bentham believed humans are egotistical is a matter of debate: maybe he left open the possibility that humans can choose to act contrary to their own WB. If so, Benthamite Lana might simply decide to not light up if her smoking would reduce aggregate WB.³⁴ To make Economicus Lana’s behavior moral, someone would have to interfere.

In Table 1 in Chapter 2, Benthamite utilitarianism is necessarily a subset of Sets I or VII and a strict subset of VII if Bentham believed everyone maximizes their own WB. If Bentham didn’t believe this, Benthamite utilitarianism is a strict subset of I. If it is a strict subset of VII, it is a different subset of VII than the welfare-economics subset, with no or minimal overlap.³⁵

The first place to look past Benthamite utilitarianism is the supposed utilitarianism of John Stuart Mill, but I will skip over Mill for now: he is complicated.³⁶ I will also jump over the influential utilitarian [Henry Sidgwick](#) (1838-1900) to the 20th century.³⁷ Simply put, there is no

³⁴ Bentham is contradictory on this issue: sometimes saying we are egoists, but sometimes giving us the ability to act contrary to our self-interest. In *The Book of Fallacies*, Bentham says, “In every human breast, rare and short-lived ebullitions, the result of some extraordinary strong stimulus or incitement excepted, self-regarding interest is predominant over social interest: each person’s own individual interest, over the interests of all other person taken together” (Bentham 1824 :392-3). Except for the insert, “rare and short-lived ebullitions, the result of some extraordinary strong stimulus or incitement excepted,” (which is often omitted replaced with “...”). This sounds like an endorsement of egoism. The insert leaves the door open a bit. But, in contrast, in *An Introduction to Principles of Morals and Legislation* he says, “There is no case in which a private man **ought** not [emphasis added] to aim to produce his own happiness and of that of his fellow-creatures; ... Every act that promises to be beneficial on the whole to the community (himself included) each individual ought to perform of himself; ... Every act that promises to be pernicious on the whole to the community (himself included) each individual ought to abstain from;...” (Bentham 1996, Chap 17, paragraph 8). There are differing interpretations as to what Bentham meant (see, e.g. Driver (2009:9-10), Dinwiddy (2004: 137-8) and Lyons (1991)

³⁵ Looking ahead, to get overlap, welfare economists must assume that only emotional WB matters, that the emotional WB of all animals matters equally, and that WB is cardinally meaningful.

³⁶ In partial explanation, Mill complicated the simplicity of Benthamite utilitarianism by arguing that higher pleasures should count more than lower ones (sex, drugs, and rock’n’roll). See, *Utilitarianism* (Mill 1867), and Reeves (2007 biography *John Stuart Mill: Victorian firebrand*. The deceased Bentham likely rolled over at Mill’s notion of utilitarianism. In 1879, Mill’s book was perceived by many as the voice of utilitarianism, but [Stanley Jevons](#) (1879) argued that if one drops the assumption that all feelings fall on the same pleasure/pain univariate measure, it is no longer utilitarianism. Quoting Stanley Jevons (1879) critique of *Utilitarianism*,

Nothing can be more plain, too, than that Mill himself believed he was dutifully expounding the doctrines of his father, [and] of his father’s friend, the great Bentham...His Essays purport throughout to be a defense and exposition of the Utilitarian doctrine.... but there is a wide gulf between what he intends and what he achieves... I [Jevons] make it my business therefore in this article to show that Mill was intellectually unfitted to decide what was utilitarian, and what was not.

³⁷ Sidgwick wrote the influential *The Method of Ethics* (1874) which was both a defense and explanation of utilitarianism (Driver (2014)). He critically pointed out the problem of total emotional WB versus average emotional WB: suggesting that the goal should be to maximize the product of average emotional WB and population size.

one modern utilitarian ethic but several overlapping ethics. For example, *act utilitarianism*, *rule utilitarianism*, and *preference utilitarianism*. Act utilitarianism, like Benthamite utilitarianism, makes the objective maximizing aggregate WB—act by act. Whereas the objective of rule utilitarianism is to identify rules—rules that apply to everyone—that will, overall, lead to the maximization of aggregate WB.³⁸ The different varieties agree that WB isn't simply pleasures minus pain. And none are Benthamite utilitarianism.

Interesting is preference utilitarianism. Peter Singer states, “my ethical position is a form of preference-utilitarianism:” “I approach each issue by seeking the solution that has the best consequences for all affected. By ‘consequences’, I understand that which satisfies the most preferences, weighted in accordance with the strength of the preferences” (Singer 1997). A noted difference between Singer and Benthamite utilitarianism is that for Singer, the objective is to fulfill an individual's *interests*, achieved by giving the individual what they prefer, even if it causes them more pain and less pleasure, or even less emotional WB. Singer defines preferences as “wants, needs, and desires” (Singer 2011). He does not advocate for maximizing aggregate WB, so while he is a consequentialist, he is a *want/desire consequentialist*. Singer does not believe fulfilling wants and desires always increases WB. His advocacy for fulfilling wants and desires, rather than increasing WB, is that wants and desires are easier to observe and measure. He counts the wants and desires of all species equally, giving all animals equal standing.³⁹ Singer defines “basic equality” as when everyone's interests get equal consideration (impartiality) and argues that equal consideration of interests is morally preferred over unequal consideration of interests. Singer (2011) describes his preference utilitarianism as an extension of the utilitarian views of his Oxford professor [R.M. Hare](#) (1919-2002).⁴⁰

In 1977, Harsanyi (1977 and 82) proposed preference utilitarianism. He pointed out (as many psychological studies confirm) that individuals don't always maximize, subject to their constraints, their WB: he argued individuals violate the assumptions of NCT based on misbeliefs, emotions, etc. He distinguishes between what he calls *true preferences* (an ordering

³⁸ One argument for rule over act utilitarianism is that act utilitarianism provides no justification for moral rights or obligations—anything goes if it increases aggregate WB (Harsanyi 1977 and 82, Sec. 9). Whereas the general rules produced by rule utilitarianism provides guidelines (rights and obligations) for behavior. So, for example, rule utilitarianism likely warrants the banning of slavery, while in act utilitarianism whether enslaving someone is moral or immoral is determined case by case. The other difference is that an individual undertakes an act taking as given the behavior of others, whereas a chosen rule applies to everyone, so effects the behavior of others.

³⁹ Singer is a founder of the animal-rights movement—*Animal Liberation* (1975, 4th ed 2009).

⁴⁰ For an introduction to Hare, see Price (2016). Of interest is also *Early Utilitarians* (Binmore 2021)

based on personal WB) and *manifest preferences* (an ordering consistent with behavior). Harsanyi's moral objective is to maximize WB based on true preferences—but only after the deletion of WB caused by morally objectionable acts.⁴¹

The preference utilitarianisms of Singer, Hare (1981), and Harsanyi all have a rules component. They argue that society should determine what is morally right based on rules that, if followed, would over time lead to the satisfaction of more wants and desires. But, in complicated cases, the rules will not suffice, and critical thinking will be required. They all agree that preferences based on illogical or bad information should not count. While Harsanyi excludes the fulfillment of immoral wants and desires (hate, sadism, etc.), I have not been able to find anything by Singer or Hare that indicates they agree. Singer emphasizes the equal consideration of everyone's interests, and his textbook, *Practical Ethics* (2011), says nothing about inappropriate interests.

Contrasting with preference utilitarianism, welfare economists are only concerned with welfare/WB, not want and desire fulfillment per se. And preference utilitarians like Singer include the preferences of non-humans. And many welfare economists reject the notion that preferences and WB have intensities that can be compared across individuals. Another difference is that preference utilitarianism gives everyone equal weight, as does Benthamite utilitarianism, while welfare economics does not require this.

Welfare economics does not distinguish between true and manifest WB: self-assessed WB and WB are the same?⁴² Preference utilitarianism explicitly admits the possibility that you don't know what is best for you. One can be a preference utilitarian without assuming everyone experiences their HRAP, whereas NCT and welfare economics assume the individual always experiences their HRAP. Preference utilitarians admit flawed choosing in terms of WB. If they believe someone makes a flawed choice, they will advocate for paternalism (limiting and directing the individual's behavior against their wishes). For example, imagine that your teenage daughter is inclined to cut herself (not uncommon). Welfare economics says let her cut unless the cutting negatively affects others. But a preference utilitarian would advocate sending her to therapy and throwing away the knives—even if her cutting affects only her. Or imagine that I,

⁴¹ It isn't clear whether such subtractions are possible. How would one determine what a malicious racist's ordering would be if they were neither a racist nor malicious?

⁴² Self-assessed WB is sometimes called *subjective WB*: from the subject's perspective.

when given the opportunity, poke pins in my eyes and fail to save for retirement. A preference utilitarian might stop the poking and forced saving. But if the poking does not affect others, the welfare economist has no grounds for interfering. (In determining whether to restrict my behavior, a preference utilitarian should account for the fact that forcing me to save for retirement conflicts with my desiring freedom.)

Some economic ethicists agree with Singer that the moral objective is satisfying more wants and desires rather than increasing WB. As I characterized Singer, they are want/desire consequentialists. In Table 1: want/desire consequentialists are restricted to VII and IX. If they admit choices that violate preferences, as most do, they are in Set IX. Hausman (2010) critiques the divide between WC and want/desire consequentialism.

I end this comparison of utilitarianism and welfare economics with what each implies about redistributing wealth. A moral imperative of Singer's preference utilitarianism is that once the basic needs of you and your loved ones have been met, you should allocate all remaining time and income to reduce the suffering of those individuals who are suffering the most, irrespective of their relationship to you, or their species—to do otherwise is wrong. This conclusion follows from his assuming (1) preferences have intensity, (2) preferences can be compared across individuals (even between spiders and humans), (3) a dollar spent will fulfill the most wants and desires if it's spent helping those who are currently suffering the most,⁴³ (4) one can choose to not pursue their own interests, and (5) the moral objective is to increase aggregate want and desire fulfillment. In his textbook *Practical Ethics*, Singer admits he is asking a lot. His examples deal with our moral imperative to help desperately poor people. But his assumption that everyone should count equally, independent of species, implies a moral imperative to reduce animal suffering. Reducing the suffering of food animals could be cheaper than serving the interests of desperate humans).

Welfare economists would not advocate Singer's moral imperative because they would not, as we have discussed, accept all his assumptions. For one, they would reject (4): economicus can't do this; he must maximize his personal WB (or desire fulfillment) subject to his constraints. So, for the welfare economist, behaving more morally must be imposed on the individual.

⁴³ This assumption is akin to what economics defines as *diminishing marginal utility*: something we teach as a law to students in principles of microeconomics.

If a welfare consequentialist, or welfare economist, adopts (3) and the moral objective is to maximize aggregate WB, they would conclude that it is right (moral) that the State redistributes from the well-off to the worst off.⁴⁴ But many reject (3).⁴⁵ Given all these differences, welfare economists are not preference-utilitarians.

Welfare economists are also not rule-utilitarians: welfare economists typically think about whether a specific act or policy would increase aggregate or societal WB rather than suggesting ethical rules that would, on average, increase aggregate or societal WB.⁴⁶

Besides not being utilitarians (Benthamite, rule, or preference), welfare economists differ from other modern welfare consequentialists^{vi}

Some modern welfare consequentialists reject the notion that an individual can WB-order all possible paths, arguing that there are different kinds of WB and that not all are commensurable. One could believe that only welfare consequences matter, but it is impossible to compare all the different types. But, as noted in Chapter 1, the commensurability of WB is the foundation of NCT and welfare economics—pervasive non-commensurability would be a blow to both.⁴⁷

In contrast, many welfare consequentialists who believe all paths can be WB-ordered reject the additional assumption that individuals always maximize their personal WB. They reject it either because they believe in mistakes or believe a few individuals choose their path to maximize aggregate WB. [If individuals can do the latter, how to live an ethical life is a component of WC.]

Some modern welfare consequentialists (e.g., a former chief economist of the World Bank, [Kaushik Basu](#)) argue that WB isn't only emotional and life-satisfaction WB. Instead, they assert that WB also increases with knowledge, the extent to which one is fulfilling their

⁴⁴ Everyone would account for the fact that progressive tax rates could influence the incentive to work.

⁴⁵ Bentham also advocated for the redistribution of wealth from the best off to the worst off. While Bentham did not articulate diminishing marginal utility in the full modern sense, he did feel that an additional dollar spent by someone well off would increase net pleasure less than if were spent by a worst off, and Bentham's goal was to maximize aggregate pleasure minus pain. Quoting Bentham, "the more remote from equality are the shares [of wealth] possessed by the individuals in question, in the mass of the instruments of felicity, --the less is the sum of the felicity produced by the sum of those same shares" (Bentham 1839: 271).

⁴⁶ The aim of *social-choice theory* (List 2013) is to identify **rules** (functions) to identify what is morally preferred based on its welfare consequences. For example, there is the well-researched question of how different voting schemes (rules for decision making) succeed and fail to distinguish between right and wrong policies.

⁴⁷ NCT survives limited non-commensurability but with wounds. For example, if all the kinds of WB associated with the consumption of market goods are commensurable, but that these kinds of WB are not commensurable with those associated with non-market experiences such as religion and time with friends. Chapter 8 has more details.

capabilities, and even simply having more capabilities—even if they fail to increase emotional or life-satisfaction WB.⁴⁸

According to this view of WB, you can have more even if you don't feel it or realize it. And you could be better off even though you feel worse if, for example, the newly acquired knowledge makes you anxious and depressed.

One can, of course, be an economic ethicist who adheres to an ethic where the WB of society's members is important, but WB isn't the only thing of importance (II or V in Table 1). But then, one is neither a welfare economist nor a welfare consequentialist.⁴⁹ This is OK: no rule says economists who care about parsing right from wrong actions must only care about WB. What one is called if one cares about WB, but not only WB, depends on how much they think WB should matter relative to other things, such as being virtuous or fulfilling your duties. Consider calling said ethicists *NotOnlyWB ethicists*.

We now turn to research on how regular people judge right from wrong. If you view yourself as a welfare consequentialist or a welfare economist, it is enlightening to know that few accept your view that only welfare consequences matter. Also disturbing to an applied welfare economist like myself is many people, while concerned with their own WB, do not believe that maximizing societal WB is the moral objective or that P.I.s are always right. Summarizing by paraphrasing Aristotle, if the intellectuals who study ethics and the guy-on-the-street reject welfare economics, I, as a lifelong welfare economist, should stop and pause.

The moral judgments of regular people: they don't ascribe to the ethics of welfare economics^{vii}

People don't determine whether an act is right or wrong solely on its effect on WB. For them, the WB of almost everyone else plays no role in deciding right from wrong. Research in moral psychology indicates that typical Western research subjects judge morality (of an act or a person) using two criteria: (1): justice/fairness (treating others in proportion to their actions and intent—they should get their just desserts),⁵⁰ and (2): caring for others and protecting them from harm.

⁴⁸ The view, that WB is increasing in capabilities, even if they are unrealized, is widely associated with Sen and Nussbaum (Sen 1980, Nussbaum 2006 and 11, Basu and Lopez-Calva 2011, Robeyns 2016)

⁴⁹ One could even be economic ethicist who believes that how an act influences WB is irrelevant to determining whether the action was right or wrong. The moral imperative could be fulfilling the will of God. My friend Mike is in this category.

⁵⁰ If justice and fairness depend on intent, and not just consequence, it is inconsistent with WC.

Harming others is wrong; caring for others is right, those who harm should be punished, and those who care for others, and reduce harm, should be rewarded.

Non-western adults and religious conservatives add two additional criteria: (3) loyalty to your community/group, including respecting its authority, including fulfilling your obligations to the group, and (4) an ethic of holiness (humans are God's creations and are morally obligated to not degrade that holiness). Describing this perspective:

People are bearers of something holy or god-like and have moral obligations to not act in ways that are degrading to or incommensurate with that holiness. The ethic of divinity requires that people treat their bodies as temples, not as playgrounds, and so personal choices that seem to harm nobody else (e.g., about food, sex, and hygiene) are sometimes moralized (for a further elaboration of moral foundations, see Haidt & Graham 2007). In sum, the high-socioeconomic status (SES), secular Western populations that have been the primary target of study thus far, appear unusual in a global context, based on their peculiarly narrow reliance, relative to the rest of humanity, on a single foundation for moral reasoning (based on justice, individual rights, and the avoidance of harm to others; cf. Haidt & Graham 2007). (Henrich, Heinz, and Norenzayan 2010)

One way to observe and measure moral diversity is to present research subjects with descriptions of different behaviors and ask how they judge them (right or wrong?). Judgment varies with the respondent's socio-economic status, political orientation (liberal vs. conservative), and religiousness. Researchers also ask what criteria the respondent considered when judging an act right or wrong. The research objective is to determine whether the criteria vary across individuals.

[Jonathan Haidt](#) (2012), his co-authors, and others have tested respondents' reactions to *harmless acts* (acts that cause no physical or psychological damage, except possible offense—acts that occur in private). Three examples (Haidt, [Koller](#), and Dias 2003):

A woman is cleaning out her closet, and she finds her old [American or Brazilian] flag. She does not want the flag anymore, so she cuts it up into pieces and uses the rags to clean her bathroom.

A man goes to the supermarket once a week and buys a dead chicken. But before cooking the chicken, he has sexual intercourse with it. Then he cooks it and eats it.

A brother and sister like to kiss each other on the mouth. When nobody is around, they find a secret hiding place and kiss each other on the mouth passionately.

Surveying in three cities (two in Brazil, plus Philadelphia), the authors asked about such activities. They found that high socioeconomic Americans view such acts as causing no harm and are not immoral—just activities that might go against social convention. In contrast, the poor and the religious (more so in Brazil) find such acts immoral simply because they invoke disgust or disrespect. Moreover, western democracies typically legislate to reduce harm, not to eliminate

harmless acts. [There are exceptions, such as laws against homosexual acts between consenting adults.] In contrast, religious states such as Iran and the Vatican ban many harmless acts.

Graham, Haidt, and Nosek (2009) identify five moral intuitions: harm/care, fairness/reciprocity, in-group/loyalty, authority/respect, and purity/sanctity. Conservatives give equal weight to each criterion, whereas liberals use only the first two. In their 4th Study, respondents were asked what they would have to be paid to voluntarily do different sorts of distasteful things. The following are their examples for each criterion:

(Harm) Kick a dog in the head, hard

(Fairness) Say no to a friend's request to help him move into a new apartment after he helped you move the month before.

(In-group) Burn your country's flag (nobody else sees you do it)

(Authority) Make a disrespectful hand gesture to your boss, teacher, or professor.

(Purity) Get a blood transfusion of 1 pint of disease-free blood...from a convicted child molester.

Their Figure 24 shows the results. (Each number on the vertical scale corresponds to much larger dollar amounts than the previous number.)

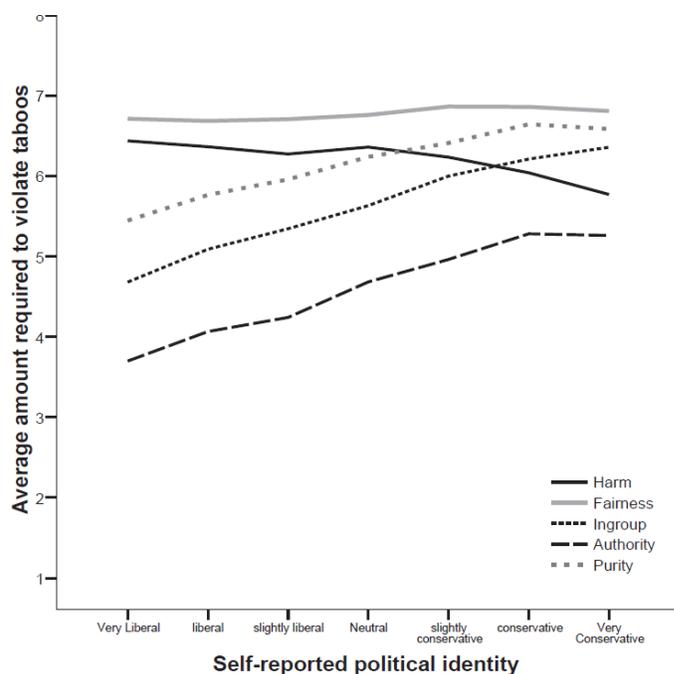


Figure 3, Chapter 2: from Graham, Haidt, and Nosek (2009)

Everyone needed the largest amount of money to be unfair, and the amount did not vary by political orientation (the top solid line). The more conservative the respondent, the less money needed to harm. And, when it comes to violating your group, authority, or purity, the more conservative, the more money required.

It is difficult to reconcile these ethical criteria with WC, especially if each gets equal weight. Criterion 2: *caring for others and protecting them from harm* is the only one with an obvious link to societal WB. Following Criterion 1: *justice/fairness (treating others in proportion to their intentions and actions—they should get their just deserts)* can decrease societal WB. I should care for others only if their behavior indicates they deserve care. None of the criteria explicitly say that increasing the WB of yourself and others is morally preferred.

Trying to reconcile the four criteria to WC and welfare economics, one might awkwardly argue that observing what the observer regards as disloyalty, disrespect, or injustice matters to the observer only because it reduces the observer's WB. And caring for others, who need and deserve care, is right only to the extent it increases the WB of the recipient and the caregiver. Put simply, the argument would be justice, disloyalty, etc. are only of moral relevance because they affect WB. Many people would disagree with this explanation. Consider the football players who knelt in protest when the U.S. National anthem was played. Did they do it only because it increased their WB, and did others only object because the kneeling decreased their WB? —both groups would probably say no.

A welfare consequentialist who isn't a welfare economist (e.g., Harsanyi) might, alternatively, try to reconcile the four criteria fully or partially with WC by arguing that WB obtained by being disloyal, disrespectful, or unjust should not count in the determination of societal WB. A welfare economist would shudder at this suggestion.

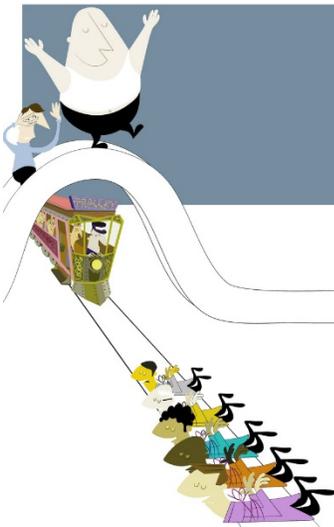
The introduction-to-microeconomics text I used, [Krugman and Wells \(2016\)](#), concludes that a P.I. is a right act. And, given a societal goal, achieving it at a lower cost is morally preferred over achieving it at a higher cost. These are two pillars of welfare economics but are not in sync with how many of us judge specific policies and behaviors. The four criteria imply an individual will conclude a P.I. is wrong if they view the gainers as undeserving (e.g., lazy) and conclude a Pareto Deterioration right if the losers deserve punishment. The same justice issues pertain to whether society should strive to minimize the costs of achieving a goal; are the

recipients of the savings deserving? There are two interpretations: (1) If WB is reduced by an act because it violates someone's sense of justice and fairness, it isn't a P.I.—so, there are few or no P.I.s. Or, if a policy is a P.I. (no one's WB declined), it is still wrong, meaning a P.I. can be wrong.

Additional evidence that people are not welfare consequentialists: trolley problems



Figure 4, Chapter 2: trolley problems



Philosophers like thought experiments, and a favorite is the *Trolley Problem*. The findings indicate that we often behave inconsistent with WC. The hypothetical-trolley scenario has numerous variations, but at issue is whether it is right or wrong for you (or someone else) to kill (or let die) one person to save five. The brake-less trolley is barreling down the tracks, soon to

derail on a curve killing all five onboard. You're standing on a footbridge with a fat guy. If you trip or push him, he falls on the track, and his large body stops the train. [Your body is too small to stop the train.] Or, an empty trolley is barreling down the tracks and will run over five people crossing the tracks unless you throw a switch that will divert the trolley onto another track where it will only run over one person. If everyone at risk is a productive member of society, a welfare consequentialist would always throw the guy or the switch. If presented with the switch story, many subjects say they would throw the switch, but if presented with the fat-guy version, few are willing to trip or push the fat guy, even though switching and pushing lead to the same number saved. Since welfare consequentialists care only about consequences, they would always switch or push.

Or a third version: you are standing by the track next to a switch. The fat guy is on the footbridge. If you throw the switch rather than diverting the trolley onto the other track, it trips the trap door the fat guy is standing on, and he drops to his death on the track. When you were both on the bridge, you might not have pushed because you were afraid you would get into a life-or-death struggle with someone bigger. But that worry is eliminated with the trap door. Still, few are willing to trip the trap door. But every welfare consequentialist always would. If you would not, you are not a welfare consequentialist.

The British philosopher and radio host [David Edmonds](#)' explanation goes to intent. The intent is to kill the fat guy if you hope to stop the trolley with his body—he needs to die for the trolley to stop—his death could not be construed as *collateral damage*. In contrast, if one switches the trolley to the other track, there is no intent to kill anyone, and once the trolley is on the other track, the five are saved even if no one is on the other track. The poor guy crossing the other track is simply collateral damage—no one is responsible for killing him. In summary, we judge actions both on consequence and process. Tweaks to the scenario are critical.⁵¹ For example, whether it is you or someone else deciding.

⁵¹ If a subject's serotonin levels are enhanced by the commonly prescribed SSRI [citalopram](#) they are less inclined to push the fat guy, particularly if they are an empathetic person ([Crockett et al. 2010](#), [Crockett 2012 and 13](#), and [Siegel and Crockett 2013](#)). [SSRI for *Selective Serotonin Reuptake Inhibitor*.] The decision to kill one to save five can be made using the executive-functioning part of your brain, thinking through the consequences, and then choosing the best outcome, which is to kill one to save five. Or it can be driven by your moral empathic-inclination to not harm an innocent person—"I can't kill someone." How much your choice is based on moral inclination versus consequence depends on the type of person you are, the amount of time you have to decide, your emotional

Conclusions and qualifications:

Questions addressed: What are the ethics of welfare economics? What is utilitarianism? And what are their ethical distinctions and similarities? The ethics of welfare economics is a unique type of WC and quite different from utilitarianism. The latter comparison is complicated by there being various forms of utilitarianism and by a hesitancy amongst moral philosophers to self-identify as utilitarians. Who is what and why depends on how welfare economists and utilitarians are defined and identified and how WC is defined. There is no book of official definitions, and you might object to some of my definitions. This is OK—but I hope you come away averse to suggesting that welfare economists are utilitarians but believing welfare economics is a type of WC, quirky like utilitarianism, but different. And a sort that other welfare consequentialists shun.

What to conclude about welfare economics? Welfare economists should question why few besides economists subscribe to it. And they should not be surprised that everyone does not accept the policy recommendations it implies. A vital distinction between welfare economics and other forms of WC (including utilitarianism) is welfare economists assume individuals always choose their HRAP. Much research in psychology and behavioral economics indicates this assumption is wrong—Mark Thayer’s recent Nobel in Economics confirms this. That welfare economics dictates that actions that do not affect others should, ethically, not be messed with depends on an assumption often violated.

An economist concerned with ethics and WB might ask how economics might inform and develop a WC ethic that does not adhere to the specific trapping of welfare economics. Alternatively, an economist concerned with ethics and WB, but not only WB, might adopt an ethic based on welfare consequences and intent (or duty, rights, or whatever). Or, an economist concerned with ethics might adopt an ethic where WB effects play no role in determining right from wrong; for example, one could adopt a Kantian ethic. As Hicks said, “one also has the right to an economics free of utilitarian assumptions.” And an economist can remain a scientist who is only trying to model—without judgment—behavior.

state, and the situation’s emotional salience. Greene et al., (2001 and 4) found, using fMRI imaging, that when the moral decision is more up-close-and-personal, there is more activation in the emotional-processing centers of your brain (e.g. the amygdala and posterior cingulate gyrus), so more conflict between emotion and reasoning. The more emotionally salient the situation (pushing the fat guy is more personal and emotional than flicking a switch), the more likely your decision is based on the moral imperative to not harm, decreasing the probability you will kill the one to save the five. Serotonin is thought to strengthen your natural aversion to doing harm.

But then again, maybe utilitarian means someone who talks a lot about utility.

Chapters 11 and 12 will continue our conversation on ethics. Now I return to behavior and choice, starting with whether economic could be the product of evolution.

ⁱ Sources:

1. Samuel Fleischacker (2013) Stanford Encyclopedia of Philosophy (hereafter, SEP on *Adam Smith's Moral and Political Philosophy*. Edward N. Zalta (ed.), <http://plato.stanford.edu/entries/smith-moral-political/>
2. Daniel Hausman (1998) Economics, philosophy of, 1998, doi:10.4324/9780415249126-R005-1. Routledge Encyclopedia of Philosophy, Taylor and Francis, <https://www.rep.routledge.com/articles/thematic/economics-philosophy-of/v-1>
3. Daniel Hausman and M. McPherson (1993) Taking ethics seriously: economics and contemporary moral philosophy. *J. Econ. Lit.* 31(2): 671-731. <http://www.jstor.org/stable/2728513?origin=JSTOR-pdf>
4. Deidre McCloskey (2008) Adam Smith, the Last of the Former Future Virtue Ethicists, *Hist. of Political Econ.* 40(1): 43-71. Also, in Jeffrey Young, ed., *The Elgar Companion to Adam Smith*, 2010. At <http://www.deirdremccloskey.com/docs/smith.pdf>
5. Edward Morey (2018) What are the ethics of welfare economics? And, are welfare economists utilitarians? *Intern. Econ. Rev.* 65(2): 201–230. DOI: 10.1007/s12232-018-0294-y
6. Vilfredo Pareto (1909) *Manual of Political Economy* [Manuale di economia politica] English translation Schwier AS,N.Y., Augustus M. Kelley, 1971.
7. John Rawls (1971, revised 1999) *A Theory of Justice*, revised edition, Harvard University Press.
8. Amartya Sen (1979) The welfare basis of real income comparisons. *J. Econ. Lit.* 17(1):1-45.
9. Amartya Sen and Bernard Williams (1982) Introduction: utilitarianism and beyond in *Utilitarianism and Beyond* (eds. Amartya Sen and Bernard Williams), Cambridge, Cambridge University Press.
10. Adam Smith (1759) *The Theory of Moral Sentiments* at http://books.google.com/books/about/The_Theory_of_Moral_Sentiments.html?id=xVkOAAAAQAAJ
11. Adam Smith (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by S. M. Soares. MetaLibri Digital. Library, 29th May 2007. https://www.ibiblio.org/ml/libri/s/SmithA_WealthNations_p.pdf
12. Leif Wenar (Spring 2017) John Rawls, SEP, Edward N. Zalta (ed.). <http://plato.stanford.edu/entries/rawls/>
13. Jonathan Wolff (2010) *On John Rawls' A Theory of Justice*. *Philosophy Bites*. <http://philosophybites.com/2010/02/jonathan-wolff-on-john-rawls-a-theory-of-justice.html>

ⁱⁱ Sources:

1. Daniel Hausman (2010) Hedonism and welfare economics, *Econ. and Phil.* 26:321-44.
2. Nien-he Hsieh (Spring 2016) Incommensurable Values, *SEP*, Edward N. Zalta (ed.). <http://plato.stanford.edu/entries/value-incommensurable/>
3. Martha Nussbaum (2012) Who is a happy warrior? Philosophy, happiness research and public policy, *International Rev. of Econ.* 59: 335-361.
4. Joseph Raz (1986) *The Morality of Freedom*, Oxford: Clarendon Press.
5. Peter Singer (1997) A philosophical self-portrait in *The Penguin Dictionary of Philosophy* (ed. Thomas Mautner), London, Penguin Books.
6. Peter Singer (2011) *Practical ethics: third edition*, Cambridge, Cambridge University Press.
7. Walter Sinnott-Armstrong (Winter 2015) Consequentialism, *SEP*, Edward N. Zalta (ed.). <http://plato.stanford.edu/archives/win2015/entries/consequentialism/>
8. L.W. Sumner (1996) *Welfare Happiness and Ethics*, Oxford, Oxford University Press

ⁱⁱⁱ Sources:

1. Kenneth Arrow (1961/63) *Social choice and individual values*, N.Y., Wiley.
2. John Broome (1988) What's wrong with poverty? *London Rev. of Books* 10(10):16-17.
3. Harsanyi J (1982) Morality and the theory of rational behavior. In: Sen A, William B (eds) *Utilitarianism and Beyond*. Cambridge University Press, Cambridge. Originally published in 1977 in *Soc. Res.* 44(4).
4. Jonathan Wolff (2010) *ibid.*

5. John Harsanyi (1995) Nobel address, *from Les Prix Nobel, the Nobel Prizes 1994*, Tore Fragsmyr (ed.), Stockholm, Nobel Foundation. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1994/harsanyi-bio.html
6. Harold Hotelling (1938) The general welfare in relation to the problems of taxation and of railway and utility rates. *Econometrica* 6:243-69.
7. David Hume (1758) *Essays and treatises on various topics*, London, A Millar in the Strand. https://books.google.com/books/about/Essays_and_Treatises_on_Several_Subjects.html?id=EHQRAAAIAAJ
8. R.F. Kahn (1935) Some notes on ideal output. *Econ. J.* 45(117): 1-35. <https://www.jstor.org/stable/2224577>
9. Frank Knight (1944) Realism and relevance in the theory of demand. *J. Pol. Econ.* 52(4): 289-318. <https://www.jstor.org/stable/1826961>
10. Paul Krugman and Robin Wells (2014) *Microeconomics 4th ed.* Worth Publishers
11. I.M.D. Little (1950) *A critique of welfare economics*. Oxford, Oxford University Press.
12. Christian List (Winter 2013) Social Choice Theory, *SEP*, Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2013/entries/social-choice/>
13. A.C. Pigou (1932) *The economics of welfare, 4th ed.* London, Macmillan, London
14. Richard Posner (1983) *The Economics of Justice*. Harvard, Harvard University Press
15. John Rawls (1971, revised 1999) *ibid.*
16. Lionel Robbins (1938) Interpersonal Comparisons of Utility: A Comment, *The Econ. J.* 48(192): 635-641.
17. Adam Smith (1759) *ibid.*
18. Tibor Scitovsky (1951) The state of welfare economics. *Am. Econ. Rev.* 41(3): 303-315. <http://www.jstor.org/stable/1802104>
19. Leif Wenar (Spring 2017) *ibid.*

^{iv} Sources:

1. J. Harsanyi (1982) *ibid.*
2. J. Harsanyi (1995) *ibid.*
3. Richard Reeves (2007) *John Stuart Mill: Victorian firebrand*, London, Atlantic Books.
4. Amartya Sen (1970) The impossibility of a Paretian Liberal, *J. of Political Econ.* 78: 152-57.

^v Sources:

1. Jeremy Bentham (1789) *Introduction to the Principles of Morals and Legislation*, first published 1789. www.earlymoderntexts.com/assets/pdfs/bentham1780.pdf
2. Jeremy Bentham (1824) *The Book of Fallacies* in the *Unfinished Works of Jeremy Bentham Vol. II* (John Bowring ed.). https://play.google.com/store/books/details/The_Book_of_Fallacies_from_Unfinished_Papers_of_Je?id=0eXO WN_vtc4C
3. Jeremy Bentham (1839) *The Collected Works of Jeremy Bentham, Volume 2: Part I* (John Bowring ed.). https://books.google.com/books?id=7eoJAAAIAAJ&dq=Bentham+the+more+remote+from+equality&source=gbs_navlinks_s
4. Jeremy Bentham, Chapter 17 in *The Collected Works of Jeremy Bentham* (eds. J.H. Burns and H.L.A. Hart), Oxford University Press, 1996. <http://books.google.com/books?id=GcnssaU-1wgCG>
5. Ken Binmore (2017) *Early Utilitarians*. Springer.
6. Fyodor Dostoyevsky (1884 in Russian) *Notes from the underground*, English version Project Gutenberg, posting date 2008, release date 1996. Produced by Judith Boss. <http://www.gutenberg.org/ebooks/600>
7. John Dinwiddy (2003) Bentham: Selected writing of John Dinwiddy, (ed. William Twining), Stanford, Stanford University Press. https://books.google.com/books?id=8_L13JxfTvgC&printsec=frontcover&dq=Bentham+Selected+writhing+of+John+Dinwiddy&hl=en&sa=X&ved=0ahUKEwiK_pg4snRAhVBMGMKHaYaBoYQ6AEIHjAA#v=onepage&q&f=false
8. Julia Driver (Winter 2014) The History of Utilitarianism, *SEP*, Edward N. Zalta (ed.). <https://plato.stanford.edu/entries/utilitarianism-history/>
9. Thomas Eisner (2003) *For love of insects*, Cambridge, Harvard University Press. <https://books.google.com/books?isbn=0674018273>
10. Anthony Gottlieb (2016) *The dream of enlightenment: the rise of modern philosophy*, N.Y. and London, W.W. Norton and Company.

-
11. R.M. Hare (1981). *Moral thinking: its levels, method, and point*. Oxford, N.Y. Clarendon Press Oxford University Press.
 12. John Harsanyi (1982) *ibid.*
 13. John Harsanyi (1995) *ibid.*
 14. David Hume (1758) *ibid.*
 15. Stanley Jevons (1879) *John Stuart Mill's philosophy tested. IV.—Utilitarianism, The Contemporary Review vol. 36: 522-538.*
https://books.google.com/books?id=M39HAQAAMAAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=falsehttp://plato.stanford.edu/entries/value-incommensurable/
 16. Christian List (Winter 2013) *ibid.*
 17. David Lyons (1991 revised) *In the Interest of the Governed: A study of Bentham's philosophy of utility and law*, Oxford, Clarendon Press Oxford.
 18. John Stuart Mill (1859) *On Liberty*, eBooks@Adelaide 2009 at http://ebooks.adelaide.edu.au/m/mill/john_stuart/m645o/ with the text from *Harvard Classics, Vol. 25*, P.F. Collier and Son, 1909.
 19. John Stuart Mill (1879) *Utilitarianism, third edition*, London, Longmans, Green, Reader, and Dyer. At <http://books.google.com/books?id=FtkIcBOZuCsC>
 20. John Stuart Mill (1987) *Bentham*, in *John Stuart Mill and Jeremy Bentham, Utilitarianism and other essays* (ed. Alan Ryan), London and N.Y., Penguin.
 21. Anthony Price (Winter 2016) Richard Mervyn Hare, *SEP*, Edward N. Zalta (ed.).
<https://plato.stanford.edu/archives/win2016/entries/hare/>
 22. Richard Reeves (2007) *John Stuart Mill: Victorian firebrand*, London, Atlantic Books.
 23. Henry Sidgwick (1874) *The Method of Ethics*, The Project Gutenberg Ebook (2014)
<https://www.gutenberg.org/files/46743/46743-h/46743-h.htm>
 24. Peter Singer (2011) *ibid.*
 25. Peter Singer (1997) A philosophical self-portrait in *The Penguin Dictionary of Philosophy* (ed. Thomas Mautner), London, Penguin Books.
 26. Peter Singer (1975, 4th edition 2009) *Animal liberation: A new ethics for our treatment of animals*, Harper Collins.
- vi Sources:
1. Kaushik Basu, Luis Lopez-Calva (2011) Chapter 16: Functioning and capabilities. In Arrow K, Sen A, Suzumura K (eds) *Handbook of Social Choice and Welfare Vol 2*. [https://doi.org/10.1016/S0169-7218\(10\)00016-X](https://doi.org/10.1016/S0169-7218(10)00016-X)
 2. Nussbaum M (2006) *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, Harvard University Press.
 3. Martha Nussbaum (2011) *Creating Capabilities*, Cambridge, Harvard University Press
 4. Ingrid Robeyns I (2016) *The capability approach*, *SEP*, Edward N. Zalta (ed.),
<https://plato.stanford.edu/archives/win2016/entries/capability-approach/>
 5. Amartya Sen (1980). Equality of what? In: McMurrin S (ed) *The Tanner Lectures on Human Values, Vol. 1*. Cambridge University Press, Cambridge
- vii Sources:
1. Molly Crockett (2012) Molly Crockett on Brain Chemistry and Moral-Decision Making, *Philosophy Bites*.
<http://philosophybites.com/2012/07/molly-crockett-on-brain-chemistry-and-moral-decision-making-originally-on-bioethics-bites.html>
 2. Molly Crockett (2013) Models of Morality, *Trends in Cognitive Sci.* 17(8).
<http://dx.doi.org/10.1016/j.tics.2013.06.005>
 3. Molly Crockett, Luke Clark, Marc Hauser, and Trevor Robbins (2010) Serotonin selectively influences moral judgment and behavior through effects on harm aversion, *Proc. of the National Academy of Sci.* 107(40): 17433–17438. [10.1073/pnas.1009396107](https://doi.org/10.1073/pnas.1009396107)
 4. David Edmonds (2013) David Edmonds on Trolley Problems, *Philosophy bites*.
<http://philosophybites.com/2013/09/david-edmonds-on-trolley-problem.html>
 5. David Edmonds (2013) *Would you kill the fat man? The trolley problem and what your answer tells us about right and wrong*, Princeton, Princeton University Press. <http://press.princeton.edu/titles/10074.html>

-
6. Jesse Graham, Jonathan Haidt and Brian Nosek (2009) Liberals and conservatives rely on different sets of moral foundations, *Personality Process and Individual Differences* 96(5): 1029-1046.
 7. Jonathan Haidt (2012) *The Righteous mind: Why good people are divided by politics and religion*, N.Y., Pantheon Books.
 8. Jonathan Haidt and Jesse Graham (2007) When morality opposes justice: conservatives have moral intuitions liberals may not recognize, *Soc. Justice Res.* 20:98-116.
 9. Jonathan Haidt, Silvia Koller and Maria Dias (2003) Affect, culture and morality, or is it wrong to eat your dog, *J. of Personality and Soc. Psych.* 65(4): 613-28.
 10. Joseph Henrich, Stephen Heinz and Ara Norenzayan (2010) The weirdest people in the world, *Behav. and Brain Sci.* 33: 61-135.
 11. Paul Krugman and Robin Wells (2014) *ibid*
 12. Jenifer Siegel and Molly Crockett (2013) How serotonin shapes moral judgment and behavior, *Annals of the N.Y. Academy of Sci.* 1299:42-51. [10.1111/nyas.12229](https://doi.org/10.1111/nyas.12229)

Chapter 3: Could economicus be the product of evolution rather than the creation of God? Unlikely

08142022

Evolution is driven only by the survival of genes/traits; its objective is not to produce gene containers that are happy or error-free¹

Evolution has two components: (1) genes generate traits in their containers (us and other living things). These traits increase or decrease the container's probability of passing its genes to the next generation. This causes the population stock of some genes to increase and others to decrease. And (2) when genes are transmitted, transcription errors occur; these are usually disastrous for survival but occasionally fortuitous.

Evolution takes place at the gene level rather than at the level of the organism—living things are simply the copying machine. As the eminent evolutionary biologist [Richard Dawkins](#) explains in “[The Selfish Gene](#),”¹

They are in you and me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. They have come a long way, those replicators. Now they go by the name of genes, and we are their survival machines.

Evolution does not require happy gene containers nor implies that the containers will become happier over time. Consider lust, the great contributor to sexual angst in teenage boys; this evolutionary trait increases reproduction, but rather than being happy, the container is, in the words of Elvis, "All shook up". The aim of evolution isn't to maximize our WB: always happy and satisfied isn't an appropriate mindset for surviving in a world of hungry, brutish beasts—not for the beast nor the prey.

Naturally, animals (and plants) suffering or dying because they lack the resources to survive are less likely to reproduce and nurture their children. But, a lot of WB isn't required. If things are too good pre-kids, you won't have any, which wouldn't serve your genes.

A caution: a strand of genes typically influences a set of traits, but that does not mean each trait, by itself, makes it more likely the container will successfully reproduce; genes produce side-effect traits. Some side-effects won't influence reproduction, and others, taken

¹ In 2006, the book's thirtieth anniversary was celebrated with a [conference](#) at the London School of Economics.

alone, will reduce the probability of successful reproduction. The textbook example: genes evolved to produce a trait, sickle-shaped blood cells. This happened in sub-Saharan Africa, where malaria was common and deadly. The sickle shape provides protection: the parasite needs round blood cells to prosper. But with the trait comes sickle-cell anemia, a negative side-effect.

A secondary trait more germane to our topic is striving to acquire more resources, even after one has enough to successfully reproduce. A strand of genes that conveys the desire to survive and reproduce will increase the probability of those genes surviving. But this desire becomes a side effect if it does not turn off when one has enough to survive and reproduce—and for gene transmission, there is no advantage to turning the desire off.

Evolution, since it's typically incremental, gravitates towards a local best rather than the bestest best (the global best)ⁱⁱ

Evolution wouldn't produce the perfection of economicus; it proceeds from its current state in small steps. Recipients of drastically mutated genes (a big step) typically don't survive. Consider, for example, a species' current system for night vision. If better night-vision will increase survival and reproduction, and if better night-vision is a tweak away, evolution will improve night vision. Figure 1 shows that through many small steps, evolution could move from the current position to a local max to its left, but it won't take us to the best possible night-vision (the global maximum).²

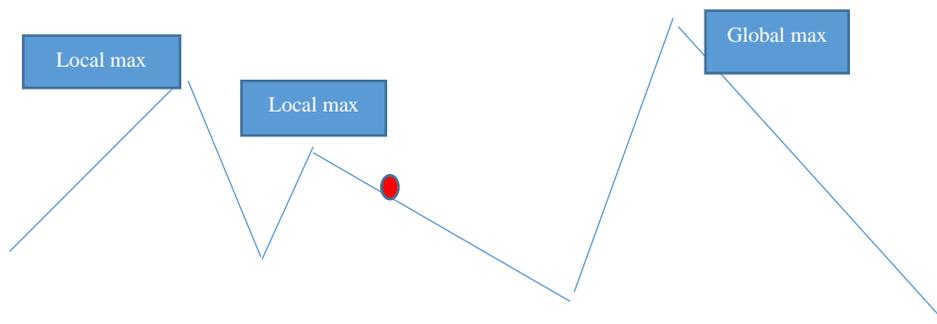


Figure 1, Chapter 3: Night vision acuity on vertical axis, genetic mix on horizontal

red dot is current genetic mix

If a better system is utterly different from the current one, evolution is unlikely to get us there. This holds for our brain system as well as our vision. N.Y.U. psychologist [Gary Marcus](#)

² Any marginal shift to the right would reduce survival rates.

describes our brain, a product of evolution, as a “kluge”: n. Slang, "A clumsy or inelegant solution to a problem”, in his amusing and informative book *Kluge: The Haphazard evolution of the human mind*. While it serves us well in many ways, our brain isn't a best mechanism; evolution isn't about perfection, but what the economics Nobel laureate [Herbert Simon](#) calls *satisficing*, doing well enough.

Humans suffer the limitations of evolution. One is that humans mix beliefs with logic and inference. Marcus conjectures that formal logic, to the extent we practice it, isn't a product of evolution but something we learn in school and sometimes find useful. In the 1930s, the Russian psychologist [Alexander Luria](#) (1902-1977) went to villages in the mountains of Central Asia and interviewed indigenous, uneducated villagers. He told them that all the bears in a particular Siberian town are white. He then told them their neighbor had visited that town and saw a bear. Then he asked, “What color was the bear?” The typical answer was, “Ask my neighbor”. Logic implies white. The psychologists [Amos Tversky](#) (1937-96) and Kahneman describe Linda:

Linda is 31-years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable? 1. Linda is a bank teller. 2. Linda is a bank teller and is active in the feminist movement. Many modern literate respondents illogically choose 2: the probability of two events occurring simultaneously can't be greater than the probability of only one of them occurring. We get it wrong because Linda sounds like someone who might be a feminist. In a 2017 book, *The enigma of reason*, [Hugo Mercier](#) and [Dan Sperber](#) argue that humans did not evolve to follow the rules of logic. Instead, we evolved to reason (create and give reasons/justifications), so other humans would accept our ideas and plans—we evolved to convince, not to be logical. And we evolved to be convinced by the reasons of others only when it is in our interest. In this view, reasoning evolved to increase cooperation, not logical thinking.

NCT emerged and developed in 19th and early 20th century Britain, a Christian creationist world where man was unique and in God's imageⁱⁱⁱ

In 19th-Century England, only Christians could be college professors. Developers of NCT include [William Stanley Jevons](#) (1835-1882), [Francis Edgeworth](#) (1845-1926), Marshall, and the Nobel Laureate [John Hicks](#) (1904-1989). If NCT was a religion, the works of these economists would form its Canon.

Jevon’s diary and correspondence demonstrate that the originator of marginal utility believed in God and found no conflict between his Christian faith and science. Edgeworth (the creator of indifference curves) was a friend of Jevons.³ Marshall is the father of consumer’s surplus; his father was “a devout Evangelical”. Robert Frantz, in his 2005 book *Two Minds: Intuition and Analysis in the History of Economic Thought*, notes that Marshall,

...integrated his economics into his ideas about religion, character, duty, and intuition... Marshall also considered the joy brought about from religion as the ‘highest joys of which men are capable.

Hicks advocated for an NCT where the ordering of paths has no cardinal properties (Assumption 2b). He was raised a Baptist. These gentlemen were aware of Darwin’s theory but also seeped in creationist Christianity, so it shouldn’t surprise that they adopted a theory of behavior where man is perfection—in God’s image.⁴ I don’t know the depths of their Christian faith, but their NCT makes assumptions consistent with Christianity’s view of man, assumptions evolution does not imply.

Creationism and NCT are symbiotic bedfellows^{iv}

The late 19th Century drive towards mathematical formalism and logic in choice theory, led by Jevons, Marshall, and the Austrian [Carl Menger](#), melded effortlessly with the notion that man is an optimizer—a goal was to make economics more science-like. Human behavior is difficult to mathematically model if human behavior results from the awkward process called evolution. Human behavior is easier to model if God created man, and man maximizes subject to his constraints.

The next chapter defines happiness and asks where it came from and why. Whether more emotional WB/happiness should be the societal goal was discussed in Chapter 2, and I will return to the topic in Part III.

ⁱ Sources:

1. Richard Dawkins (1976) *The Selfish Gene*, N.Y.: Oxford University Press
2. Alan Grafen and M. Ridley (2006) *Richard Dawkins: How A Scientist Changed the Way We Think*, N.Y.: Oxford University Press.

ⁱⁱ Sources:

1. Gary Marcus (2008) *Kluge: the haphazard evolution of the human mind*, N.Y., Houghton Mifflin Harcourt.

³ An indifference curve for Path *j* identifies all paths that are ranked the same as Path *j*.

⁴If humans are both logical and created in God’s image, then God must be logical, so not a holistic Yin-Yang kind-of-guy.

-
2. Hugo Mercier and Dan Sperber (2017) *The enigma of reason*, Cambridge MA, Harvard University Press.
 3. Amos Tversky and Daniel Kahneman (1983) Extension versus intuitive reasoning: the conjunction fallacy in probability judgement, *Psych. Rev.* 90(4): 293-315. [doi:10.1037/0033-295X.90.4.293](https://doi.org/10.1037/0033-295X.90.4.293).

iii Sources:

1. Roger Frantz (2005) *Two Minds: Intuition and Analysis in the History of Economic Thought*, Springer.
2. Daniel B. Klein and Ryan Daza (2013) John R. Hicks [Ideological Profiles of the Economics Laureates] *Econ J. Watch* 10(3): 366-377. At <http://econjwatch.org/ancillary/IPEL.html>
3. Bert Mosselmans (2015) SEP on *William Jevons*. <http://plato.stanford.edu/entries/william-jevons/>
4. Wikipedia, Alfred Marshall, http://en.wikipedia.org/wiki/Alfred_Marshall

iv Sources:

1. Margaret Schabas (1989) Alfred Marshall, W. Stanley Jevons, and the mathematization of economics, *Isis* 80 (1): 60-73.

Chapter 4: Wanting versus liking, and happiness: an electric14al/chemical stew

08142022

A brief history of happiness and WB¹

In Chapter 1, I distinguished between WB and its components: emotional WB, life-satisfaction WB, and pleasurable sensations—but in the literature, the word “happiness” is often used as a catchall. Emotional WB and life-satisfaction WB is modern terminology, so for my brief history, I will simply talk about happiness's history, noting how it has changed over time. Then, in the neuroscience of happiness section, I will revert and use “happiness” as a catchall for emotional WB. Unfortunately, neuroscientists mainly study non-humans, so they don't much study life-satisfaction WB: it is difficult to assess a rat's life-satisfaction WB.

Thomas Jefferson inscribed an American's Right to the pursuit of happiness.¹ In 1843, the cranky, contrarian essayist, [Thomas Carlyle](#), pointed out²

Every pitifulest whipster [whippersnapper, wannabe hipster] that walks within a skin has his head filled with the notion that he is, shall be, or by all human and divine laws ought to be, 'happy.' His wishes, the pitifulest whipster's, are to be fulfilled for him; his days, the pitifulest whipster's, are to flow on in an ever-gentle current of enjoyment, impossible even for the gods.

Why is happiness a societal goal? —it hasn't always been. This chapter reviews how the conception of happiness has evolved. A modern view is that it's a chemical state of mind.

Whether more is always better is left for later. [Jim Holt](#) (an American philosopher and essayist) imagines the history of happiness with bumper-sticker slogans:

Happiness= Luck (Homeric), Happiness=Virtue (classical), Happiness=Heaven (medieval), Happiness=Pleasure (Enlightenment) and Happiness=A Warm Puppy (contemporary).

If the WB of Assumption 9a is defined as happier—it does not need to be—we need to consider what happier means.

¹ The right to pursue it, not the right to it.

² Carlyle coined economics the “dismal science”.

The pre-Enlightenment Western view of happinessⁱⁱ

An ancient view of happiness

Aristotle (384–322 BC) distinguished between pleasure and *eudaimonia*, often translated as happiness, though scholars translate it as “human flourishing” or “living well”: others think it best to not translate it. For Aristotle, the goal was *eudaimonia*, a way of life (pursuing certain activities), rather than pleasurable sensations and positive emotions—one was a happy camper if they died having lived a meaningful life. This required a life of *arête* (excellence) in all activities, including “morally virtuous activities”. You had to experience a life of *reasoning*. Why? Because humans have the unique ability to reason; it is part of our nature. Things are best utilized if they do what they were uniquely designed to do: hammers should pound nails, and humans should reason.

Different activities produce different physical and mental sensations, and it is difficult to separate them from the activity that produced them. Pleasurable sensations are the ones that feel good. The activities that promote *eudaimonia* might, or might not, produce pleasure. Aristotle had nothing against pleasure: it just wasn’t the greatest good.

Aristotle outlined two incompatible paths to *eudaimonia*: the practical social/political life and the intellectual life. For the first, one needs the practical virtues—the appropriate amounts (not too much, not too little) of courage, temperance, generosity, honor, friendliness, and wit, to name a few. In contrast, the intellectual virtues include a commitment to studying philosophy or science and educating the young to pursue truth and reason.

At the same time, Aristotle makes it clear that in order to be happy one must possess other goods as well—such goods as friends, wealth, and power. And one's happiness is endangered if one is severely lacking in certain... advantages—if, for example, one is extremely ugly, or has lost children or good friends through death. But why so?... Aristotle's reply is that one's virtuous activity will be to some extent diminished or defective, if one lacks an adequate supply of other goods. Someone who is friendless, childless, powerless, weak, and ugly will simply not be able to find many opportunities for virtuous activity (Shields 2014)

So, for Aristotelians, *eudaimonia* requires luck, position, and resources, so only for the elite. Animals can’t have it (they can’t reason), and one must be old—kids have not lived long enough. Sex, drugs, and rock’n’roll won’t do it, no matter how pleasant the sensations produced; sensations do not cause a life to be happy or unhappy. For Aristotle, living a virtuous life is right for you and society, making Aristotle an early advocate for *Virtue ethics* (one is ethical if one is virtuous). Virtue ethics is discussed in more detail in Chapter 11.

In contrast, Epicurus (341-270 BC) aimed to experience pleasurable sensations and positive emotions. But on a practical level, this was best achieved by avoiding pain and suffering—a happy life is a tranquil life. Like Aristotle, he was all for a virtuous life, but, unlike Aristotle, he thought people should be virtuous because it would make encounters with others more pleasant (cheating and lying might get one beat up or worse). There are similarities with Buddhism: trying to fulfill desires and cravings causes more craving and stress, and tranquility is as good as it gets. Epicurean happiness is a root of Benthamite utilitarianism.

*The Medieval (Middle Ages) view of happiness: happiness=heaven*ⁱⁱⁱ



Figure 1, Chapter 4: Lotario dei Segni

Paraphrasing Lotario dei Segni (Lotario of the Signs), before he became Pope Innocent III in 1198,

Men and women are simply vile. Conceived in the stench of lust and formed of the filthiest sperm, we spend our earthly days in misery, toil, and degradation. (Quoted by McMahon)

Worldly happiness was not expected; a good life was a devout, religious life. Happiness was being blessed by God in the afterlife.

Suffer for virtue's sake, medieval Christians were told, and you will be rewarded with an eternity of heavenly bliss. (Holt)

Medieval-Christian happiness is a sensation, so not Aristotelian—but one available only to dead people, a sensation achieved, eventually, by serving King, Church, and tradition, and by Stayin' Alive—life was “nasty, brutish, and short”, with an occasional drunken feast. Thomas Aquinas (1225-1274), a Dominican monk and Aristotelian scholar, allowed for a bit of worldly happiness, distinguishing between perfect happiness (*beatitudo*)—only available in heaven—and imperfect happiness on earth (*felicitas* or *beatitudo imperfecta*). If an individual has sufficient

resources and abilities, and she plays her cards right, she gets *duplex felicitas* (both kinds of happiness)

Why does society care whether you are happy? The Enlightenment and the reactions to it^{iv}

The *Enlightenment* (late 17th to late 18th Century) was a revolution in thinking, an intellectual movement that proposed reason, logic, and evidence as the tools for decision making— “The Age of Reason”—a reaction to things being done based on tradition, the Church, and the monarchy. It followed in the footsteps of the scientific revolution of the 16th and 17th centuries. [The word “enlightenment” should be interpreted as wanting to shine the light of reason on questions, not when we were necessarily enlightened by the truth.]

The Enlightenment advocated for the individual, for thinking for oneself, for self-determination, and in summary, for *autonomy*.

[The Enlightenment] The period during which modern identity was forged...People become the source of power and authority...The individual has the right to decide, personally, for all of their actions. (Tzvetan Todorov, a modern French/Bulgarian essayist)

While these ideas go back to ancient Greece, the Enlightenment combined the old ideas in new ways and applied them to the common man. The American and French Revolutions were motivated by it and are expressions of it. In the 4th century, St. Augustine divided actions into two types: a means to an end and those we directly enjoy. And the only one we should enjoy is worshipping God. The Enlightenment opened the door to enjoying worldly things.

Enlightenment names include [Francis Bacon](#) (1561–1626), [Baruch Spinoza](#) (1632-1677), [John Locke](#) 1632-1704), [Denis Diderot](#) (1713-1784), [Voltaire](#) (1694-1778), Hume, [Immanuel Kant](#) (1724-1804) and [Isaac Newton](#) (1642-1726). We in the West are children of the Enlightenment and can thank it, for better or worse, that we, as individuals, have standing and that our personal welfare counts for something. Without this revolution in thinking and importance, there would be no WC nor welfare economics.

While Enlightenment thinkers advocated for scientific thinking, that individual autonomy is of value is an ethical position, not a scientific discovery. Its endorsement was not based on research in psychology, sociology, or neuroscience; there was none. Instead, it was a negative reaction to the medieval ethic that the suffering of ordinary people was unimportant.

The start of the 19th Century saw a partial pullback from Enlightenment thought. The reaction, the Romantic Age, added back sensations, emotions, and mystery—think of English poets such as Lord Byron, Percy Shelley, and Samuel Coleridge, and painting by Eugene Delacroix, Casper David Friedrich, and J.M.W. Turner.



Figure 2, Chapter 4: Rousseau

The works of [Jean Jacques Rousseau](#) (1712-78) bridge from the Enlightenment to Romanticism. Thank him for equating welfare with happiness:

...the love of happiness is the sole principle of all human actions

N.B. the distinction between (1) giving importance to the individual and his WB and (2) defining WB as happiness. The Enlightenment caused the first, Romanticism solidified happiness as the goal—the Romantic Age was the age of feelings and emotions.

There are dark sides to individualism and the right to happiness: two are boredom and frustration. As the Norwegian philosopher [Lars Svendsen](#) notes, before the Enlightenment, life was largely determined by the constraints of tradition, religion, and staying alive. Now we must choose a *lifestyle* (Buddhist, or bass fisherman, or Buddhist bass- fisherman?),³ and we expect our choice will make us happy. So, we experience frustration and disappointment when it does not—so try another. There are potential costs to having the right and resources to worry about happiness, including boredom, indecision, frustration, and anxiety. Except for monks and nobility, boredom is a modern emotion: “A ‘privilege’ of modern man.”

An Eastern view

There is an Eastern Western happiness distinction. Buddhism, for example, teaches that the pursuit of happiness is destined to fail and that pleasures are fleeting, a belief consistent with recent research (see Chapter 5). Right process, rather than the consumption of goods, is the path

³ He defines a *lifestyle* as “essentially a set of practices maintained for a period of time.”

to Buddhist WB, making the Buddhist practice of a good life closer to the ancient Western practice than to the modern Western notion. It as an ethic will be discussed in Part III, Chapter 11. Buddhist practice as a path to WB is discussed in Chapter 5.

The late 19th Century view(s) of happiness^v



Figure 3, Chapter 4: Jeremy Bentham (embalmed)

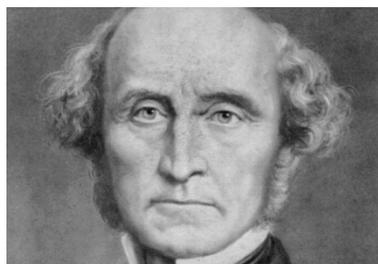


Figure 4, Chapter 4, John Stuart Mill

Bentham and John Stuart Mill defined happiness as Epicurus did, as Rousseau did. Happiness is positive minus negative feelings, where these sensations have both duration and intensity.⁴ Epicurus and Bentham, and maybe Mill, are what philosophers call *pleasure hedonists*—guys who equate good with pleasure. Bentham and Mill are typically viewed as founders of utilitarianism; whether Mill was depends on what of his you read and how much you read between the lines.

Aristotle's happiness is not rooted in WB; rather, it is a synonym for a virtuous life. *WB*, *emotional WB*, and *life-satisfaction WB* are not Aristotelian happiness.

Epicurus and Bentham equated happiness with emotional WB. Mill's view of happiness and WB is complicated: Mill considered freedom to choose a component of WB, independent of whether it increases happiness or affects behavior. At issue is whether the pleasure one gets from an activity depends on the process or circumstances that made it possible. For example, would

⁴Mill, but not Bentham, argued that while two activities might be equally pleasurable, pleasure produced by high-brow activities generates more social good than pleasure produced by low-brow activities. Mill argued the social worth of a pleasure depends on how it was produced—In Mill's view, Rousseau's pleasure from reading Shakespeare was better than his pleasure from picking his nose, and Rousseau's pleasure from scratching his butt was better than his pig's pleasure from Rousseau scratching the pig's butt—simply because his pig was just a pig. There is more on this ethical distinction in Chapter 11. You might disagree with my assessment and argue Mill simply thought high-brow activities produce more pleasure. But if you interpret Mill in that way, there is no need to judge social good based on anything other than pleasure.

forcing you to eat a piece of cake, a piece you would have chosen to eat, diminish the pleasure of eating it? [More on freedom in the section of Mill in Chapter 11.]

Perceptions influence the assessment of life satisfaction; for example, was the enjoyment experienced the result of a misperception— “I foolishly believed Wanda loved me but now realize my time with her was a sham.” Other concerns include recollection errors, false projections, expectations, self-delusion, and discounting past and future sensations. WB is complicated: economicists might order only on emotional WB, others on only life-satisfaction WB. Some researchers ask happiness questions, thinking it the best measure of WB, while others ask life-satisfaction questions. Imagine policies that increase one but not the other. Research suggests that a permanent increase in income causes a long-run increase in life-satisfaction WB but only a short-term happiness bump.

In both research and popular culture, it is often unclear what a writer or speaker means when they use the words “happiness” or “WB”—often, they have not settled on what they mean.

Independent of how one defines individual WB, a critical ethical question is whether WB is the criterion that should be used to determine whether an act or policy is right or wrong. Welfare economists say “yes”—Part III discusses moral philosophies that answer “no”.

Modern happiness/emotional WB: a chemical state of mind

Happiness is a brain state

Some things make our brains feel good; some things make us feel bad. For example, being cold or having a cold is unpleasant, and eating toast feels good. (Robert Frank, Cornell economist and N.Y. Times columnist)

The modern scientific view—a view many of my students hate—is happiness is simply a brain state, a mixture of chemical and electrical activity in the brain that causes you to feel good. [I like to exercise till exhaustion. Why? Because afterward, endorphins dance through my brain, the "endorphin rush", as effective as cocaine.]

Dualists and Materialists^{vi}

Materialists believe the mind and brain are the same—everything has a physical explanation, including our thoughts, emotions, and consciousness—the mind isn’t ethereal. In contrast, *Dualists* distinguish between the brain, a physical thing, and the mind, an immaterial spirit, believing that while much is explained by physics and chemistry, the human mind transcends the

physical.⁵ Put simply, for dualists, the mind creates feelings and thoughts that are incapable of ever being explained with physics or chemistry—so magical.⁶



Figure 5, Chapter 4: Rene Descartes 1598-1650

Rene Descartes went to great lengths to defend dualism. Why? Maybe because he wanted to defend Catholic doctrine? Maybe because he did not want to be labeled a heretic.⁷ While few neuroscientists are dualists, the guy-on-the-street won't abandon it. While he accepts physical explanations for many things, he is hesitant to believe who we are and what we feel is solely the result of a chemical/electrical stew. I think of one student, an army officer from the South, who shuddered in protest whenever I suggested he was only a "material boy". He was baffled that I would even imagine such a thing.

Many people, including Descartes, are willing to accept physicality for everything non-human.⁸ But not me. Unwilling to accept mind over matter, I am a material boy living in a material world made of physical matter with lots of chemical and electrical connections, and so are you. We are moist sophisticated robots (machines/automata); in the 1874 words of the

⁵For a modern questioning of materialism, see Thomas Nagel's 1974 essay, [What is it like to be a bat?](#)—a fun read. Nagel, a philosophy professor at NYU, argues while one can't reject the argument that the experience of mind is the result of a chemical stew, it's hard to argue it's only that, at least not without a theory of how the stew generates our sense of self and the feeling we are experiencing. But! no such theory [now] exists. He questions whether it's even possible to have a theory that equates every mental state with a specific and physical brain-state, at least within the confines of modern physics.

⁶ Dualists are not required to believe that the magic extends beyond the minds of men; they are also not required to believe in God, and one could arguably believe in God and not be a dualist. Conversely, you can believe some magic exists without believing it affects your mind.

You might prefer the word "supernatural" to "magical"; as in, outcomes not consistent with the current laws of nature.

⁷ During the enlightenment, philosophers went to great lengths (Hume being the poster boy) to not be labeled a heretic, atheist, or unbeliever. They criticized church and religion in code. Hume was not surprised when [because Rousseau's book *Emile* contradicted the notion of original sin] the city of Paris issued a warrant for Rousseau arrest. Writing to Rousseau's English host, Hume noted that Rousseau "had not the precaution to throw any veil over his sentiments; ...he could not wonder that all the zealots were in arms against him." Concern was warranted; it was only in 1697 that Britain hung its last person for blasphemy, [Thomas Aikenhead](#).

⁸ "Descartes held that non-human animals are automata" (Kirk): they have no conscious experiences, and their behavior is wholly determined by physical mechanisms. Descartes considered, but rejected, the idea of a machine that looked and behaved like a human being (a zombie).

famous British biologist and materialist [T.H. Huxley](#), I am a “conscious automata”. So are you unless you are a zombie (sentient but not conscious). I have no way of telling; I hope my wife isn’t one. Zombies raise a problem for materialism: the existence of zombies is inconsistent with materialism, and proof of a zombie would confirm dualism.⁹ Zombies can’t be happy; they can’t feel.

Science hasn’t proven all our thoughts and actions are embodied in our brain, but the alternative requires an acceptance of processes and agents (spirits) that don’t adhere to the laws of physics and chemistry. People believe in things inconsistent with the laws of physics, including many of my friends, economic colleagues, and students; many physicists do not. For example, most of us are at least a bit superstitious. Human dualism is difficult for humans to abandon—many religions and cultures teach dualism—imagining the mind transcends the brain is comforting—it makes people feel superior to animals. And besides, we don’t know how the brain does what it does—it is even hard to imagine knowing how it works—making it easy to believe there is magic.

Happiness: a brain state

Happiness is associated with a whole class of brain states, many overlapping, each state producing a different kind of WB experience. So, in theory, science could identify a set of brain states and define happy as one or more of those states.

So, what is a brain on happiness? Your brain in a nutshell^{vii}

Happiness, like all feelings, is the expression of neural activity in the brain; increased happiness is simply more of some neural activity and less of other sorts. You eat an ice-cream cone, pet a dog, and play chess because you hope these activities will change the synapses in your brain in ways you will experience as nice—happier is simply a matter of modifying your brain state. This section is about how that works. While my interest in the brain started with the psychology of choice and then progressed directly to neuroscience, there is a new field in economics called *neuroeconomics*. A few economics departments offer courses. While I initially did not encounter the neuroscience of WB and choice through the lens of neuroeconomics, this chapter draws, in

⁹ The argument: if zombies and humans are physically identical but humans have conscious experiences and zombies don’t, then conscious experience can’t be the result of body parts—humans and zombies have the same parts. Therefore, if zombies exist, human consciousness arises from the mind, not the physical brain.

part, from the textbook, *Neuroeconomics 2nd edition*, edited by the neuroeconomists [Paul Glimcher](#) at NYU and [Ernst Fehr](#) at the University of Zurich. Each chapter is written by noted experts. The other textbooks that I draw on are *Neuroscience: exploring the brain 3rd edition* by [Mark Bear](#), director of the BEAR Lab at MIT, [Barry Connors](#), Professor of Medical Science at Brown, and [Michael Paradiso](#), the Director of the Center for Vision Research at Brown. I also draw on *Cognitive neuroscience: the biology of the mind (4th ed.)* by [Michael Gazziniga](#), [Richard Ivry](#), and [George Mangun](#).¹⁰ Gazziniga is at the U. of California Santa Barbara, Ivry at U. of C. Berkeley, and Mangun at U. of C. San Diego.

A human brain contains approximately 86 billion neurons.¹¹ Collectively, neurons generate sensations and feelings¹² and embody your knowledge (your dog’s name and your skateboarding skills), memories, tastes, and goals. Neurons generate your behavior. The *cell body* of each houses its *nucleus* and power source. Neurons communicate; each neuron has a transmission tower called an *axon*, and one or more antennas (receiving towers), called *dendrites*. Axons vary drastically in length; the sciatic nerve runs the length of your leg; many axons traverse from one part of the brain to another—these are long-distance transmission lines; most are much shorter.

¹⁰ For a short introduction to the brain see “Perception Lecture note: the brain ([David Heeger](#) 2014). The notes include fantastic pictures.

¹¹ Everyone used to say 100 billion but [Suzana Herculano-Houzel](#) counted 86 billion by turning five brains into soup and counting the number of nuclei in the soup.

¹² The burning of a muscle is an unpleasant *sensation* (a physical sensation), your interpretation of that sensation is a perception that you might, or might not, find pleasing.

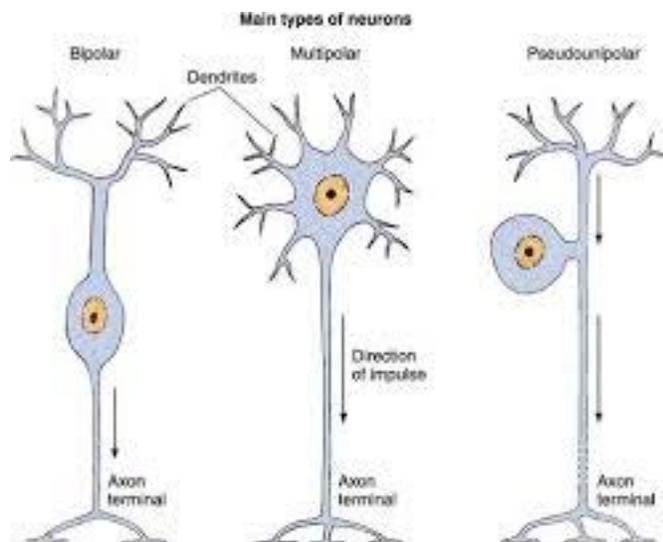


Figure 5, Chapter 4: Neurons

The end of the axon, the *axon terminal*, is stocked with chemicals called neurotransmitters (e.g., serotonin and dopamine). Each axon terminal (the end of the transmission tower) almost touches one or more antennas (the receptors on other neurons' dendrites). The gaps, called *synapses*, are about 20 billionths a meter wide. Neurons communicate across these gaps: an axon terminal releases a packet of its neurotransmitter; these cross the gaps and bond with the receptors on other neurons' dendrites.

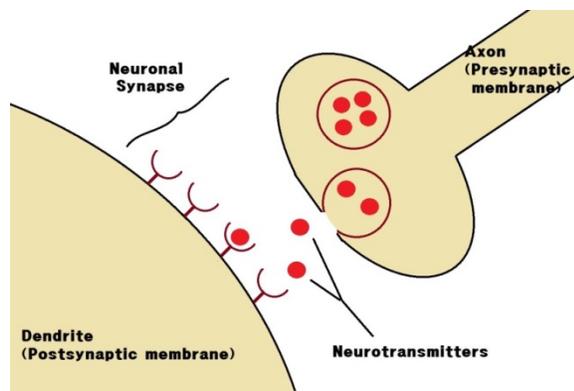


Figure 6, Chapter 4: Neurotransmitters released across a synapse

Axons can branch (divide), so synapse with many receptors; sometimes, these receptors are all on the same recipient neuron; sometimes, they are on different neurons. You have sixty

trillion synapses—give or take. Many neurons synapse with the axons of many neurons (see Figure 7),

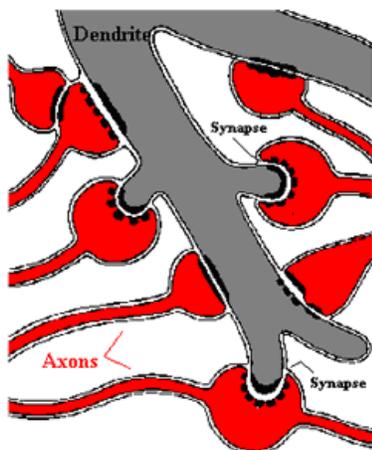


Figure 7, Chapter 4: A dendrite that synapses with a bunch of axon terminals

When an electrical impulse is transmitted down a neuron's axon, its terminal releases neurotransmitters. The technical term for this electrical impulse is an *action potential*. A neuron's axon fires (experiences an action potential) after its dendrites are the recipients of enough neurotransmitters.

There are many types of neurons and neurotransmitters.¹³ We all know about serotonin and dopamine. GABA (gamma-aminobutyric acid), norepinephrine, enkephalin, and anandamide are others that explain pleasure and reward-seeking.

Your neurons determine who you are, what you do, and how you feel. How does that work? All the details are not known, but many are. Genetics, experiences, and what you ingest affect your neurons; specifically, they affect the number and efficacy of the synapses. “Neurons that fire together wire together.” And “Neurons that fire out of sync lose their link.” *Wiring together* means the number of synapses they share increases, and each's effectiveness (efficacy) increases. A neuron firing alone will not strengthen the link between itself and a post-synaptic neuron; the link will weaken. But if many neurons simultaneously shoot neurotransmitters at the same neuron, all the links will get stronger.

¹³ Most neurons only connect with other neurons, but some also connect with muscles, glands, and other tissues. The firings of neuromuscular cells cause your heart to beat and your feet to move as you walk down the street. Nerve cells in the skin transmit touch to the brain, and neurons in the eyes start the process of seeing.

Consider all the things that can change the efficacy of a synapse: the amount of the neurotransmitter stored in the axon terminal, the amount it releases and for how long, and the receptiveness of the dendrites on the recipient neuron. These are influenced by what you have experienced: the road trip to Yellowstone when you were twelve, the toast at breakfast, and you're reading this book. There are temporary modifications (short-term memories); others last a lifetime. Learning simply modifies neural connections, and learning and pleasures work many of the same neural pathways. Like rats pushing bars for pellets, we learn behaviors that get reinforced, and pleasure is one type of reward. Your ordering of paths—if you have one—is embedded in your neurons.

Mice are afraid of cats unless a neurologist modifies their neural connections so they are not—it has been done.

You recognize an object as a cat because your brain contains a neural synaptic imprint of a cat (an *engram*, a *memory trace*). It is a cat when enough neurons in your cat imprint fire.¹⁴ My previous dog was white, so when, out of the corner of my eye, I glimpse a white pillow on the floor, I register Sofie. That we like sweets and adolescent males pursue girls rather than learning Latin is determined by neural connections and hormones.

Choosing/seeking, pleasure, and the mesolimbic dopamine system (the MDS)

Economists use the word “choose”, and neuroscientists “seek”. While we often like what we sought/chose and seek/choose what we will like, one can choose what you will not like and like what you did not pursue. Choosing/seeking is behavior; liking, by itself, isn't a behavior; it's a feeling. (When I throw the ball, Giacomo seeks it (chooses to pursue it); at academic cocktail parties, many of us seek the wine rather than the soda—what we seek is a choice.) The path you experience is the result of seeking, **not** necessarily liking.

We believe we seek/pursue a commodity or an experience because realizing it will increase our WB (we will like it). This assumption, Assumption 9a of NBT, is a foundation of welfare economics: NBT with 9a assumes the path pursued is the available path that we will **like**

¹⁴ Your hippocampus plays an important role in how you store episodic memories (what you got for your birthday). Recent research, using electrodes implanted in epilepsy patients, indicates that each episodic memory is encoded in about 2% of the neurons in your hippocampus (millions of neurons), so “each neuron contributes to the representation of only a few memories” ([Wixted](#) et al.) See also Quenqua. The good news is if you lose a few thousand neurons, you will still remember your wedding. **Need to review the section on memory in the G textbook.**

the best and the path we will experience. [Some economists say it does not matter whether we end up liking what we wanted: what matters for behavior is desires (Assumption 9b), not likes. This is discussed below.] But is Assumptions 9a true? While we often like what we sought and seek what we will like, you can seek what you will not like and like what you did not pursue.

The *mesolimbic dopamine system*, the MDS, plays a critical role. “Meso” is Greek for middle, the brain’s approximate center. “Limbic” refers to the part of the brain that includes the under chamber (the hypothalamus), the seahorse (the hippocampus), and the almond (amygdala). Included are clumps of neuron cell bodies in the VTA (*ventral tegmental area*) Fig. 7). [“Ventral tegmental,” from Greek, means “belly-covering”; it’s the area of the brain that sits on the floor of the midbrain, covering its belly.] These neurons have long axons that project into the prefrontal cortex and nucleus accumbens. Since they all emit dopamine, the clumps and their axons are called the MDS.

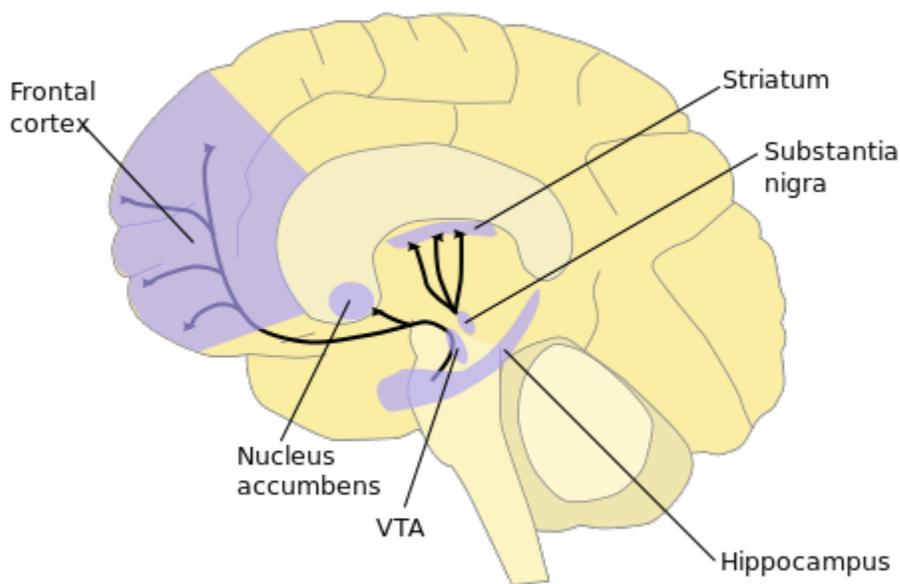


Figure 8, Chapter 4: The mesolimbic dopamine system (Kringelbach and Berridge 2012)

In Figure 8, note the axons, the black line starting in the VTA that branches out (in purple) when they enter the frontal cortex. When they fire, dopamine is released in the frontal cortex. If a rat presses a bar and gets a reward, axons in the MDS fire, and dopamine is released in the frontal cortex, causing the rat to quickly learn to keep pressing. If the bar is reward-less, pressing does

not lead to more pressing. Unless, when the rat presses, the researcher stimulates the rat's mesolimbic dopamine axons. [She implanted electrodes in these axons.] Then the rat quickly learns to keep pressing. It persists until it collapses in exhaustion. It continues even if each press causes a painful electrical shock. But! Suppose a rat is administered a drug that blocks the dopamine receptors in its frontal cortex.¹⁵ In that case, the rat never learns to push the bar even if every press causes a dopamine release, caused by either a real treat or the researcher electrically stimulating the axons.

In a 2006 study published in *Nature*, human subjects had to quickly and repeatedly choose between two symbols on a computer screen. After each choice, the individual was shown whether they had won £1. The probability of winning was always higher for one of the symbols (e.g., one was 60% and the other 40%). The objective was to see how quickly the subject figured this out. Subjects who had taken a dopamine-enhancing drug were the quickest to learn, subjects who took an inhibitor were the slowest, and those on a placebo were in between.

We, and rats, are motivated to seek/choose alternatives that cause dopamine to synapse with receptor neurons in our frontal cortex. This is a positive from an evolutionary perspective if the actions that produce a dopamine rush increase reproduction. The pursuit of sex releases dopamine. Cocaine, nicotine, and methamphetamines are fast ways to cause the dopamine to flow, causing many of us to pursue these drugs—including lab rats.¹⁶ But snorting coke does not typically confer an evolutionary advantage. [Keith Richards, an affluent father of many, is an exception to this rule.] The cravings for sugar can be as intense as the craving for cocaine. In our past, seeking sweet, high-calorie foods conveyed an evolutionary advantage. [Note that foods with high sugar concentrations easily ferment into alcohol.]

When and how much dopamine is released when one gets the reward? How much when an expected reward does not occur? How much when the individual observes a cue that makes a reward more or less likely?^{viii}

Initial answers came from an experiment with monkeys in the 1990s: if, when a light starts flashing, the monkey pushes the left bar, he gets juice. While learning the task, the dopamine

¹⁵ A drug that blocks the receptors on a neuron's dendrites from accepting (bonding with) the released neurotransmitter.

¹⁶ Here, "flow" is the firing rate of dopamine neurons. Dopamine's effectiveness is enhanced by how long it remains in the synapses (before reuptake) and the recipient neurons' receptiveness.

only flows when the monkey gets the juice. However, after he has learned to associate the light (a cue)¹⁷ with the juice, the dopamine flows not when he gets the juice but when the light flashes—when the monkey knows he will get the juice if he pushes the bar. fMRI (*functional magnetic resonance imaging*) indicates it's the same for us: after learning, dopamine isn't released by the reward itself (biting into and tasting a great burger), but prior when you know it will happen, when you realize a desire/want will be fulfilled, not when it's fulfilled.

What happens if, after the monkey learns to associate the light with the juice, the expected juice does not appear when he pushes the bar? There is still a burst of dopamine when the light flashes, but then the dopamine level drops below baseline right after the juice should have arrived—reducing his motivation to push the bar again. Alternatively, if more than the expected amount of juice is delivered, its arrival causes a second burst of dopamine—increasing his motivation to push the bar again. Dopamine release motivates behavior: one interpretation is if the reward is better than expected, more dopamine flows; if the reward is less than expected, less flows.

Summarizing: when the cue for a positive outcome occurs, there is a burst of dopamine. Then, when it's reward time, a second occurs if the reward is better than expected. Dopamine drops if the reward is worse than expected and stays the same if the reward is as expected. One theory (the error-prediction hypothesis) assumes that the magnitude and direction of the burst depend on whether the outcome was better than expected, worse than expected, or as expected.

An additional wrinkle: an experiment where there are different cues, each is associated with a different probability of the monkey getting the juice (e.g., 0%, 25%, 50%, 75%, and 100%). When cued, the amount of dopamine that flows depends on which cue is presented: the higher probability associated with it, the more information it conveys. See Figure 9 from an experiment by [Fiorillo](#), Tobler, and Schultz (2003). Their second graph, for example, shows the dopamine spike when the cue signals a 25% probability of juice, followed by a higher dopamine

¹⁷ A cue for a specific outcome is a stimulus you associate with that outcome. Pavlov trained many dogs to associate an initially neutral stimulus (e.g., a light coming on, or a sound) with either a positive or adverse outcome, changing the stimulus into what is now called a *Pavlovian cue*, a learned/conditioned cue. Put simply, a cue is a potentially-influential observation that occurs right before you make a choice. Many cues are consciously observed (you hear the phone ring), but others you only record subliminally. For a thirsty you, a cue could be you seeing someone enjoying a Pepsi rather than a Coke.

spike when the juice arrives. The five figures in Figure 9 are for a single neuron in Monkey “A”. p is the probability associated with the cue presented.

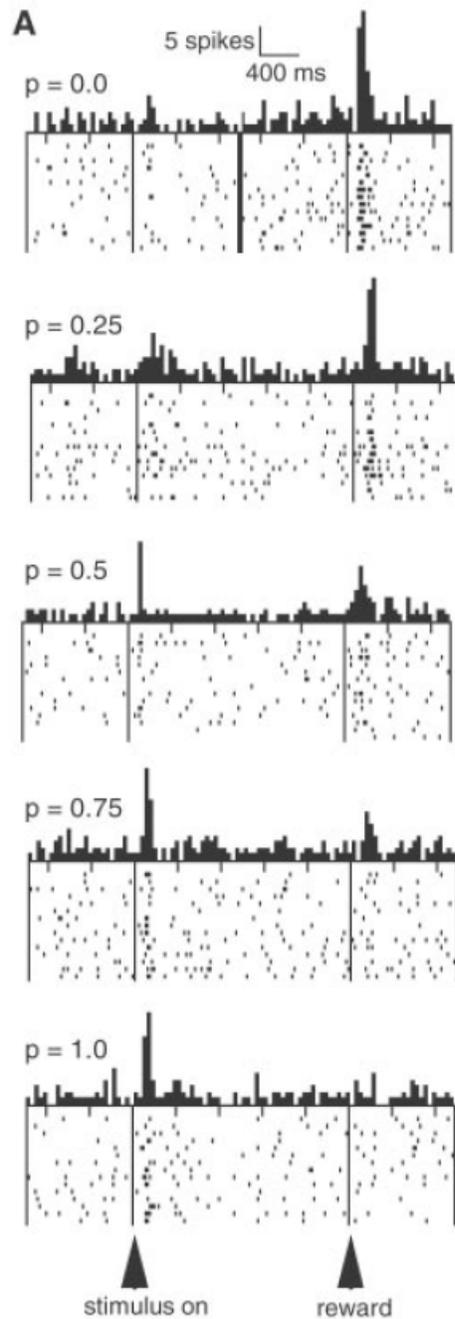


Figure 9, Chapter 4: from Fiorillo, Tobler, and Schultz, 2003

When the monkey sees the cue that signals the reward will arrive with certainty, there is a big spike when that cue appears, but there is no spike when the certain reward arrives (the 5th graph).

The lower the probability that a cue means a reward, the smaller the initial spike, but the larger the spike if the reward appears.

In a 2008 economics paper, [Andrew Caplin](#) and [Mark Dean](#) suggest dopamine “update[s] the ‘value’ humans and animals attach to different stimuli and actions, causing changes in the probabilities of acting certain ways. For them, the sequence of dopamine releases measures the value of the outcome relative to its expected value. A caution: don’t interpret their “value” to necessarily mean providing pleasure. A better word would be “reinforcing”, as in choices that result in the release of dopamine are likely to be repeated; the choice might, or might not, produce pleasure. We now turn to that issue.

Wanting versus liking: ^{ix}

It has been known for over sixty years that releasing dopamine causes wanting/seeking, leading to the reasonable conjecture that dopamine release causes both wanting and pleasure (liking). If correct, it would mean dopamine neurons are both wanting and liking neurons. This conjecture was accepted as true.

However, findings over the last few decades by the neuroscientist [Kent Berridge](#) and others contradict this conjecture. They find that the neural pathways for liking are separate from those for wanting and that choice is driven by wanting rather than liking. Choice is based on Assumption 9b (based on desire/wanting, not necessarily WB (9a))

The Berridge lab separated liking a food from wanting it: measuring liking using facial and mouth expressions (lip-smacking and Yuk face). [Using facial expressions to measure emotions goes back to Darwin’s 1873 book, [*The Expression of the Emotions in Man and Animals*](#).] If the neurologist blocks or destroys the relevant dopamine axons, a rat will not seek food: it will starve rather than walk across the cage for food. But the rat will eat with relish a sweet food placed on its tongue, expressing all the gestures a hungry carnivore exhibits when enjoying a cheeseburger—the food is liked but not wanted. Conversely, in rats whose dopamine axons have not been severed, stimulating these axons causes more dopamine to be released, which causes the craving for food to increase, but not the liking (no additional lip-smacking).

Research on humans and animals suggests that liking and wanting [seeking] are mediated by separate circuits in the brain. Berridge and his colleagues have, for example, shown that how much you like a sweet is independent of

how much dopamine is flowing. Drugs like antipsychotics that inhibit dopamine activity reduce people's desire for pleasure, but don't make that pleasure less intense (Bear et al.)

Consider the distinction between the sex drive and the pleasure of sex. As people age, their sex drive (their seeking) diminishes, but when sex occurs, it's enjoyed more than when they were eighteen. Alternatively, imagine yourself an adolescent taking an antidepressant that blocks the ability to orgasm. Yet, you're aroused and want sex—an awful, un-liked state.

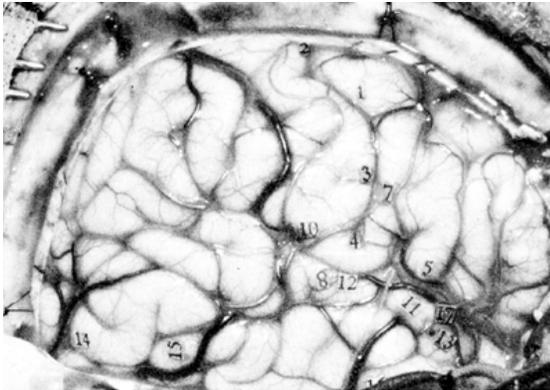


Figure 10, Chapter 4: The exposed brain of a conscious subject—Wilder Penfield
<http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/brain/brain-slides/penfield-brain.gif>

Researchers occasionally get to cut or stimulate neural pathways in humans, typically to reduce extreme seizures. If patients with stimulation implants can choose how often to self-stimulate each implant, they stimulate the sites that produce an experience they describe as a potential for pleasure that never comes. They stimulate fewer sites that deliver a pleasurable sensation.

My friend Bob recently paid a thousand dollars for a power meter for his bicycle (don't even ask). He had to have it but, as he would readily admit, having it would not increase his WB. A famous philosophical example demonstrates how desiring and liking can differ. Imagine an addictive drug that causes neither pain nor pleasure, but, once addicted, you wake up desiring it, a feeling neither pleasurable nor painful. Once addicted, if you don't take the drug within an hour of waking, you experience pain. If the drug were free and if your ordering of paths was determined by desiring rather than liking, you would choose to become addicted: why? — because you would fulfill a lot of desires for free. The example is from the Oxford philosopher [Derek Parfit](#) (1942-). You want your seeking and liking systems in sync. You don't want to be like [Marcel Proust's](#) fictional self, Charles Swann, in [Swann's Way](#) (1913), who spent his life desiring a woman he knew he would not like.

Remember why the distinction between wanting and liking is essential: if they do not line up, and choice is based on wanting rather than liking, what you choose will not always make you better off; it could even make you worse. Neuroscientists agree that wanting is driven by the mesolimbic dopamine system but somewhat disagree on how dopamine affects wanting. And the disagreement is critical to NBT, ethics, and the pursuit of happiness.

How much one wants alternative A versus B might depend on her recollection of how each previously affected her emotional WB or how much she expects each would affect her emotional WB. Or, it might depend on the desires/wants she is experiencing, which might be unrelated to the rewards she expects each alternative to produce.¹⁸ Advocates of choice theories that order paths based on experienced WB should hope that wants/desires are in sync with WB expectations and that individuals experience the WB they expected to experience.

The wanting/liking disagreement put simply: the *incentive-saliency hypothesis* vs. the *reward prediction-error hypothesis*

One hypothesis (*incentive-saliency*) is dopamine influences choice by **directly** influencing wanting. The other hypothesis (*reward prediction-error*) is dopamine **indirectly** affects choice by influencing **expected emotional WB**—And wanting is determined by expected emotional WB. There is data both in support and in conflict with each hypothesis. However, a bunch of the evidence is consistent with both.

A critical distinction is whether dopamine directly drives learning: the reward-prediction hypothesis says yes, and incentive-saliency says no. “Chapter 18: From Experienced Utility to Decision Utility” in the textbook *Neuroeconomics* critiques, compares and debates the two hypotheses. The chapter is jointly written by Kent Berridge and [John O’Doherty](#). O’Doherty, the director of the Caltech Brain Imaging Center, advocates for the reward prediction-error hypothesis and Berridge for the incentive-saliency.

¹⁸ Neuroeconomists, and interestingly non-economic neuroscientists, use the word *utility*: separating *decision utility* (the measure (of value) you use to decide which alternative to choose from *experienced utility* which is how emotional WB is affected by the alternative chosen. Experienced utility is what I have defined as emotional WB. So, terminology-wise, they separate what you will like from what you choose. They then separate decision utility into its possible components: *remembered utility* (how it felt last time), *predicted/expected utility*, and *wanting (incentive saliency)*.

The incentive-salience hypothesis

This hypothesis is associated with Kent Berridge and his lab. *Incentive salience* is fancy for serious wanting (a strong urge—conscious or not) compared to run-of-the-mill wanting. The *incentive* is to choose the most *salient* (striking, attention-grabbing).¹⁹ An initial motivator for the hypothesis was the Berridge finding. By removing or blocking dopamine, rats can be manipulated to not want what they would like (e.g., food). And they can be manipulated, with dopamine, to want what they will not like (e.g., a shock). This implies that wants are not always based on expected emotional WB. Summarizing, wanting is separable from liking, and wanting is driven by dopamine. This directed his lab to look for the brain part that produces pleasure. Two emotional/pleasure hotspots have been identified.

The sensation of pleasure

So, if dopamine isn't the pleasure neurotransmitter and the mesolimbic dopamine system, MDS, isn't the pleasure pathway, what causes the sensation of pleasure? Recent research by Berridge and others has identified two *hedonic hotspots* (the red spots in Fig. 10).

One of these lies in a subregion of the nucleus accumbens called the medial shell”; the nucleus accumbens is a small part of the MDS (see Fig. 7).

“A second is found within the ventral pallidum, the VP, a deep-seated structure near the base of the forebrain that receives most of its signals from the nucleus accumbens” (Kringelbach and Berridge 2012).

¹⁹ Salience, attention, and working memory: Paraphrasing the neurologist Amish Jha (2021), envision working memory as the whiteboard in your brain. On it is written what you are actively remembering. However, its capacity is limited, and the writings on it quickly fade if they are not continuously traced over. What is written is what you are attending to (the objects of your attention). They can be external stimuli (e.g., the stop light or a person's expression), internal (your thoughts, emotions, and long-term memories), or both.

There are millions of things that you could pay attention to, but your working memory can hold only a few of them, so your brain must ignore much of the information out there. So, what do you attend to? A complicated question! Attention is simply the mechanism that allows you to attend to some things and in-attend others.

Your attention can be highly focused like the beam of a flashlight or broadly focused like a spotlight, but it can't do both simultaneously. Either can be external or internal. And your focus can be self-directed or usurped. The scope of your focus and what it is attending to is constantly changing, which is usually a good thing from a survival perspective. For example, if you're lying in the grass taking a broad view of the night sky and then hear the rattle of a snake, your focus will immediately and involuntarily shift to narrowly focus on the sound and the image of the night sky will disappear from your working memory.

Salient objects are objects (or characteristics of objects) that draw our focus—often stuff you didn't consciously choose to focus on. The snake is a dramatic example: immediate threats to our well-being are highly salient—a good thing but only if the threat is real. Salient for you might not be for me, and its strength will vary with your physiological and emotional state. For example, seeing someone smoking a cigarette is salient to someone who has recently quit, and a Pepsi ad will push you towards drinking a Pepsi rather than a Coke even if you prefer the taste of Coke, but more so if you are thirsty. Salient objects pull your attention away from what you were attending to so often are viewed as distractions.

Why do we care about salience? We veer towards the alternative that draws our attention

The ventral pallidum is between the nucleus accumbens and the amygdala. In Fig. 10, the VP is a red dot, not the grey area where “ventral pallidum” appears.

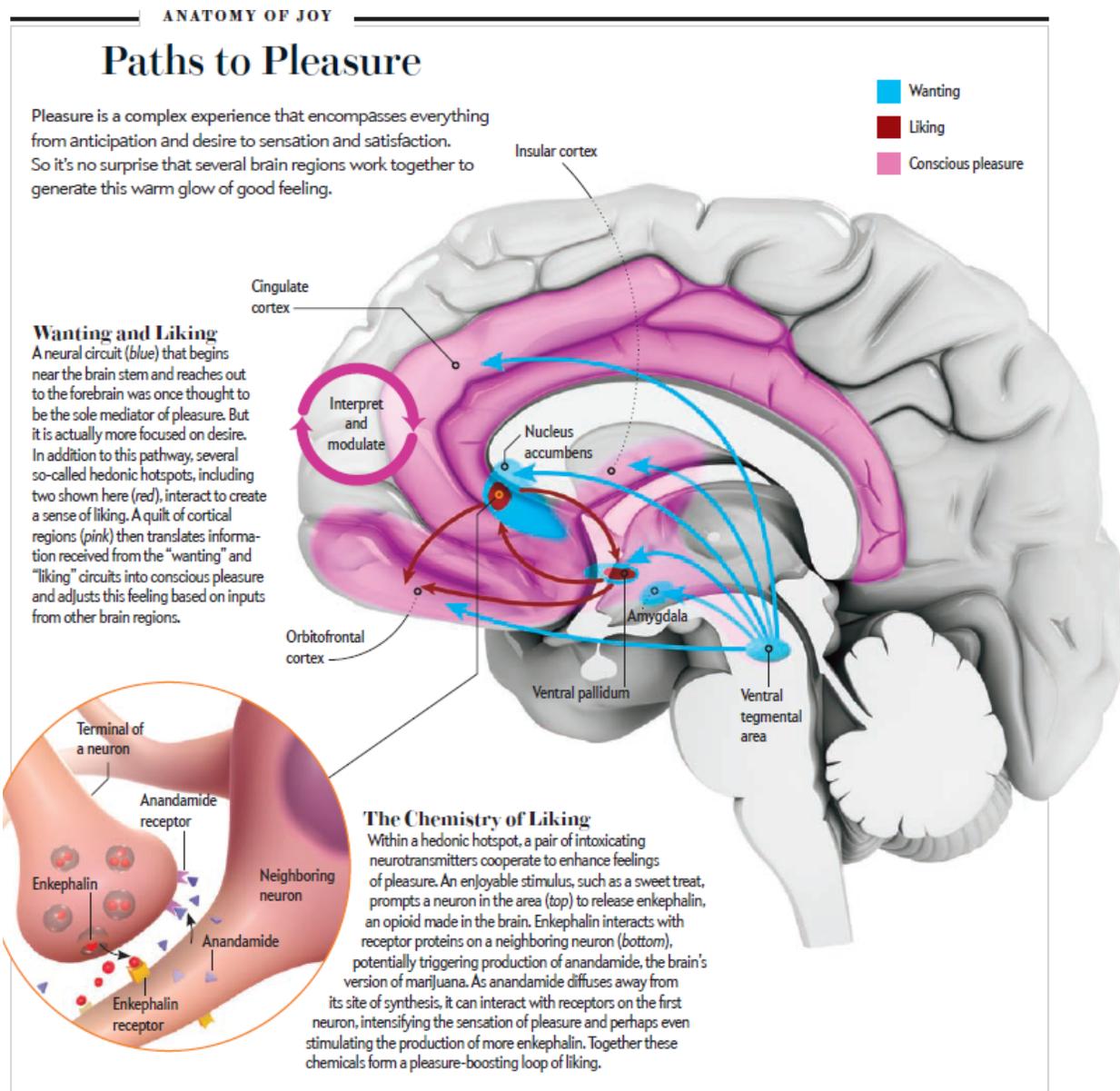


Figure 11, Chapter 4: Krangelbach and Berridge (2012) *The Joyful Mind*, *Scientific American*

Axons connect the ventral pallidum and the nucleus accumbens to each other, and both have axons projecting to the orbitofrontal cortex.

So, why are these two regions hedonic hotspots? Bathing their neurons with the neurotransmitter *enkephalin* makes a sweet taste sweeter, and so does the neurotransmitter *anandamide*. Anandamide is “the brain’s version of the active ingredient in marijuana.” [Ananda

is Sanskrit for bliss.] Releasing enkephalin causes anandamides to flow, which feedback to the axon terminal, causing more enkephalin to flow (a positive feedback loop). Destroying a rat's ventral pallidum makes a sweet taste yucky. [Many rats get brain parts destroyed in the interests of neuroscience.²⁰] Stimulating the ventral pallidum can make something yucky produce pleasure.

The VP and medial shell light up when people experience pleasure. Enkephalin and anandamide are opiates: pleasure is bathing the right neurons in the right drugs.

The axons projecting from the VP and the medial shell terminate in the *orbitofrontal cortex*, where the experience of conscious pleasure is produced (the *pleasure gloss*). The orbitofrontal cortex “adjusts this feeling based on inputs from other brain regions” (Kringelbach and Berridge 2012). Neural activity in the orbitofrontal cortex (observed by neuroimaging) is significantly correlated with the sensation of pleasure. There is evidence that stimulating this region can cause deep pleasure—where do I sign up?

Some people experience more pleasure because they consume more of what makes the enkephalin and anandamide flow. Lucky others inherit more and better enkephalin/anandamide synapses. Past experiences can also strengthen them. Twenty percent of Americans have a FAAH gene mutation that increases the flow of anandamide.²¹ The gene encodes to produce the enzyme *fatty acid amide hydrolase* (FAAH). This enzyme breaks down anandamide. The mutation causes less FAAH to be made, so less anandamide is broken down, so more of it flows. People and mice with this mutation tend to be less anxious and recover from bad experiences more quickly.²² So to experience pleasure, you want parents who are inherently good at

²⁰ The pleasure parts of human and rat brains are similar. But unlike a rat, a human has a large cortex, giving it the ability to ruminate about what they feel. But a rat does not need these higher brain regions to experience pleasure, and neither do humans. In the 1940s and '50s, thousands of people got frontal lobotomies (a procedure that intentionally destroys cortex), and while many then suffered from bad decision making, they didn't experience a decrease in their ability to experience pleasures—they still liked food and sex.

²¹ The probability of inheriting this gene mutation varies by ethnic group; while it's 21% for Americans of European descent it's 14% for Han Chinese living in China and 45% for Yoruban Nigerians (Friedman, referring to data from the [HapMaP](#) project).

²² Researchers ([Dincheva](#) et al.) recently inserted the mutated gene into mice to see how their behavior would be affected. Compared to normal mice they were more adventurous and had better connections between their amygdala (a fear-processing center) and their prefrontal cortex. Then the mice (mutated and normal) were conditioned to be afraid of a neutral tone. This was achieved by pairing it with an electric shock. Later all were repeatedly exposed to the same tone but without the shock. The mutated mice more quickly learned to not be afraid of the tone. The researchers got the same results when they compared humans with and without the gene mutation.

experiencing pleasure, and you want experiences that strengthen the pleasure synapses you were born with.

Most neuroscientists agree that wanting is typically determined by remembered and expected WB. Quoting from the text *Neuroeconomics*:

That is, we ordinarily desire an outcome to exactly the same degree that we predict the outcome will be liked, and most predictions about future experienced utility [WB] are based on how liked the outcome was in the past.

[There is a 2012 economic study that tends to support this conclusion. The study asks hypothetical choice questions; for example, “Would you choose a job that pays \$80K/yr. and you only have time to sleep 7.5 hrs./night or a job that pays \$140K with only 6 hours of sleep?” After you answer, the researcher asks which scenario would make you happier. The study estimates that in 83% of the questions asked, the subject chose the alternative they expected would generate the most happiness. If we make a leap and believe the subjects always chose their most-desired alternative, the study implies that 83% of the time, the alternative most desired was the one they expected to like the most.]

According to the incentive-salience hypothesis, sometimes (17%?) the intensity of the desiring/wanting for one of the alternatives is excessive. This happens when a cue for that alternative occurs when the individual is in a physiological state (hungry, aroused, tired, etc.) that increases their receptiveness to the cue. For example, you are thirsty and see someone enjoying a Pepsi, rather than a Coke. When a cue occurs, and you are receptive, the dopamine spike is compounded, causing you to want the alternative associated with that cue above and beyond what would be warranted by its expected effect on your emotional WB. You choose Pepsi over Coke even though you might have enjoyed the Coke more. [Remember that dopamine is released when a cue for the outcome occurs, which is before the outcome is experienced.] For example, the clock striking six is a cue for a hungry Giacomo that dinner will soon arrive and for me to drink wine. A second cue for us both is the smell of his burger cooking. These cues trigger the release of dopamine, directly and excessively increasing our desire for the alternative associated with these cues, making me overly inclined to drink and Giacomo to stand next to his bowl.

Experiments by [Cindy Wyvell](#) in the Berridge lab have measured the combined effect of a cue and physiological receptiveness. “Rats were first trained to press one of two levers to obtain sucrose pellets. They were separately conditioned to associate a Pavlovian cue (30 sec. light) with free sucrose pellets” (Wyvell and Berridge 2000). Then on test day, the rats had

different amounts of amphetamines injected into their nucleus accumbens, “putting the rat’s brain in an elevated state of dopamine activation.” The rats were then set free to press, or not press, one or both levers. No sucrose pellets were rewarded or given, but the cue was presented at intervals. All the rats pressed both levers, but more the lever that they had associated with the pellet. When the light went on (the cue), they all pressed the sucrose lever more than when the light was off but did not increase how often they pressed the other lever. The number of additional sucrose-lever presses increased with the amphetamine dose. Since the amphetamines did not affect the number of presses between cues, they only influenced cue-specific behavior. With a separate test, she showed that the amphetamines did not increase liking, so the rate of lever pushing was not increased by liking, only by wanting. To summarize, the cue increased sucrose-lever pushing and more so the more the rat’s physiological state was amped up. The more the rat was amped up, the more salient the cue.

One might view incentive salience as expected WB multiplied by a constant k , so when k equals 1, the individual only wants what they expect to like, but if k is more, or less than 1, wanting deviates from liking. And the extent a cue affects k depends on the individual’s physiological state. The implication for NBT is that what the individual chooses will not always make them better off. And this is more likely to occur when they experience cues and are in a receptive physiological state. You often make flawed choices when you are tired, aroused, intoxicated, etc., and a cue for the bad alternative appears.

Consider the 83% estimate noted above, keeping in mind that the choices were hypothetical in that study. There was also no cueing (no pictures of comfortable beds were flashed on the screen). And we can presume the subjects were at their physiological baseline (e.g., not tired). This makes me conjecture that in the real world (where cues are numerous and people are often not at their physiological baseline), wanting and expected liking line up much less than 83% of the time. Whether getting it wrong at least 17% of the time is a big deal is a matter of interpretation. Some choices (like Coke or Pepsi) probably don’t have much effect on long-term WB, but others do (e.g., having unprotected sex without birth control).

Add references to other salience models of choice. In a recent paper, Franz and List proposed a choice model where choice is based on the salience of each alternative (more salient alternatives are ranked higher), so it is a model consistent with the incentive-salience hypothesis. They formulated the model to explain why you have the ordering you have and, maybe more

importantly, why your ordering changes over time (the salience of an alternative can vary even as its effect on WB remains constant). As previously mentioned, there isn't much literature in economics on the creation of orderings.

The reward prediction-error hypothesis:

This hypothesis, proposed by [Wolfram Schultz](#) at Cambridge, is dopamine flows when the expectation of future rewards needs correcting. Specifically, it flows whenever new information indicates the current expectation is wrong. New information comes in two forms: the magnitude of your last reward and sometimes a cue before the next reward is expected. A temporary increase in firing rate (a positive spike) signals expect more (your current expectation is too low). A negative spike signals expect less. No spike signals no need to correct.

When a cue associated with an alternative appears, it tells your brain that you should associate a larger reward with that alternative. And, the taller the spike, the more you need to revise upward the expected reward. (see Figure 9). If there is no cue, there is no pre-reward spike because, since the last reward, there is no new information.

A positive dopamine spike when the reward arrives signals that it was better than expected, so you should expect even more going forward. A negative spike when the reward arrives signals the opposite.

In summary, the spikes start a process that brings expected WB more in line with the WB that will be experienced. [Recollect that the axons of these dopamine neurons project widely.] Discussions of this hypothesis typically use the word *learning*: you learn, through reinforcement, better what to expect. When the prediction error is large, there is more to learn.

The primary evidence for the hypothesis that dopamine flow is a gauge of the prediction error is that the dopamine spike when a cue is presented is larger the more the cue guarantees the reward. And, the observation that the magnitude of dopamine spike when the reward occurs (up, down, or not at all) depends on how much the reward deviates from the expected reward. These two observations are evidence for the hypothesis in that they are what it predicts. However, they are not proof because they are also consistent with other hypotheses.

A problem with testing the reward prediction-error hypothesis is that while neuron firing rates can be directly observed (what Fiorillo did), prediction errors can't. So they are estimated

based on assumptions, assumptions that not everyone accepts. Many studies show that firing rates are correlated with the estimated prediction errors, but no current studies prove that firing rates are correlated with the actual prediction errors.

In addition, Berridge points out that the above experiment would produce results consistent with the prediction-error hypothesis even if it is wrong. This will happen if the incentive-salience hypothesis (his hypothesis) is correct. And, if during the experiment, the physiological states of the subjects are all the same and constant (e.g., all the rats are equally thirsty or not). Wanting will be mislabeled as prediction error. So, according to Berridge, a distinction between wanting and prediction error will only appear if there is variation across subjects in their physiological states. But studies have not included such variation.

A finding that conflicts with the reward prediction-error hypothesis is that rats can learn without dopamine. For example, even after most of their dopamine neurons have been destroyed, they can learn to dislike a taste they liked (Berridge and O’Doherty 2013).²³

That said, many neuroscientists support the reward prediction-error hypothesis.

The reward prediction-error hypothesis implies that dopamine does not directly drive wanting but indirectly influences it by updating expected emotional WB. With respect to Assumption 9a (paths are WB-ordered) and Assumption 9b (paths are want/desire ordered), the reward prediction-error hypothesis and its supporting data are consistent with paths being ordered based on desiring/wanting (9b)—not in WB (9a). But where the desire is correlated with expected WB. A neurological mechanism continuously uses new information to update expected WB. If the reward-prediction error hypothesis is correct, your ordering is based on expected WB, not actual WB, so things can still turn out badly. You expected the beautiful house with the beautiful wife would make you happy, so you got both, but you might have been better off living elsewhere with someone else—a point made by the [Talking Heads](#).

Summarizing:

Both hypotheses hypothesize that choice is driven by wanting/desiring, often by the expected WB associated with each alternative. The link between wanting and expected WB is weaker if

²³ But it seems that some types of learning do require dopamine (Berridge and O’Doherty 2013).

the incentive-salience hypothesis is correct. Specifically, the link weakens when cues proceed the choice, and the chooser is in a physiological state that makes her hyper-receptive to the cues. In such cases, the incentive-salience hypothesis predicts you will often choose something other than what you would have liked the most. In the reward-prediction hypothesis, dopamine drives learning (changing expected values when new information arrives). In incentive-salience, dopamine drives wanting, not learning.

The neurological data supports the conjecture that choices are based, at the decision point, on wants rather than likes (Assumption 9b, not 9a). I am no expert in judging between the two hypotheses. Still, I find the incentive-salience hypothesis appealing and intriguing because it is more in line with models and data from psychology that indicate we often act against our self-interest (see Chapters 6 and 7). It is also in line with literature on addiction that suggests how much you want something does not necessarily reflect how much you will like it. In one interesting study (Leyton et al., 2005),

cocaine users were given a drug that lowered their dopamine levels. In the lowered dopamine state, cues indicating the availability of the drug were rated as less desirable. When given the drug, however, the users' feelings of euphoria and the rate of self-administration were unaffected. That is, with reduced dopamine, study participants still liked the drug in the same way (reinforcement was unchanged), even though they didn't particularly want it. (Gazzaniga, Ivry, and Mangun 2014).

Or the reverse, the recovering addict encounters a positive cue while in an elevated physiological state, which makes them temporally want the drug more than they will like it. So, they take it even though they know that their past use was a mistake and know that taking it now is a mistake ([Bernheim](#) and [Rangel](#) 2004)

As a choice modeler welfare economist, I find the incentive-salience hypothesis both more intriguing and more troublesome because, if correct, you and I don't always maximize our WB. (The reward prediction-error hypothesis is more in line with us each maximizing our WB, so I guess that I should hope the incentive-salience hypothesis is eventually rejected. That said, seeing apple carts upset spikes my emotional WB—when I don't have too many apples in the game.)

What about economists and ethicists who don't care about WB? If you are an economist who only wants to predict behavior, Assumption (9b) is consistent with both hypotheses and the neurological data.

If you are an ethicist (economic or otherwise) for which want/desire fulfillment is the ethical objective (rather than maximizing WB), go with 9b: it is consistent with both hypotheses

(both imply 9b). And it means that individuals pursue want fulfillment, the ethical objective, from their perspective.

Findings in neuroscience are coming fast and furious, and new research could negate or confirm the results reported above. One possibility is that dopamine plays two roles: in one part of your brain, it might be a gauge of prediction error, and in another, a gauge of salience. Gazzaniga, Ivry, and Mangun (2014) describe a study by [Matsumoto](#) and [Hikosaka](#) (2009) that recorded firing rates for dopamine neurons in the brain stem:

One subset of dopamine neurons responded in terms of valence [reward or punishment]. These cells increase their firing rate to stimuli that are predictive of reward and decrease their firing rate to aversive stimuli (Figure 12.19a). A greater number of dopamine neurons, however, were excited by the increased likelihood of any reinforcement, independent of whether it was a reward or a punishment, and especially when it was unpredictable (Figure 12.19b). The first response class is similar to what would be expected of neurons coding prediction errors, the second to what would be expected of neurons coding salience or signaling things that require attention.

In closing, I leave it to the reader to decide when and if it would be better for NCT to be explicit about the basis for ordering. My preference is to be explicit, and the neurological evidence indicates it is based on wants, not likes.

A more macro view: pleasures, emotions, temperaments, and happiness^x

Consider four sorts of happiness: (1) a pleasurable sensation (the taste of Valrhona chocolate, coming in from the cold), (2) a positive *emotion*, (3) being in a good *mood*, and (4) having a good *temperament*. Happiness is one of the *effects* experienced by humans and a few other animals. Affect is the general term for emotions, feelings, and moods.

Emotional WB is a suite of emotions, but what is an emotion? And how are they formed?

An emotion is a conscious experience. As was discussed in Chapter 1, while there is no universally agreed-upon definition of an emotion, psychologists generally agree that it has at least three parts: a “physiological reaction to a stimulus, a behavioral response, and a feeling.” Feelings are subjective in that I can’t know how what I am feeling differs from what you are feeling or what a bat, or another animal, is feeling. Giacomo is happy (or appears so) when playing in the snow, catching things, and pouncing on stuffed animals—the sight of an airborne stuffed animal is a stimulus. Most of us imagine slugs are incapable of emotions, but we don’t know for sure.

Emotions and moods are of limited duration, and the duration of an emotion is shorter than a mood but more intense. And, unlike emotions, moods don’t have an “obvious trigger”.

The fear experienced when you hear a snake in the grass is an emotion, whereas being anxious for long periods, for no reason, is a mood, a mood disorder. A semi-permanent mood is a temperament, a component of your personality.

Emotions such as disgust, fear, anger, happiness, sadness, and surprise are considered basic emotions: they are, arguably, hardwired, shared by everyone, and characterized by the same facial expression in every culture. The Big-Bad Wolf knew Goldie Locks was afraid when he saw the whites of her eyes.²⁴



Figure 12, Chapter 4: The faces of emotions

More complex emotions include bored and embarrassed; they are learned and vary culturally. Romantic love is a complex emotion, one not associated with any facial expression. Current thinking is that there isn't a one-stop emotion center in your brain—not surprising given all the different emotions and that many brain regions serve multiple roles—but the amygdala and hypothalamus play a significant role in many emotions. Emotions have two dimensions: positive or negative (pleasant or unpleasant) and intensity. For example, a little sad can be pleasant.

²⁴ If B.B. Wolf had had his two amygdala surgically removed, he would likely not have recognized Goldie's fear: individuals with severe amygdala damage do not process fear as do you and I do, in part because they do not automatically focus on eyes (See [Adolphs et al. 2005](#)).

Emotion formation

If the goal is to make someone happier, we want to enhance their ability to form the emotion happy. Emotions have three components: a physiological response, a behavioral response, and a feeling. There are several contrasting theories of emotion formation; they differ in the order the components occur—which comes first. Two primary and historical theories of emotion are the *James-Lang Theory* and the contrasting *Cannon-Bard Theory*. Put simply, The James-Lang theory hypothesizes that the first response to the stimulus is physiological (e.g., sweating, increased heart rate), and the conscious feeling we experience—the emotion experienced—is a reaction to these physiological changes. So, the unpleasant sensations I experienced when I arrived at the home of my high-school prom date (terror and anxiety) were caused by my body's sweating and shaking and my inability to swallow. [[William James](#) (1842-1910) was both a noted philosopher and the father of psychology in the U.S. He suffered from depression and insomnia and used chloroform as a sleep aid. His brother was the novelist [Henry James](#). [Carl George Lang](#) (1834-1900) was a Danish physician. They separately, but within a year of each other, proposed the theory.] In contrast, the Cannon-Bard Theory ([Walter Cannon](#) and Phillip Bard) hypothesizes that your brain processes the physiological change and the conscious feeling simultaneously. But the feeling processes faster, so first you feel nervous and then sweat.

There are other theories of how emotions are formed. One class is *Appraisal Theories*: the sexy grizzly bear sitting alone at the bar winks at you; first, you do a quick harm/benefit risk appraisal (often unconsciously).²⁵ Your appraisal has two components: what does the wink mean for me, and can I cope? Based on your appraisal, you feel fear, sexual arousal, disgust, or whatever. This feeling might be associated with physiological changes, but that isn't required. Then, based on which feeling is experienced, you run, offer to buy the bear a drink, or remain alone in the corner with your Bud Light. Appraisal theories explain why you and I might experience different emotions when we are winked at. Appraisal theories sharply contrast with the James-Lange theory, where it all starts with a physiological response to a stimulus.

The Schachter and Singer Theory (1962) blends Appraisal theories with the James-Lange theory. It hypothesizes that (1) emotions are influenced by both the physiological and cognitive reactions to the stimulus, (2) the physiological response sometimes proceeds the thinking part, and (3) people identify the emotion and gauge their degree of arousal based on how

²⁵ William James liked scary-bear examples.

physiologically they are revved up. Their theory was proposed to explain the results of their famous experiment where they injected subjects with adrenaline (a potent stimulant), telling half of them that they were being injected with adrenaline. The other half were told they were being injected with vitamins. Each subject then interacted with an individual who behaved either angrily or euphorically. [The subject thought the other person was another subject, but it was a research assistant.] After interacting with the other individual, subjects told they were injected with vitamins attributed their emotional state to the individual's behavior. Those told they were injected with adrenaline attributed their emotional state to the adrenaline.

Another theory, the *Construct Theory* of emotions ([Lisa Barrett](#)), hypothesizes that many emotions are cultural constructs. In this view, embarrassment, for example, isn't a hard-wired emotion but instead constructed.

in the same way that they see color or the way they perceive behaviors in others... Color is a continuum of wavelengths, yet ... experienced categorically....

Physical movements are continuous in the same way that the visible light spectrum is continuous. People are constantly moving and doing things—that is, they are constantly engaging in “behavioral actions.” Yet, behaviors are perceived as discrete acts ...

People are compelled by their own experiences to believe that emotions exist as natural-kind entities, yet a century of research has not produced a strong evidentiary basis for this belief. To date, there is no clear, unambiguous criterion for indicating the presence of anger or sadness or fear. (Barrett 2006)

According to Construct Theory, the bear's wink causes a physiological change. Then you assess your past experiences with bears (winking and not) and other bear winkers. Then you categorize the experience, in your language, as fear, attraction, embarrassment, or whatever. Then you act.

Summarizing the theories of how emotions form, the jury is still out, but all start with a stimulus that causes a physiological change, a behavioral change, and a feeling.

While happy is considered a primary emotion, it can have different forms: joyful, gleeful, delighted, euphoric, etc. Happy is either one type of emotional WB or a synonym for it, and how it arises depends on which theory of emotional WB you subscribe to. Happiness that persists without an obvious trigger is better described as a mood than an emotion, but, for simplicity of exposition, include a happy mood in emotional WB.

Your brain in a mood

A necessary condition for feeling happy is the absence of negative emotions and moods such as anxious or depressed. Anxiety and depression might or might not have an identifiable trigger.

Most of the mood research is on mood disorders, specifically anxiety disorders and depression.

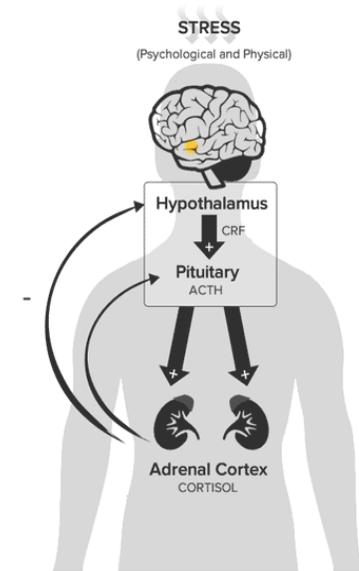
Anxious

For many, unsolicited anxiety is a barrier to emotional WB. In addition to precluding happiness, it affects our behavior: we avoid situations that provoke it; we do things to reduce it, and we don't do things we would have done if we weren't anxious. When I called Patricia to ask her to the prom, I was so anxious I had a flow chart of what to say in response to everything she might say. When she said yes, the flow chart became useless, and I hung up in a panic. Our prom was a disaster.

When you perceive danger, you experience fear and the flight-or-fight response (our response to stress), making you more likely to survive the threat. Your amygdale receives an external stimulus (you hear a gunshot) that shouts danger. The warning is passed through your hypothalamic and pituitary glands to your adrenal glands, which react by pumping *cortisol* (a stress hormone) into your bloodstream.²⁶ Specifically, your hypothalami release CRH, the *corticotropin-releasing hormone*; it binds with receptors in your pituitary glands. This causes each pituitary gland to release ACTH (the *adrenocorticotropic hormone*); it binds with receptors in your adrenal glands. The cortisol channels energy away from needs that aren't immediate, such as growth, digestion, and immunity. This hypothalamus/pituitary/adrenal system is called

²⁶ Cortisol shouldn't be confused with *adrenaline*, another hormone released by your adrenal gland when a threat is detected; adrenaline increases heart rate, sweating, and muscle tension.

the *HPA axis*. This neurological and hormonal pathway that kicks in when your body registers the gunshot is your *excitatory system*—the pathway that generates excitement. See Figure 13.



As part of the inhibitory system, the hippocampus monitors the amount of cortisol emitted. If it's deemed too much, a message is sent to the pituitary gland through the hypothalamus. The message is to cut down on the cortisol release. Specifically, the message to inhibit is sent when GABA neurons release the neurotransmitter GABA. In the post-synaptic neuron, the GABA attaches to neurons in the hippocampus with GABA receptors, causing these neurons to fire. Their firings start the chain of events that lead to less cortisol being released, which is a negative feedback loop.

Both the excitatory and inhibitory messages pass through the hypothalamus. Ideally, both systems are balanced, and you only freak out when freaking out is appropriate. However, genetics, prior experiences, and learning can cause a stress response when there is no danger. These experiences can be low-level and chronic (a *generalized anxiety disorder*) or short and extreme in short (a *panic disorder*). Either way, the amygdala works too much or the hippocampus too little.

Three drugs reduce anxiety: benzodiazepine (e.g., Valium), alcohol, and SSRIs. They do this by modifying neural synapses. Valium and alcohol temporarily reduce stress by enhancing the inhibitory system; specifically, the GABA receptors in the hippocampus become more receptive to GABA after the Valium or alcohol has attached to them. Like many who suffer from a generalized-anxiety disorder, I make my GABA receptors more receptive by self-administering wine. A nickname for Valium is “[Mother’s little helper](#),” which is the name of a 1996 Rolling Stones song about how Valium got your mom through the day.

No one is sure why SSRIs reduce anxiety, but they often do. Selective Serotonin Reuptake Inhibitors keep the released serotonin in the synaptic gap longer by inhibiting its

reabsorption by the axon terminal that emitted it. Most of the serotonin-emitting neurons in your brain have cell bodies in one of the nine Raphe (ridge) Nuclei, an area at the base of your brain (see Figure 14). Their axons project into many parts of the brain (one axon can synapse with hundreds of thousands of neurons spread all over the place, and the effects of serotonin releases can last for hours). These neurons influence many brain activities, including emotions and moods such as happiness, fear, and aggression.

While an SSRI immediately causes serotonin to remain longer in the synapses, it, unlike alcohol and Valium, does not immediately reduce anxiety; that can take weeks. (This is why your doctor might prescribe Valium for a few weeks even though you have started on Prozac.) So, it's not the increased serotonin per se that reduces the anxiety. One relevant finding is that the increased serotonin causes, in the hippocampus, an increase in the number of receptors for the neurotransmitter glucocorticoid. Speculation is this strengthens the inhibitory system so reduces the release of cortisol.

Eating chocolate causes serotonin to flow. You can reduce anxiety by learning not to be inappropriately stressed. When there is a threat, other things are also happening around us; these can become misleading cues: things that aren't dangerous (*benign stimuli*—maybe it was raining when you heard the gunshot). You learn, incorrectly, to go into stress mode when one or more of these other things occur, even when there is no threat—you have learned to associate rain with danger. Like a rat that experiences a sight or sound cue paired with something unpleasant, we learn that when that sight or sound occurs, we imagine something terrible will happen, and our mind and body prepare. *Cognitive Behavioral Therapy* (CBT), and its extensions, can break the link: you learn to no longer associate the stimulus with the threat. Like drugs, re-learning changes the number and effectiveness of your synapses.

Besides unwarranted anxiety, there is warranted stress: fear and anxiety caused by being at risk. Many people are in circumstances with a high probability of injury, death, sickness, and hunger. Warranted anxiety can be reduced by reducing these probabilities. For example, knowing the kids won't starve will lessen the anxiety of most parents. [Independent of how you feel about Obamacare, assuring access to healthcare will reduce anxiety. On the other hand, Obamacare has put those who hate it in a bad mood. An important question is how long these bad moods will last. The anxiety reduction will continue if the access continues.]

The Diffuse Modulatory Systems of the Brain

- The Serotonergic Raphe Nuclei

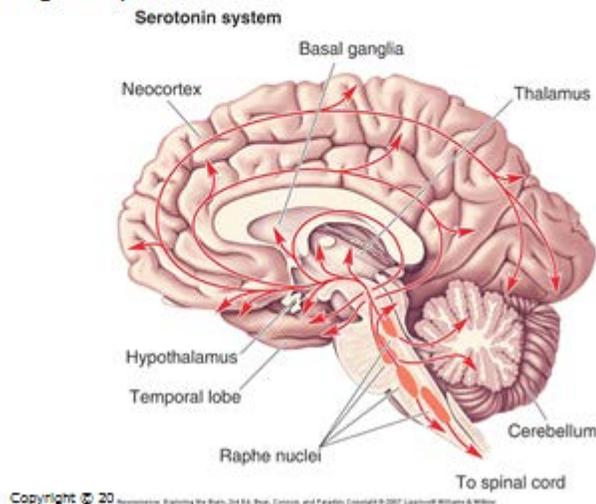


Figure 14, Chapter 4: The diffuse serotonin system

Depressed

If you are depressed, you are not happy, and we all occasionally experience depression. Unfortunately, many suffer from chronic depression, making it a leading cause of disability and unhappiness. “Perhaps as many as 20% of the population will suffer a major, incapacitating episode of depression during their lifetime” (Bear, Connors, and Paradiso 2007). Like anxiety, depression causes behavioral changes: a shutting-down, a decreased responsiveness to stimuli—an ice-cream cone is no longer worth the trouble.

Anxiety and depression often go together, and SSRIs are effective for anxiety and depression. Why are they effective for depression? One answer is the *monoamine hypothesis of mood disorders*. There are two monoamine neurotransmitters: serotonin and *norepinephrine*. Explaining: your brain contains three mood-critical *diffuse modulatory systems*. Figures 11 and 14 illustrate the diffuse modulatory systems for dopamine and serotonin. Each consists of a clump of cell bodies, all with axons spewing their neurotransmitter far and wide. The serotonin system modulates the diffusion of serotonin; the dopamine system diffuses dopamine. There is also a diffuse modulatory system for norepinephrine; its clumps are in your brainstem.

The monoamine hypothesis is that depression results from a deficit in either the serotonin or norepinephrine diffuse modulatory system. The evidence that SSRIs work as an antidepressant through these two diffuse systems is (1) SSRIs keep the effects of serotonin release working longer (inhibit re-uptake), (2) another effective antidepressant drug (*imipramine* (Tofranil)) inhibits the re-uptake of both serotonin and norepinephrine, (3) a few drugs for other ailments cause depression by messing up the release of serotonin, and (4) drugs that inhibit the enzyme MAO—MAO destroys serotonin—are antidepressants. These drugs are called MAO inhibitors; they effectively treat depression but are now rarely prescribed. They can cause severe high blood pressure when consumed with red wine and aged cheese.

A different but overlapping theory of depression is the *diathesis-stress hypothesis*. Depression, like chronic anxiety and other mood disorders, is caused by hyperactivity in the hypothalamus/pituitary/adrenal system, the HPA axis. This system is implicated in all theories of anxiety and depression. The adjective “stress” because stress distorts the HPA; the adjective “diathesis” because mood disorders have a genetic component; some people are predisposed. “Animal studies are highly suggestive” (Bear, Connors, and Michael Paradiso 2007) that hyperactivity in the HPA results in depression. There is also strong evidence that infants who experience nurturing and stimulation are better able to respond to stressors in adulthood, making them less likely to develop mood disorders. Lousy pre-natal care and neglecting the baby can lead to an adult with a hyperactive HPA.

In summary, genes, monoamines, and early childhood experience are implicated in mood disorders. So, what does all this have to do with happiness and emotional WB? A way to increase emotional WB in adults is to make sure babies are nurtured and stimulated. In adults, depression and chronic anxiety can be treated with CBT and drugs. Electroconvulsive therapy (inducing a seizure by zapping your brain with electricity) can effectively treat major depression, but it also zaps your short-term memory—think of it as setting the reset button. **(maybe footnote here noting recent research wrt psychedelic drugs.)**

Pervasive happiness (a happy person) and WB

The happiness from an experience/stimulus is typically fleeting, and these experiences are intermittent, not pervasive. So, neurologically, what causes persistent emotional WB? Not being

chronically anxious nor depressed is necessary, but not sufficient. Having strong enkephalin/anandamide synapses also helps.

By many definitions, WB requires not just the experience of happiness (emotional WB) but also a sense of meaning in your life (life-satisfaction WB). One relevant finding is that high reported happiness is correlated with feeling you have a meaningful life, suggesting a neurological connection between the two.

The neurobiology of persistent happiness is sketchy, and the neurobiology of meaningfulness is an informed scribble. But experiencing meaning in your life must have a neurobiological foundation. For example, there is a suggestion that the brain mechanisms that give you pleasure from chocolate also cause persistent meaningfulness.

Based more on observation than on neurobiology, we each have our own quasi-set level of emotional WB/happiness, which we can easily deviate from in the short-run but not in the long run—a hypothesis Adam Smith subscribed to. This is discussed in Chapter 5. Studies on identical twins separated at birth indicate that happiness has a genetic component. One large study finds that approximately 33% of the variation in happiness across individuals is explained by genetic variation. The paper “Genes, Economics, and Happiness” finds that a feature on a serotonin transporter gene positively correlates with reported happiness.²⁷ People that inherit this feature grow stronger happiness-related synapses.

Defining happiness in non-biological ways is difficult—but, like porn, you know it when you see

Dictionary synonyms for modern happiness include pleasure, joy, bliss, felicity, and contentment, but these are not all the same. The physical feelings associated with contentment differ from what it feels like to laugh. Happiness varies on multiple dimensions, and for each, there is the question of intensity and how long it lasts: short and intense or less intense but longer-lasting. Possible antonyms include unhappiness, sadness, grief, pain, and depression.

²⁷ It's the 5-HTT gene and “much is known about the way different versions of this gene influence transcription, metabolism, and signal transfers between genes. . .efficient variants of this gene have been shown to moderate the influence of stress on depression, . . .and [are] linked to optimism” (De Neve et al.). Keep in mind that it's unlikely a single gene can explain all the variation in a complex feeling, like happiness, and that huge samples are required if the researcher hopes to find links between specific genes and behavior. Quoting further, “We do not claim that 5-HTT determines happiness, nor do we exclude the possibility that several other genes may play a role in accounting for the influence of genes on happiness.”

Which words should be synonyms or antonyms is a tricky business. For example, pain often leads to unhappiness, but pain is neither necessary nor sufficient for unhappiness.

Happiness is the sublime moment when you get out of your corset at night—Joyce Grenfell (1919-1978)—actress, comedienne, and singer-songwriter.

What we call happiness in the strictest sense of the word comes from the (preferably sudden) satisfaction of needs which have been dammed up to a high degree—Sigmund Freud.

Most economists talk about preferences, not happiness nor WB

Recollect assumption 9a vs. 9b (an ordering based on WB vs. desire/wanting). Given that pleasure (liking) and desiring (wanting) have different neurological pathways that are not necessarily in sync, whether your ordering is based on WB (9a or desire (9b) is important. Either way, whether path *a* is ranked higher than path *b* depends on their respective electrical/chemical brain mixtures.

When economists think about you being better off, they typically think of this being achieved by you experiencing a higher-ordered path of goods and experiences. But what if there are other more direct ways to feel better—drugs, for example? I ask my students whether they would take a drug that would make them happier. Most say "No." [I'm always surprised, given that many college students are prescribed SSRIs, and most drink alcohol.] When I ask for an explanation, the first response is typically, "There would likely be bad side effects." When I ask what they would do if they knew there were no undesirable side effects, many students continue to say they wouldn't take the drug. But their reasons change to, "It would be wrong/unnatural/dishonest to achieve happiness in this way." This implies they care not only about happiness but also about the process that achieved it—an outcome, meaning they care about both consequences and process (see Chapter 2 and Part III).

Another way to make you better off is to change your ordering so the path you are experiencing moves up in the orderings. Economists typically avoid this possibility, but marketers embrace it—also Buddhists.

Is an economist concerned with policy a WB doctor?

Most people, including most economists, don't think of economists as WB doctors. Whereas mental-health therapy typically tries to increase WB one patient at a time, macroeconomic policy tries, I imagine, to simultaneously increase the WB of thousands, if not millions. As an

undergraduate, I switched from psychology to economics. The welfare objective of clinical psychology and welfare economics is the same, to improve the lot of their patients. [Of course, not all psychologists and economists care about their patients; many simply study behavior models for the sake of the models, not the patients.]

If the goal is to increase WB, my goal as an environmental welfare economist should be to manage the environment to achieve more WB, not necessarily preserve or clean it up. Trashing the environment can be fun, but I have to consider that trashing it now will make the future unhappy.²⁸ If better is defined explicitly as happier (greater emotional WB), the goal of all economic policy should be to make the members of society happier. Achieving it by implementing policies encouraging encephalin, anandamide, and other happiness neurotransmitters to flow.

Welfare economists who accept Assumption 9a are WB doctors, but not necessarily emotional WB doctors.

This chapter considered “What is happiness?” and outlined ways to reduce unhappiness but didn’t consider whether more happiness is morally superior to less. That conversation was started in Chapter 2 and continues in Part III. Next, Chapter 5 looks at the evidence, pro and con, for whether consuming more increases or decreases, WB.

ⁱ Sources:

1. Thomas Carlyle (1963) *Past and Present*, Boston, Estes and Lauriat Publishers (first published in 1843). http://books.google.com/books?id=QqoPAAAAAYAAJ&printsec=frontcover&source=gb_s_ge_summary_r&cad=0#v=onepage&q&f=false
2. Jim Holt (2006) Oh Joy: a review of *Happiness: A History* by Darrin M. McMahon, *N.Y. Time Book Rev.* February 12. <http://www.nytimes.com/2006/02/12/books/review/12holt.html>
3. Darrin McMahon (2005) The pursuit of unhappiness, *N.Y. Times*, December 29. <http://www.nytimes.com/2005/12/29/opinion/29McMahon.html>

ⁱⁱ Sources are:

1. Myles Burnyeat (2007) Myles Burnyeat on Aristotle on Happiness, *Philosophy Bites*. <http://philosophybites.com/2007/11/miles-burnyeat.html>
2. Roger Crisp (2015) SEP on *Well-being*. <http://plato.stanford.edu/entries/well-being/>
3. The Internet Encyclopedia of Philosophy (IEP) on *Epicurus*. <http://www.iep.utm.edu/epicurus/>
4. David Konstan (2014) SEP on *Epicurus*. <http://plato.stanford.edu/entries/epicurus/>
5. Richard Kraut (2014) SEP on *Aristotle’s Ethics*. <http://plato.stanford.edu/entries/aristotle-ethics/>

²⁸ I used to own a snowmobile (we did not plow the road to the cabin); it emitted more pollution than a fleet of cars. Zooming through the snow was bliss.

-
6. Martha Nussbaum (2012) Who is a happy warrior? Philosophy, happiness research, and public policy, *Intern. Rev. of Econ.* 59: 335-356.
 7. Christopher Shields (2014) SEP on *Aristotle*. <http://plato.stanford.edu/entries/aristotle/>
- iii Sources:
1. Thomas Aquinas (1265-74) *The Summa Theologica*, translated by Fathers of the English Dominican Province, Benziger Bros. Edition, 1947. At <http://www.ccel.org/a/aquinas/summa/FS.html>
 2. Holt (2006) *ibid.*
 3. P. Kreeft (1990) *Summa of the Summa*. Ignatius Press.
 4. Ralph McInerney and John O'Callaghan (2015) SEP on *Thomas Aquinas (Moral Doctrine)*. <http://plato.stanford.edu/entries/aquinas/>
 5. Darrin McMahan (2006) *Happiness: A history*, Atlantic Monthly Press. A great book.
- iv Sources:
1. William Bristow (2011) SEP on *Enlightenment*. <http://plato.stanford.edu/entries/enlightenment/>
 2. Holt (2006) *ibid.*
 3. Darrin McMahan (2006) *ibid.*
 4. Jean-Jacques Rousseau (1754) *A discourse on a subject proposed by the Academy of Dijon: What is the origin of inequality among men, and is it authorized by natural law?* (Translated by G.D.H. Cole, public domain, <http://www.constitution.org/jjr/ineq.htm>).
 5. Lars Svendsen (2005, in English) *A Philosophy of Boredom*, London, Reaktion books (Translated by John Irons. First published in Norwegian in 1999 by *Universitetsforlaget*, Oslo)
 6. Tzvetan Todorov (2010) Tzvetan Todorov and the Enlightenment Today, Philosophy Bites. <http://philosophybites.com/2010/01/tzvetan-todorov-on-the-enlightenment-today.html>
 7. Wikipedia on *Romanticism*, <http://en.wikipedia.org/wiki/Romanticism>
- v Sources:
1. Roger Crisp (2015) *ibid.*
 2. Dan Haybron (2011) SEP on *Happiness*. <http://plato.stanford.edu/entries/happiness/>
 3. Richard Reeves (2007) *John Stuart Mill: Victorian firebrand*, London, Atlantic Books.
- vi Sources:
1. John Burton, ed. (1746) *Life and correspondence of David Hume*, Volume 2. Edinburgh. <https://books.google.com/books?id=Gvg5AAAAcAAJ>
 2. Tim Crane (2007) *Mind and Body*, Philosophy Bites. <http://philosophybites.com/2007/09/tim-crane-on-mi.html>
 3. Robert Kirk on *Zombies*. <http://plato.stanford.edu/entries/zombies/>
 4. Howard Robinson (2012) SEP on *Dualism*. <http://plato.stanford.edu/entries/dualism/>
 5. Daniel Stoljar (2015) SEP on *Physicalism*. <http://plato.stanford.edu/entries/physicalism/>
- vii Sources:
1. Mark Bear, Barry Connors and Michael Paradiso (2007) *Neuroscience: exploring the brain: third edition*, Lippincott, Williams and Wilkins.
 2. Michael Gazziniga, Richard Ivry, and George Mangun (2014) *Cognitive neuroscience: the biology of the mind 4th ed.*, W.W. Norton.
 3. Paul Glimcher and Ernst Fehr editors (2013) *Neuroeconomics: 4th ed.: Decision making and the brain*, Boston, Elsevier: Academic Press. This is the standard neuroeconomics textbook.
 4. David Heeger: Lecture notes and slides on "The Brain." <http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/brain/brain.html>
 5. Suzana Herculano-Houzel (2016) *The human advantage: a new understanding of how our brain became remarkable*, Cambridge MA, MIT Press.
 6. Mathias Pessiglione, Ben Seymour, Guillaume Flandin, Raymond J. Dolan, and Chris Frith (2006) Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behavior in Humans, *Nature* 442(7106): 1042-1045.
 7. Douglas Quenqua (2014) Where personal memories take root, *N.Y. Times*, June 23.
 8. John Wixted, Larry Squire, Yoonhee Jang, Megan Papesh, Stephen Goldinger, Joel Kuhn, Kris Smith, David Treiman and Peter Steinmetz (June 2014) Sparse and distributed coding of episodic memory in neurons of the human hippocampus, *Proc. of the Nat. Acad. of Sci. (PNAS)*. www.pnas.org/cgi/doi/10.1073/pnas.1408365111
- viii Sources:
1. Mark Bear, Barry Connors, and Michael Paradiso (2007) *ibid.*

2. Andrew Caplin and Mark Dean (2008) Dopamine, Reward Prediction Error, and Economics, *Quart. J. of Econ.* 123(2): 663-701. It reviews many of the relevant findings.
3. Matteo Colombo (2014) Deep and beautiful: the reward prediction error hypothesis of dopamine, *Studies in Hist. and Phil. of Biological and Biomedical Sci.* 45: 57-67.
4. Nathaniel Daw and Philippe Tobler (2013) “Chapter 15: Value learning through reinforcement” in Paul Glimcher and Ernst Fehr editors (2013) *ibid.*
5. Christopher Fiorillo, Philippe Tobler and Wolfram Schultz (2003) Discrete coding of reward probability and uncertainty by dopamine neurons, *Sci.* 299: 1898-1902.

^{ix} Sources:

1. Mark Bear, Barry Connors, and Michael Paradiso (2007) *ibid.*
2. Daniel Benjamin, Ori Heffetz, Miles Kimball and Alex Rees-Jones (2012) **XXX** *Amer. Econ. Rev.* 102(5): 2083-2110. <http://dx.doi.org/10.1257/aer.102.5.2083>
3. Douglas Bernheim and Antonio Rangel (2004) Addiction and cue-triggered decision processes, *Amer. Econ. Rev.* 94(5):1558-90.
4. Kent Berridge (2005) Simple Pleasures. *Psych. sci. agenda* 18(11).
5. Kent Berridge and Wayne Aldridge (2008) Decision utility, the brain and the pursuit of hedonic goals, *Soc. Cognition* 26(5): 621-646.
6. Kent Berridge and Morton Kringelbach (2013) Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology* 23: 294-303. This article is an academic version of the *Scientific American* article, providing references to all the relevant studies.
7. Kent Berridge and Terry Robinson (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.*28(3): 309-369. <https://www.ncbi.nlm.nih.gov/pubmed/9858756>
8. Roger Crip (2015) *ibid.* (section *Desire Theories*)
9. Nathaniel Daw and Philippe Tobler (2013) *ibid.*
10. Iva Dincheva, Andrew Drysdale, Catherine Hartley, David Johnson, Deqiang Jing, Elizabeth King, Stephen Ra, et al. (2015) FAAH genetic variation enhances fronto-amygdala function in mouse and human. *Nature Commun.* 6. doi:10.1038/ncomms7395
11. Duke University (2012) With altered brain chemistry, fear is more easily overcome, *Sci. Daily* June 12. www.sciencedaily.com/releases/2012/06/120612101328.htm
12. Dietrich Franz and Christian List (2016) Reason-based choice and context-dependence: an explanatory framework. *Econ. and Phil.* 32 (2):175-229. ISSN 0266-2671
13. Richard Friedman (2015) The Feel-Good Gene, *N.Y. Times Sunday Rev.* March 6. <http://www.nytimes.com/2015/03/08/opinion/sunday/the-feel-good-gene.html>. Friedman is a professor of clinical psychiatry at Weill Cornell Medical College.
14. Michael Gazziniga, Richard Ivry, and George Mangun (2014) *ibid.*
15. Ozge Gunduz-Cinar, Matthew Hill, Bruce McEwen and Andrew Holmes (2013) Amygdala FAAH and anandamide: mediating protection and recovery from stress, *Trends in Pharmacological Sci.* 34(11): 637-643.
16. Amish Jha (2021) *Peak Mind*, Harper Collins.
17. Morton Kringelbach and Kent Berridge (2012) The joyful mind, *Sci. Amer.* 307(2): 40-45. Unless noted otherwise, quotes are from this article.
18. Marco Leyton, Kevin Casey, Scott Delaney, and Chawki Benkelfat (2005) Cocaine craving, euphoria, and self-administration: A preliminary study of the effect of catecholamine precursor depletion. *Behav. Neuroscience* 119, 1619–1627. https://www.researchgate.net/publication/7352753_Cocaine_craving_euphoria_and_self-administration_A_preliminary_study_of_the_effect_of_catecholamine_precursor_depletion
19. K.P. Macpherson et al. (2012) Convergent translational evidence of a role for anandamide in amygdala-mediated fear extinction, threat processing and stress-reactivity, *Molecular Psychiatry* 18: 813-823. [10.1038/mp.2012.72](https://doi.org/10.1038/mp.2012.72)
20. Masayuki Matsumoto and Okihide Hikosaka (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841. <http://www.nature.com/nature/journal/v459/n7248/abs/nature08028.html>
21. Derek Parfit (1984) *Reasons and Persons*, Oxford, Clarendon Press.
22. Marcel Proust (1913) *Swann’s way: In search of lost time, Volume 1* (translated by Lydia Davis), Penguin Classics.

-
24. Wolfram Schultz (1998) Predictive reward signal of dopamine neurons. *J. of Neurophysiology* 80(1), 1–27. <http://jn.physiology.org/content/80/1/1>
 25. Wolfram Schultz, Peter Dayan, and Read Montague (1997) A neural substrate of prediction and reward. *Sci.* 275, 1593–1599. <http://science.sciencemag.org/content/275/5306/1593>
 26. Cindy Wyvell and Kent Berridge (2000) Intra-accumbens amphetamine increases the salience of sucrose reward: enhancement of reward “wanting” without enhanced “liking” or response reinforcement, *J. of Neuroscience* 20(21): 8122-8130.

^x Sources:

1. Ralph Adolphs, [Frederic Gosselin](#), [Tony Buchanan](#), [Daniel Tranel](#), [Philippe Schyns](#), and Antonio Damasio (2005) A mechanism for impaired fear recognition after amygdala damage. *Nature* 433: 68-72.
2. Lisa Barrett (2006) Solving the emotion paradox: categorization and the experience of emotion. *Person. and Soc. Psych. Rev.* 10(1): 20-46. <http://www.affective-science.org/pubs/2006/Barrett2006paradox.pdf>
3. Mark Bear, Barry Connors and Michael Paradiso (2007) *ibid.*
4. Jan-Emmanuel De Neve, Nicholas Christakis, James Fowler and Bruno Frey (2011), Genes, economics and happiness. *J. of Neuroscience, Psych., and Econ.* 5(4): 193-211.
5. Michael Gazzaniga, Richard Ivry, and George Mangun (2014) *ibid.*
6. Gazzaniga, Michael; Ivry, Richard B.; Mangun, George R.. *Cognitive Neuroscience: The Biology of the Mind* (Fourth Edition) (Page R-32). W. W. Norton & Company. Kindle Edition.
7. [Stanley Schachter](#) and [Jerome Singer](#) (1962) Cognitive, social and physiological determinants of emotional state. *Psych. Rev.* 69, 379–399.

Chapter 5: Does consuming more increase WB? Is getting rich the path to happiness?

08162022

Money is better than poverty, if only for financial reasons (Without Feathers 1975)

Those who say that money can't buy happiness don't know where to shop (Anonymous)

Does consuming more goods and services increase your WB? Does expanding your set of paths available to you? Your available paths vary in the amounts of goods consumed; some, for example, have lots of leisure, but little is spent on goods and services. While in others, you work more, have more income, and buy more stuff. One's income limits one's expenditures, so newly available paths with more income make more consumption possible. But this does not imply you will choose to consume more. Warren Buffet could consume more goods and services than he does. Many others spend all they make, so they need to switch to a path with more income if they want to consume more.

This chapter begins by reviewing the data and research on the influence of income on WB. Complicating is that some surveys ask about happiness, others about emotional WB, or life-satisfaction WB, or the emotion the subject is currently experiencing. Some ask about something that isn't quite any of these things. So, keeping track of what is asked in each survey is crucial. Of course, this would be less of an issue if income has the same effect on emotional WB and life-satisfaction WB, but this does not seem to be the case.

My read of the research and evidence on WB (reviewed below) is that you will be better off if expanding your set of available paths allows you to go from a path that does not meet your basic needs to one that does: feeding the hungry increases their WB, so does giving those without shelter a roof, and so does giving those who face dangers a haven. However, once your basic needs are met, the evidence is mixed on whether more income will increase your long-run emotional WB—there will be a short-term happiness bump. And, what you do with more income is important.

But the research also indicates that, at a given time and place, the poor tend to be less happy than the rich (report lower levels of emotional WB), even if their basic needs are met. This seems, at first, a quandary: more does not make most of us happier, at least not for long, but, at a

point in time, more income relative to others does. Said another way, the evidence suggests that over time a group's average emotional WB does not increase substantially as its average income increases. But, among group members at a point in time, those with less are less happy. How can both be true, and why?

The evidence on the relationship between income and happiness: the big picture¹

The data, at first blush, can be confusing. In a 2008 article in the *Journal of Economic Literature*, [Andrew Clark](#), [Paul Frijters](#), and [Michael Shields](#) put the findings of many studies in four categories (they use the word “happiness” to summarize the findings):

1. Country-level data shows average happiness remaining flat over time despite large increases in average income per capita.

But, when one looks at data on the individual level, the results indicate that

2. Across individuals at a point in time, reported happiness is higher for individuals with higher income. Again, this effect is strongest in poor countries.
3. Panel data (data on income and reported happiness for the same individuals at different points in time) finds that an increase in an individual's income increases that individual's happiness, at least in the short run, but, again, more so in poor countries.
4. And, when one looks at large samples of aggregate data across countries and time, and one corrects for country fixed-effects (idiosyncrasies across countries), the higher a country's per-capita GDP, the happier it tends to be. Average happiness in a country is also positively related to how fast its per-capita GDP is growing relative to like countries.

I add two additional findings to their list:

5. The effect of income on life-satisfaction WB is likely different than its effect on emotional WB/happiness.
6. And both average emotional WB and average life-satisfaction WB are more sensitive to a decrease in national income than an increase.

Research on the relationship between income and WB is ongoing, so these six findings might not summarize what we know in five years. My goal is to provide a sense of the history of the research and to discuss the fundamental issues and questions concerning the data, including: (A) how to measure emotional WB/happiness and life-satisfaction WB and how to measure changes in their levels, (B) absolute income versus relative income, (C) that the data are diverse and not always comparable, and (D) WB is determined by many things besides income and consumption, making it challenging to identify their specific influence.

In Sections 3 and 4 of this chapter, I look at different theories of what determines WB and ask how well they explain the six categories of findings—but for now, consider only the data.

Data and studies on the relationship between income and WB

A bit of history: the Easterlin effect/Paradoxⁱⁱ

In 1974, **Richard Easterlin** pioneered the economic study of happiness, publishing, "Does economic growth improve the human lot? Some empirical evidence". Its notable finding, now known as the *Easterlin Paradox*, is that while U.S. aggregate income increased significantly between WWII and the early 1970s, reported happiness remained relatively flat—the Easterlin data is reported “happiness”. This trend has continued: while real per-capita income increased nearly 100% between 1970 and 2004, there was no corresponding increase in average happiness. This finding upset a few economists: it suggests that economic growth isn’t the path to increased WB. The Paradox is reflected in the following two figures:

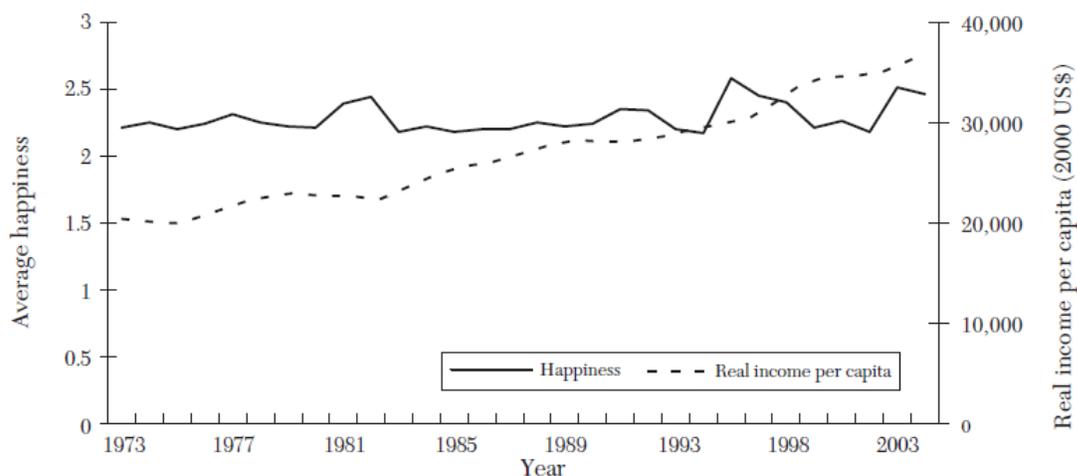


Figure 1. Happiness and Real Income Per Capita in the United States, 1973–2004

Source: World Database of Happiness and Penn World Tables. Happiness is the average reply to the following question: “Taken all together, how would you say things are these days? Would you say that you are...?” The responses are coded as (3) Very Happy, (2) Pretty Happy, and (1) Not too Happy. Happiness data are drawn from the General Social Survey.

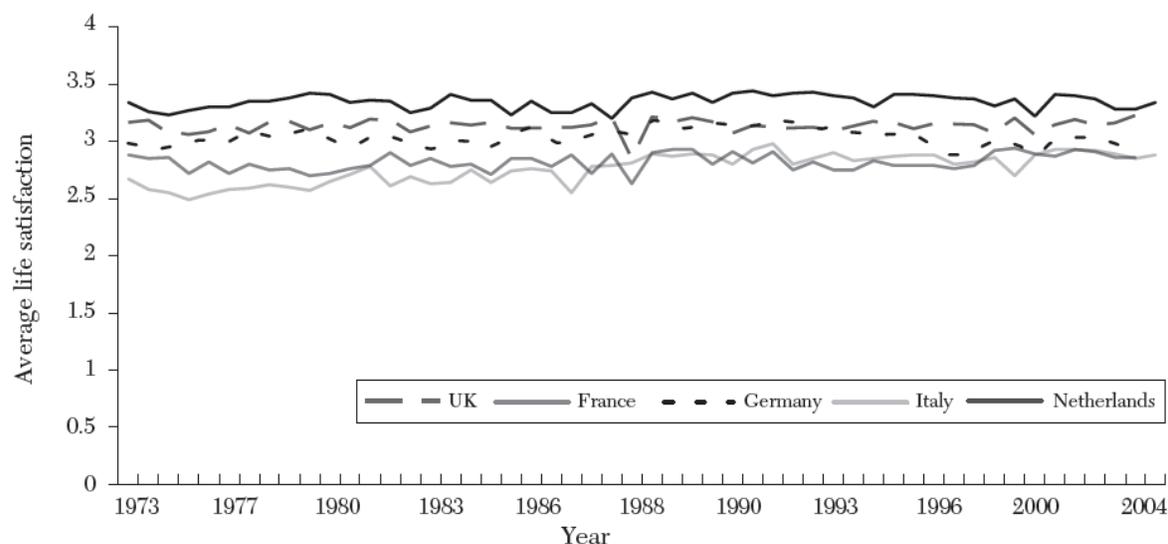


Figure 2. Life Satisfaction in Five European Countries, 1973–2004

Source: World Database of Happiness. Happiness is the average reply to the following question: “On the whole how satisfied are you with the life you lead?” The responses are coded as (4) Very Satisfied, (3) Fairly Satisfied, (2) Not Very Satisfied, and (1) Not at all Satisfied. Life satisfaction data are drawn from the Eurobarometer Survey.

Figure 15: Source is Clark, Frijters, and Shields (2008)

Note that their Figure 1 plots happiness while their Figure 2 plots satisfaction, suggesting that both average emotional WB and average life-satisfaction WB have remained flat as incomes have risen. Psychologists and sociologists get the same result, and the paradox holds for Japan.¹ In 1974, Easterlin concluded that average happiness does not vary much between rich and poor countries, and economic growth isn’t the path to greater happiness. He recently moderated this conclusion.

While Easterlin found no temporal relationship between per-capita income and average happiness, he found a positive relationship between one’s happiness and one’s relative income at a given time. This relative-position effect has since been seen by many. For example, based on [Eurobarometer](#) data from 1975-86 on 109 thousand Europeans, the economist [Andrew Oswald](#) reports that individuals in the lowest income quartile are less happy than those in the top quartile (“Very happy” 19% vs. 28% and “Not very happy” 27% vs. 13%).² More recent studies, using

¹ This growth-without-happiness paradox has been replicated in other countries and time periods by economists (e.g., Blanchflower and Oswald 2004), psychologists (e.g., Diener and Oishi), political scientists (e.g., Inglehart and Klingemann) and sociologists (e.g., Veenhoven).

² For an Oswald talk on the economics of happiness see <https://www.youtube.com/watch?v=Ggcezx4ZNLE>

different reference groups and different measures of relative position, continue to find that relativeness matters for emotional WB/happiness.³

[John Helliwell](#) at the University of British Columbia has analyzed data on life-satisfaction WB, finding the relative-income effect different in OECD countries versus developing countries. Specifically, in developing countries, the effect remains strong through all 10 deciles (each 10% income category from the lowest to the highest); for example, moving from the 80th income percentile to the 90th still significantly improves life-satisfaction WB. But in OECD countries, increasing relative income does not significantly improve life-satisfaction WB once one reaches the 40th percentile. This is surprising because moving from the 10th to the 9th decile in an OECD country requires a vast increase in absolute income.

Measuring WBⁱⁱⁱ

Most measures of WB are self-reported (“What is your level of happiness on a scale of 1 to 5”). Happiness is subjective.⁴ In 1974 (and often still), many economists dismissed Easterlin’s findings because the data is not observed behavior but self-reported feelings—many economists dismiss stated feelings and views. This begs the question of what observed behaviors could one use to determine WB—facial expression, body language? While neuroscientists use facial expressions to assess liking, can you imagine an economist measuring WB this way? Charles Darwin suggested both as windows into the emotions of animals. After winning an Oscar, [Anne Hathaway](#)’s expression was described in *The New Yorker* as joy as only a child can express. Still, her expression pales compared to Alexander Gamme’s [reaction](#) when he discovered a candy bar he had stowed on the way out on the 86th day of his solo trek to and from the South Pole. But, emotional WB is not always so expressive. At the other extreme, one can observe behaviors indicative of extreme unhappiness, such as crying, frowning, shaking, lethargy, and suicide attempts, but unhappiness, like happiness, is often unexpressed. [Blanchflower](#) and Oswald (2008) find, using country-level data, a negative relationship between average happiness and the

³ See Dolan, Peasgood, and White; Dorn et al.; Ferrer-i-Carbonell; and Luttmer.

⁴“Subjective” as in determined within the “subject” (the answering individual) so determined, at least in part, by what is going on within the subject’s brain. In contrast, objective is “it is what it is.” Taste and color are subjective, size in contrast is objective (as in Box B is bigger than Box C). See Sumner.

proportion of the population with hypertension.⁵ Hypertension and the hormone cortisol correlate with stress, making happiness less likely.

A researcher could use the assessments of friends and family (which happen to be significantly correlated with self-reported happiness), but many economists would reject these as unobjective.

Brain activity is objective in that it can be externally observed, at least in theory. And a specific pattern could be deemed a measure of emotional WB if most subjects report they are happy when this pattern is observed. Then, we would not have to ask, only monitor their brain patterns in real-time. In the future?⁶

Instead of directly asking about happiness, a few researchers have asked about the frequency and intensity of sensations and emotions experienced in the last hour or day, experiences such as joy, laughter, smiling, sadness, stress, and anger. The researchers then convert the answers into a level of happiness (emotional WB).⁷ How to do this is only now becoming a research question, research motivated more by policymakers than economists: policymakers are asking for direct estimates of WB.⁸

Given that NBT assumes that the individual chooses their highest-ranked available path, economists are in a circular quandary when assessing WB. For example, when I buy a giant TV, it must increase my WB. Otherwise, I would not have bought it; if I say, "Buying it did not increase my WB," I must be wrong.

⁵ Countries with high hypertension happen to be countries where many people are overweight. This raises the question of how happiness and being overweight are related. I know of no studies that look at this relationship, but many of us imagine we would be happier if we lost a few.

⁶ See the discussion of hedonic hotspots in Chapter 4.

⁷ Taking a middle approach to measuring emotional WB, the [Health and Retirement Study](#) at the U. of Michigan asks, "Please tell me if each of the following was true for you much of the time this past week: a) You felt you were happy b) You felt sad c) You enjoyed life d) You felt depressed."

⁸ In a 2014 paper, [Daniel Benjamin](#) and other economists, propose an index for calculating your WB. It is a weighted average of your answers to questions that ask you to tradeoff different aspects of WB; for example, would you prefer to be less anxious or somewhat happier? Such questions are different from the standard, "How happy are you?" Or "How happy would this make you?" The weights in the index are estimated using the answers to many such questions by many people. The estimated index is therefore a WB index for an average individual, not for a specific individual. The authors conservatively pitch their model as an initial step in the long process of building a WB index.

Admittedly there are serious issues with comparing self-reported WB across individuals and even for the same individual at different points in time. Our short-term memories of sensations and emotions can be faulty. Culture, language, and expectations differ across individuals, so they complicate comparisons. For example, Italians report lower happiness than Germans (see Fig. 5), but this might only be due to a cultural difference in how feelings are expressed. How happy you report to be might be a function of how happy you think you have the right to be or whether happiness is viewed as societally desirable.⁹ Statistical models estimated with panel datasets are starting to be used to control for such idiosyncrasies (see categories 3 and 4 above), clarifying the picture.¹⁰ But first, let's look only at the time-series data.

Time-series data and the difficulties in determining the influence of income^{iv}

To identify the influence of a change in aggregate income on either average emotional WB or life-satisfaction WB, one must control for all the changes that might affect WB. These include age composition, health, social institutions, and gender and racial equality. This is difficult. For example, studies suggest that an individual's happiness is U-shaped by age, decreasing in one's twenties but increasing for old farts—like me. So, since many populations are aging, one must also determine how the changing age composition affects happiness to assess the influence of aggregate income on happiness.

Occasionally situations arise where an exogenous shock causes a large change in a country or region's GDP. For example, in east Germany, during the decade after reunification, both life-satisfaction WB and real income jumped, raising the question of whether the income increase caused the happiness increase. Part, but not all, of the jump can be attributed to more individuals fulfilling their basic needs. Explaining the rest of the increase in life satisfaction requires separating the influences of increased income from increased freedoms and expectations. One must also factor in that incomes in east Germany rose relative to their neighbors: average income in Poland and west Germany did not rise.

At a given point in time, the unemployed are, on average, not as happy as the rest of us.^v

But this might not be because they have less to spend.

⁹ For more on the measurement of happiness see Diener 2000.

¹⁰ Data on income, self-reported happiness, and other explanatory variables, for the same individuals at multiple points in time.

Nineteen percent of all the individuals in the Eurobarometer data reported being “Not very happy”; but 33% of the unemployed are in this category. The unemployed are anxious and distressed. Interestingly, the highly-educated unemployed are more unhappy than the lowly-educated unemployed—a disconnect between expectations and outcome? The poorly educated are less surprised by unemployment, so they experience less distress when it occurs? Oswald suggests that most unemployment distress isn’t caused by the loss of income but by a loss of self-esteem, a sense of failure, and embarrassment; being unemployed reduces one’s status. Helliwell finds that the negative effect of unemployment on life-satisfaction WB is on par with a one-unit drop in one’s health (health measured on a five-point scale).

One study in detail: the happiness bump^{vi}

Seven-thousand eight-hundred and seventy-one Germans were followed from 1985 to 2000, with data collected yearly on their income and happiness. The sample consists of randomly selected individuals living in the western part of Germany for whom there were at least five years of data. This is panel data, the beauty of which is that it is the same people, year after year, rather than sampling different people every year. So one can rule out the possibility that answers are changing because this year’s respondents differ from last year’s.¹¹ The data has been analyzed by [Rafael Di Tella](#), [John Haisken-DeNew](#), and [Robert MacCulloch](#); their Fig. 3 supports the Easterlin Paradox. Average national income is in Deutch Marks.

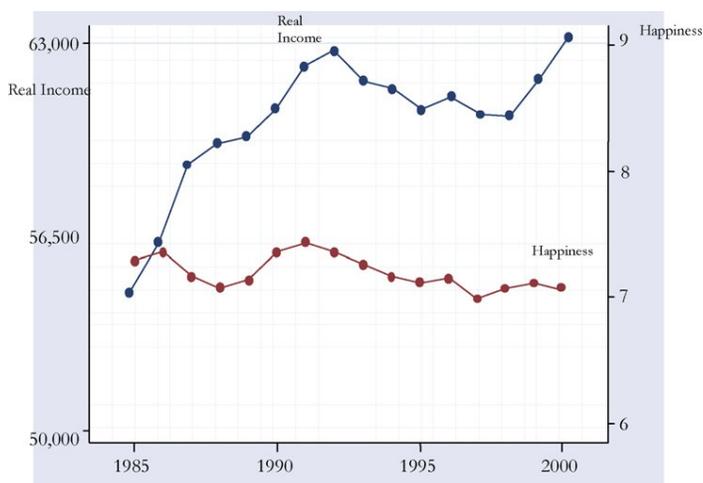


Figure 16, Chapter 5: Fig. 3 (Di Tella et al.) The graph plots average real income and happiness (on a 0-10 scale) for a group of 7812 individuals who are followed from 1985 to 2000

¹¹ It does not control for age; the sample is necessarily aging.

“Happiness” in the graph is the average happiness for the sample in that year. Nothing in this graph suggests an increase in average income will cause an increase in average happiness.

The authors estimate a statistical model(s) of everyone’s reported happiness as a function of their current income and job status; job status and income in prior years; socioeconomic characteristics including marital, employment, and educational level in each year; plus, gender and political orientation. The model includes individual-specific constants (individual “fixed effects”) and year-specific constants, so it controls for the fact that some people are just naturally happier and that things change from year to year in ways the researcher can’t always observe. The results are summarized with two graphs. A permanent 50% in an individual’s income is estimated to cause a temporary bump in the individual’s happiness, but after three years, much of the happiness gained has been lost.

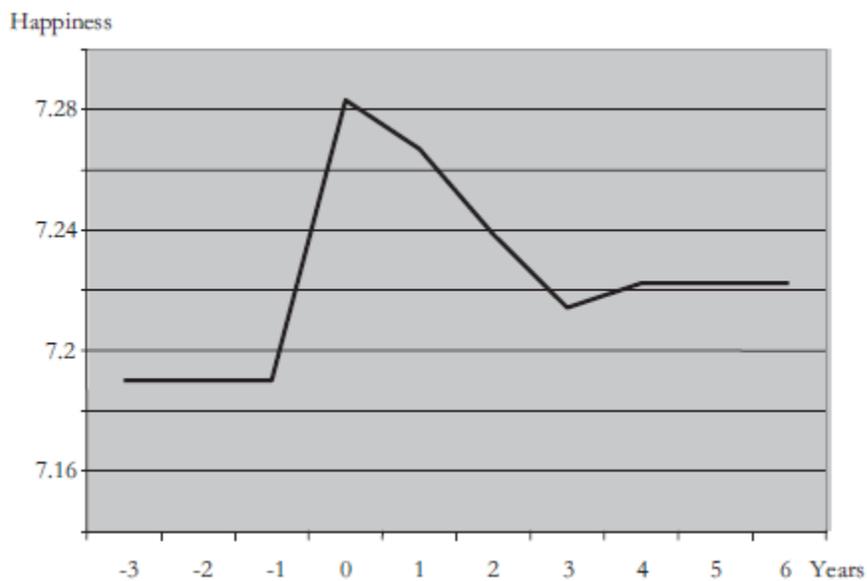


Figure 17, Chapter 5: Fig.1 (Di Tella et al.) Happiness after a shock at $t=0$ that causes income to rise by 50%

In contrast, for most individuals, most of the happiness increase that results from a permanent increase in status (an increase in relative position) persists for at least six years.

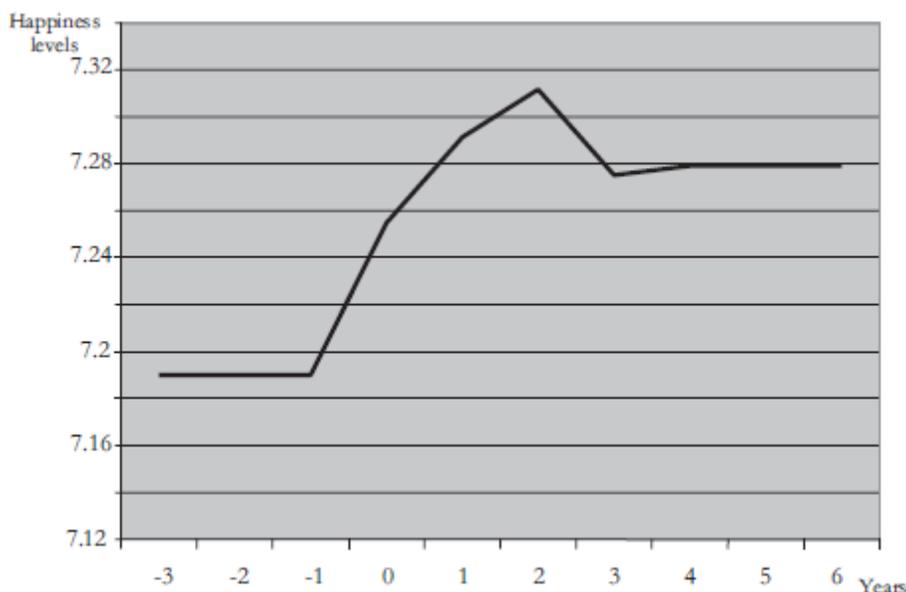


Figure 18, Chapter 5: Fig. 2 (Di Tella et al.) Happiness after a shock at $t=0$ that causes status to rise by 50%

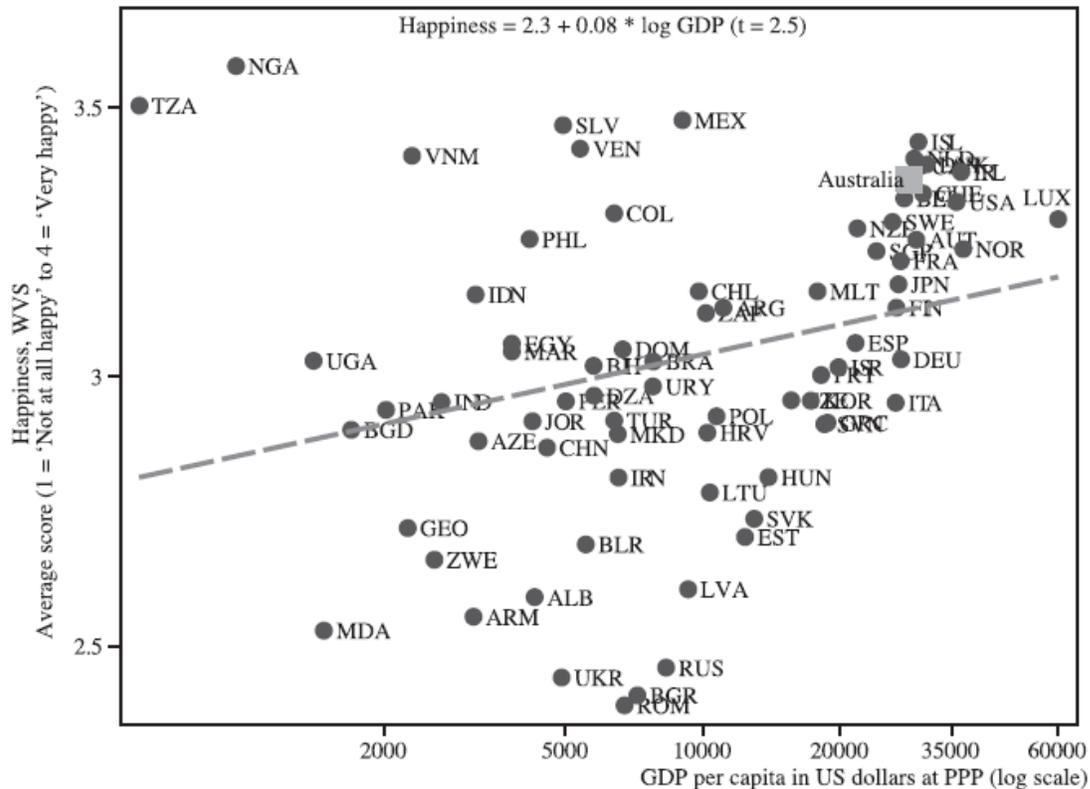
Summarizing their findings, most people get used to (adapt to) having more income and consumption, but the happiness increase from an increase in their status persists—there is little adaptation to increased status.¹²

When one looks at large data sets across countries and corrects for idiosyncrasies across those countries, average happiness increases with per-capita income, but only slightly

This is the fourth category of finding. The higher the per-capita GDP in a country relative to other countries, the higher its average happiness tends to be—people in more affluent countries tend, on average, to be somewhat happier (see Fig. 18, which indicates there are a lot of other determinants). And, if two countries have similar per-capita incomes, the one where per-capita is growing faster tends to be happier—high growth rates relative to others increase happiness.

¹² There are fascinating exceptions; those on the political right grow accustomed to an increase in their status, but the effect on their happiness of an increase in absolute income persists. Why?

Figure 5 Happiness and GDP Per Capita, World Values Survey



© 2006 The University of Melbourne, Melbourne Institute of Applied Economic and Social Research

Figure 19, Chapter 5: source Leigh and Wolfers (2006) Happiness and the human development index: Australia isn't a paradox, *The Australian Economic Review* 39(2)

These category four findings need to be squared with the Easterlin paradox—that increased aggregate income does not significantly increase a country's average happiness. Keep in mind: (1) Richer countries tend to be more democratic, tend to have more stable institutions, and individuals tend to have more rights and freedoms. All of these could increase both emotional and life-satisfaction WB. And (2), the fourth category of findings can be explained by relative position.

The fifth finding: life-satisfaction WB increases (and decreases) with income:^{vii}

Kahneman and Deaton (2010) investigated a Gallup panel of Americans; each individual was repeatedly asked to separate their emotional WB from their life-satisfaction WB. Increases in yearly income increased emotional and life-satisfaction WB, but only for incomes less than \$75K (2010 dollars). Beyond that, additional income grows life satisfaction but not emotional WB. Kahneman and Deaton define emotional WB as the “frequency and intensity of experiences of

joy, stress, sadness, anger, and affection that makes one's life pleasant or unpleasant." Whereas they define life-satisfaction WB as "the thoughts people have about their life when they think about it."¹³

And finally, the sixth finding: both emotional WB and life-satisfaction WB are more sensitive to a decrease in national income than to an increase:^{viii}

Using the same data as Kahneman and Deaton, [Jan-Emmanuel De Neve](#) and his coauthors find that an individual's emotional and life-satisfaction WB are more sensitive to a decrease in national income than an increase: a decrease of X will lower WB more than an increase of X will increase it. This can explain why, in the long run, average happiness did not increase even though real per-capita income rose substantially. Over the business cycle, the average individual's WB drops more when aggregate income declines than it rises when income grows. In the authors' words, "...recessions can rapidly undo the WB gains from longer expansionary periods and lead to an insignificant relationship between national income and average WB when considered in the long run.

Complementing their primary finding, they find that self-reported unhappiness (stress, worry, reduced joy) increases when aggregate income declines, but it does not wholly abate when income rebounds. My discussion of the endowment effect in Chapter 7 details a 2013 study by [Christopher Boyce](#) and coauthors that supports these conclusions.

A hypothesis: your WB is determined more by your relative position than by your absolute level of consumption

"Relative position" can be your position relative to others or your current position relative to your former self. For individuals whose basic needs are met, relative position can explain both (1) why a permanent income increase does not cause a permanent increase in WB, emotional or life satisfaction; and (2) why an increase in relative income (status) can cause a long-term increase in both emotional and life-satisfaction WB. And (3) relative position effects are consistent with, and could explain why a loss of income decreases emotional WB more than an equal income increases it.¹⁴

¹³ They use "WB" to refer to what I am calling life-satisfaction WB.

¹⁴ Looking ahead, Chapter 7 discusses the *endowment* effect: a loss of income decreases WB more than an income increase of the same magnitude increases WB. The main explanation for the endowment effect is loss aversion—we are averse to losses. The endowment effect explains, in part, why an increase in average National income over the

The relative-income (status) hypothesis^{ix}

We all have strength enough to endure the misfortunes of others (Francois de La Rochefoucauld, French essayist 1613-1680) I might add, *and the rich have the most endurance.*

What everyone most aims at in ordinary contact with his fellows is to prove them inferior to himself (Arthur Schopenhauer, German philosopher 1788-1860)

You would be happier if you continue to consume the same path of goods and services, but your neighbors and peers come to consume less, making you relatively better off. And, you would be less happy if your consumption remains constant and they increase their consumption: you are relatively worse off.¹⁵

The relative-income hypothesis is consistent with the finding that, at a given point in time, people at the bottom of the scale are less happy than those at the top—those at the bottom are, by definition, “poor”. Data indicate that the poor perceive the world as largely beyond their control, a world controlled by those higher up the socioeconomic ladder. This perceived lack of control—a correct perception—causes stress. We all experience stress, but those higher up the ladder have more control over their stressors and more resources to deal with them. One result is that the poor don’t live as long, suffering more from diabetes and hypertension. Their quality of life is further depressed because the rest of us have little empathy. Research by Dacher Keltner, Michael Kraus, and others shows that we pay less attention to those lower down the scale and more to those higher up. Those at the bottom of the scale are largely invisible to those at the top—an interpersonal empathy-gap. When individuals of different status meet, the higher-status

long-run does not necessarily cause an increase in average emotional WB: economic contractions decrease emotional WB more than expansions increase emotional WB. To re-acquire the level of emotional WB you had before the great recession, the subsequent expansion will need to raise your income substantially above what it was pre-recession.

¹⁵ One could hypothesize the opposite. one might be happier if there are many people consuming paths with more consumption than yours because their existence makes you believe you will eventually be consuming more: observing others with more fuels the American Dream? This alternative could explain why many poor Americans do not want the rich taxed more because they hope to be rich one day.

Recent research indicates Americans seriously overestimate the potential for upward mobility. For the rich, the overestimation justifies their wealth, and for the poor, it gives them more hope than there is. Subjects in one study were asked to estimate the probability that a randomly selected individual currently in the bottom 20% income category (the poorest quintile (1/5)) would move up. For moving from the bottom quintile into the top three (top 60% of incomes) the average estimate was 45% would move up, but only 30% do. Conservatives overestimate more than liberals. In another study, respondents overestimated the probability by five-fold that a college student came from the bottom quintile. Michael Kraus, Shai Davidai, and David Nussbaum conclude that “Taken together, these sets of results suggest that belief in the American dream is woefully misguided when compared to objective reality.” (NYT 2015)

individual signals disregard for the other person by talking at them, not listening to them, and looking past them. A 2013 *N.Y. Times* article sums it up, “[Rich people just care less](#)”. When we do pay attention, we higher up the scale tend to believe those lower are lower for a good reason, that “social class has an underlying, inherent or natural foundation” (Kraus and Keltner 2013). We feel our higher rank is justified. People at the bottom, in contrast, are more likely to think that relative position is the luck of the draw.¹⁶

The influence of relative position and status harps back to Rousseau. He blamed property rights, markets, commerce, and the Enlightenment—beyond what you would find in a village—for creating an unnatural desire to want more than others. This results in unhappiness. The institutions that allow us to consume more lead to misery because we acquire a desire (an unnatural desire) for a higher relative position.¹⁷

Rousseau was emotional and passionate¹⁸—a bit kinky—and felt emotions should play a larger role. He distinguished between desires and needs, finding desires evil. Echoing Rousseau, the American economist, sociologist, and critic of capitalism, [Thorsten Veblen](#) (1857-1929), in his famous 1899 book *The Theory of the Leisure Class*, harangued relative consumption, calling

¹⁶ A 2010 paper ([Piff et al.](#)) covering four studies found lower-class individuals to be more generous, charitable, trusting, and helpful because “of a greater commitment to egalitarian values and feelings of compassion.

¹⁷ Rousseau did not believe the desire genetic; for him, it started when a man first enclosed a field and said “Ceci est a moi.” His is an interesting conjecture: property rights, the institution that has made it possible for us to consume great amounts of goods and services, make us unsatisfied; property rights bring inequality because with them comes the potential to concentrate them. Paraphrasing the philosopher [Melissa Lane](#), the Enlightenment propelled the idea that we all could have a higher standard of living (progress) if we commercialized society (more, buying, selling, profits and property rights, more competition). Hume, Smith, and Rousseau all believed commercialization would increase consumption. But Rousseau argued it would not bring happiness because it creates a system based on desires (Assumption 9b), including the desire to have more than others. Paraphrasing Rousseau, commercialization changes man from a virtuous independent (a noble farmer in a small village) to a pursuer of desires who competes with others for self-esteem in the eyes of others. Commercialization is driven by pride, vanity, and esteem-seeking; it creates a world where you can never be satisfied, an unnatural world. Rousseau argues that the market system makes us unhappy because it causes inequality and neighbor-lusting. Long before Marx and Hegel, Rousseau was criticizing the market system. Economists tend to dismiss Rousseau's economic views, but it is the case that once basic needs are met, increased consumption does not always increase emotional WB.

[An aside: Rousseau was an early environmentalist arguing that commercialization would cause significant environmental degradation (large, polluted cities and the destruction of nature). He harped for a return to nature, not a return to “acorn grubbing”: a return to a simple agricultural society organized around the family and the village.]

¹⁸ Hume, referred to Rousseau as the “wild philosopher” noting Rousseau “is plainly mad after long having been maddish.” Rousseau thought a lot of people, including Hume, were out to get him (all were not paranoid delusions). He seemed to have visions, and he did drop all his five children, one by one, off at the foundling asylum. In the 1890 *American J. of Psychology* there is a article titled “The insanity of Jean-Jacque Rousseau”: it is an early example of psychological profiling, stating “that Rousseau was insane is generally, but not universally, admitted.” The article presents the evidence.

it "pecuniary emulation". Veblen is responsible for the term "conspicuous consumption"—showing off with goods—arguing that demonstrating superiority with goods replaces showing off with excess leisure (a leisure class) as an economy matures.

Who do we compare with? And do we compare only in terms of income?

Why do we compare? Social psychologists have *social-comparison theory*: we use other people to gather information about ourselves (Festinger 1954). Comparison is part of our quest to make ourselves feel better.¹⁹ Social comparison theory and the data indicate that you compare with those like you, those you know: neighbors, relatives, friends, co-workers, your ethnic group, and your age cohort. Data from the [National Survey of Families and Households](#) finds that happiness is negatively correlated with your neighbors' income, and the effect is stronger the more you socialize with them.

As noted, a jump in personal income temporarily increases happiness (emotional WB). While part of the happiness bump is attributable to consuming more new stuff, part is because we have improved our position relative to others, including our past self. But, like the effect of consuming new stuff, this self-effect must be temporary because our new self will become our old self.

Residents of a region or country also compare their per-capita GDP and its growth rate with these measures in other places. This affects happiness; another country can be part of one's reference group. We get a happiness bump from living in a more affluent and faster-growing country, and this bump persists if our country's relative advantage persists. Although residents of the U.S., Europe, and Canada are likely affected by each other's per-capita income, it is unlikely that their WB is affected by the per-capital income of Liberia.

While our happiness is influenced by our relative income, it is also influenced by our relative position on other scales

But this has not gotten as much attention. Tim Wadsworth, a sociologist at the University of Colorado, finds that happiness increases in how much sex you have relative to the average for

¹⁹ Quoting Morey and Kritzberg, "Motives for the drive include self enhancement, perceptions of relative standing, maintaining a positive self-image, and closure...Studies indicate that the inclination to socially compare is positively correlated with (1) low self-esteem and neuroticism, (2) a strong interest in others and what they feel, and (3) having a 'high chronic activation of the self'..."

your age/gender cohort.²⁰ Besides sex and money, people compare based on specific commodities (e.g., cars, houses), athletic accomplishments, the accomplishments of one's children, and the attractiveness of one's partner.²¹

Loss of White privilege

You see a common pattern in the Capital insurrections. They are mainly middle-class to upper-middle-class whites who are worried that, as social changes occur around them, they will see a decline in their status [emphasis added] in the future." ([Robert Pape](#) as quoted in the NYT 06/06/2021).

In the aftermath of the Capital attack, the polling firm Echelon Insights decided to ask voters a simple question: Do they think the goal of politics is more about "enacting good policy" or "ensuring the country's survival as we know it?" ... nearly half [of Republicans] said it's about survival. (Brooks 2021)

A February [2021] [Economist-YouGov poll](#) asked Americans which statement is closest to their view: "It's a big, beautiful world, mostly full of good people, and we must find a way to embrace each other and not allow ourselves to become isolated" or "Our lives are threatened by terrorists, criminals and illegal immigrants, and our priority should be to protect ourselves." ... Over 75% of Biden voters chose "a big beautiful world." Two-thirds of Trump voters chose "our lives are threatened."

The University of *Chicago Project on Security and Threats* (CPOST) analyzed the demographics of the home counties of the 337 (to date) Americans arrested or charged in the January 6th Capital siege. Ninety-five percent are white, and 85% are male. Only 10% have ties to right-wing groups, and approximately half are from Blue counties. Their significant finding: "Counties with the most significant declines in the non-Hispanic white population are the most likely to produce insurrectionists who now face charges." (*Washington Post* 04/06/2021).²² That is, from counties where the proportion of non-whites is growing the fastest. CPOST (February and March 2021) found that a primary motivator is the fear of the *Great Replacement*: the view that minorities and immigrants want to take over the country (*NYT*).²³

²⁰ Of course, a positive correlation between sex and happiness does not imply that more sex results in more happiness: maybe happy people have more sex.

²¹ I spend a great deal of time training, so I can bike faster and longer than my biker friends, or at least not be too far behind. My comparison group used to be all Boulder bikers; then I reduced it to females and older males; now it is only senior citizens.

²² "... the [U.S.] counties that saw the greatest decline in White population had an 18% chance of sending an insurrectionist to D.C.; while the counties that saw the least decline had only a 3% chance."

²³ Some in white America are experiencing an "actual or potential loss of dominance, a sense of resentment at this loss which is bound up with issues of entitlement—the undeserving are taking what we deserve" ([Reicher](#) and [Ulusahin](#) 2020). These authors call it "dominant group victimhood".

If we compare multiple dimensions, a critical question is, at what rate will you substitute higher status on one scale for a reduction on another?

For example, I can spend more time publishing research papers to increase my professional status, but that would leave less time for training to maintain my position in my cycling gruppo. What's my rate of exchange? Many parents sacrifice the McMansion to send their children to a prestigious college—a few have risked jail by bribing a USC rowing coach. Capital insurrectionists risked jail to support white privilege. [Many even advertised their crimes on social media.] Alternatively, could those who suffer racial discrimination be made whole with money (e.g., payments to Black Americans)?

Differences in happiness, by income, in the same society at a point in time could, in part, be the result of expectations and aspirations:^x

Comparing to your past and future self:

You are unhappy because you feel underpaid relative to your peers and ability. You are disappointed because it does not meet your expectations, expectations driven by your assessment of your ability, relative ability, and sense of what is fair.

Aspirations (what you aspire to and want to achieve) are influenced by what others have. For example, a peer with a higher income can make you getting a raise more plausible. Research by the Swiss economist [Alois Stutzer](#) and others indicates that high-income aspirations and expectations reduce happiness. Aspirations are often thwarted, and more so for the poor. A few are poor because they lack aspirations (the happy poor?), but most of us, including the poor, aspire. Thwarted expectations cause frustration; frustration leads to discouragement—not a pretty picture. In animals, stress that can't be controlled leads to *learned helplessness*.²⁴ The picture is quite unpleasant if you expect that your aspirations will come true, and they don't. Danes appear happier than other Europeans. It is suggested that this is because they largely have income equality, making it unrealistic for most in Denmark to aspire to a higher relative income.

²⁴ An animal repeatedly subjected to an unpleasant stimulus they can't control learns to stop trying to avoid the stimulus, even when avoidance becomes possible.

If emotional or life-satisfaction WB is determined by relative status, what would increase aggregate WB?^{xi}

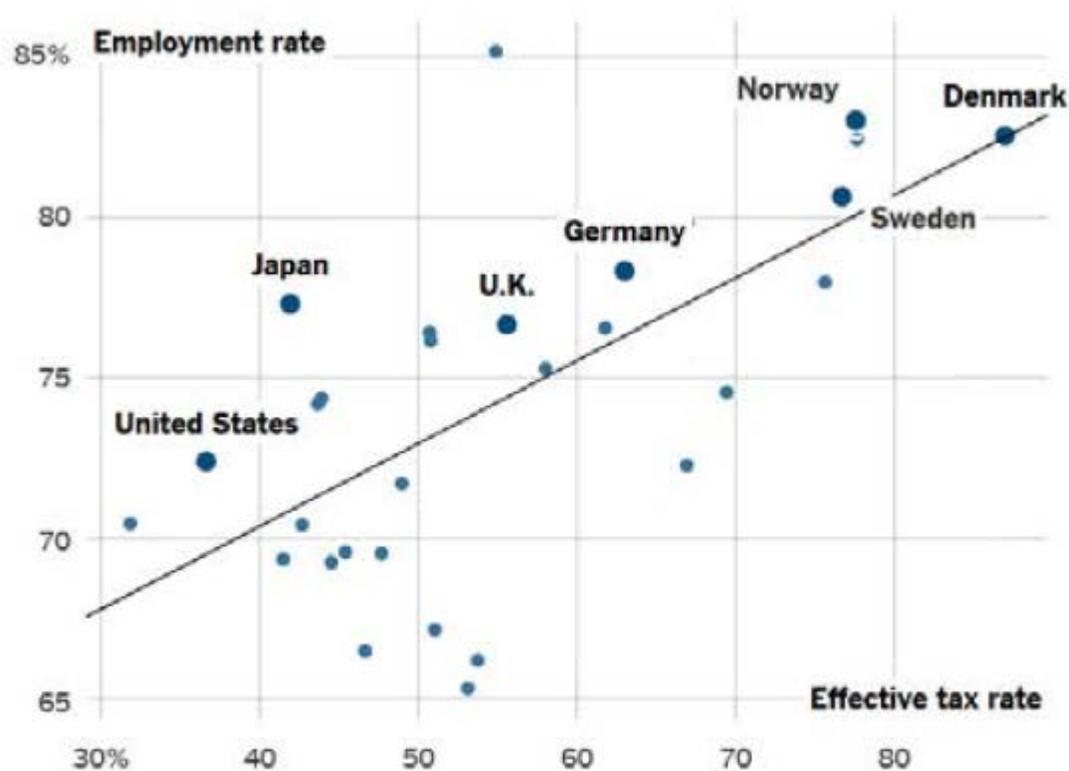
This section is about what sorts of policies would increase aggregate WB if relative-status effects exist. It isn't about whether invoking such policies is the right thing to do. And while people care about their own WB, it isn't clear they care much about increasing the WB of others.

If relative position in society is more important than absolute income in determining WB, increasing aggregate income will not necessarily increase aggregate WB. Depending on how the increase is distributed, it might even decrease. When my income increases but not yours, it reduces your WB: your relative income has declined. So, if unconstrained in my pursuit of a higher income, I will strive too much, from your perspective, to get ahead. I inefficiently work too much from society's perspective—so do you; Robert Frank discussed this in 1997.

If your earning more increases your WB but reduce the WB of others, what WB increasing policies might be pursued? A standard economic solution would be to raise marginal income-tax rates at the high end (make the tax more progressive) to encourage you to work less and earn less. [Like taxing cigarettes to discourage too much smoking: using the tax to increase the private cost of smoking (or working) so it equals the societal cost.]

At the high end, after-tax income will decline because taxes have increased and possibly also decline because those at the high end might work less: they would keep less of every additional dollar, so they might have less incentive to work. But there is little evidence that more progressive tax rates would cause the rich in the U.S. to work less or earn less. Given the current marginal tax rates in the U.S., increasing progressiveness would not decrease the amount we work. Moreover, tax rates in the U.S. are historically low and low compared to other countries. Detailed analysis by the French economist [Thomas Piketty](#) (author of [Capital in the Twenty-First Century](#)) and the Berkeley economist [Emmanuel Saez](#) indicates that much higher marginal tax rates in the U.S. would not retard economic growth.

Consistent with their conclusion, the Scandinavian countries have high tax rates and high rates of labor-force participation.



Employment rate among people aged 20 to 59; effective tax rate is more precisely known as the "participation tax rate," capturing both tax burden and loss of welfare benefits for a worker earning average wages.

Source: O.E.C.D. data analyzed by Henrik Jacobsen Kleven of the London School of Economics

Figure 20, Chapter 5: Source is Neil Irwin, [A big safety net and strong job market can coexist](#), *N.Y. Times* Dec. 17, 2014

The U.S. has one of the lowest effective tax rates and one of the lowest labor-force participation rates. [As an aside, Denmark, with one of the world's highest labor-force participation rates, ranks near the top on measures of happiness and life satisfaction. This suggests work might even increase happiness if the stress of working and raising a family is reduced by a safety net for childcare and health services.]

So, what would happen to relative income positions in the U.S. if taxes were made more progressive without affecting total tax revenues?²⁵ The rich would pay more and the poor less,

²⁵ More progressive taxes do not mean that tax revenues will necessarily increase; they might even decrease. If tax revenues are affected, what happens to emotional WB will depend on what is cut from or added to the budget.

but no one's income rank (first, second, . . . , last) would change. The income distribution would only squish—all of us would move closer, income-wise, to both our richer and poorer peers.

Whether this squishing of the income distribution would increase or decrease WB depends on how it affects perceived status and expectations: the rich would not be as rich as they were compared to you, but the poor would not be as poor. How income disparity affects average WB is critical but difficult to nail down. For example, Scandinavian countries rank high in happiness rankings, but why? Is it because of their more progressive tax rates, their more robust safety nets, or because their populations are ethically and racially more homogenous?²⁶ Or maybe it is all three.

Another critical but unanswered question is how average WB is affected by wealth distribution compared to income distribution. Ways to squish the wealth distribution (as compared to the income distribution) include wealth and inheritance taxes.

Many developed countries restrict hours of work (35-hour workweeks, 6 weeks of mandatory vacation). Increasing the cost of working more makes sense on efficiency grounds if people gain status by simply working more than their peers, creating a situation where we compete by who works more—picture new associates at a high-powered law firm.

Relative consumption is also influenced by public commodities (those consumed by all²⁷) and by government services such as education and healthcare. For example, resources allocated to building McMansion benefit primarily its inhabitants in absolute and relative terms, but resources allocated to increasing national security make us all more secure. Government-provided healthcare increases the real income of the poor more than the rich. In a market economy, allocating more resources to produce private goods will typically lead to more consumption disparity than allocating those resources to producing public goods or government-provided services.

To be clear, I am not claiming that increased income equality and spending a larger proportion on public goods will, for sure, increase WB. I am pointing out that relative-

²⁶ When I was a professor in Norway, I was in a 90% marginal tax bracket.

²⁷ A public commodity is a commodity that once produced is consumed/experienced by all, whether they like it or not. A standard example is national defense. Fighting the war in Iraq was a public commodity, a public good for some, a bad for others. Polar bears saved from extinction is a public commodity: if they are saved for you, they are saved for me, whether I like it or not.

income/status effects should not be ignored if the intent is to increase WB. Income inequality is rising, particularly in the U.S: 60% of the increase in US national income in the 30 years after 1977 went to just the top 1% of earners.

In summary, relative effects are critical and complicated. But most modern economists ignore them, and all economists ignore relative effects on dimensions other than income/consumption. NBT can accommodate relative effects if my ordering of paths depends on what I consume and what those in my reference group are consuming.

Other data and theories suggest income and wealth are not critical determinants of happiness nor WB

This section relates four different perspectives on the causes of WB, starting with genetics and what happened when you were a baby:

One's happiness has a set point from which it is, long run, difficult to deviate?^{xii}

Adam Smith, a former Chair of Moral Philosophy at Glasgow University, thought so:

The mind of every man, in a longer or shorter time, returns to its natural and usual state of tranquility. In prosperity, after a certain time, it falls back to that state; in adversity, after a certain time, it rises up to it. —The Theory of Moral Sentiments (1759/1853, p.149)

If we each have our own fixed baseline happiness, our *fixed point*, and our basic needs are met, market outcomes, government policies, and making more money can't make you happier, at least not for long. Nothing can—they also can't permanently decrease your WB.²⁸ By this view, even the death of a loved one only affects WB for a while. See Figure 1. Emotional WB is on the vertical axis. It takes, on average, six months for an adult to rebound from the loss of their mother but five years if they lose a child.

²⁸ Forty years ago, [Paul Brickman](#), a young and rising psychologist, was the first to investigate the fixed-point hypothesis. He interviewed lottery winners (\$50K to \$1M) and individuals paralyzed in accidents (Brickman, Coates, and [Janoff-Bulman](#) 1978). The interviews were within a year of the event. The victims were less happy now than they were before their accidents, but not as much as one would expect, and the winner assessed their current level of happiness as the same as it was before their winning. A few years later, Brickman jumped off a ten-story building—after years of depression. He accepted that his fixed point was depression, and there was no long-run escape (Senior 2020)?

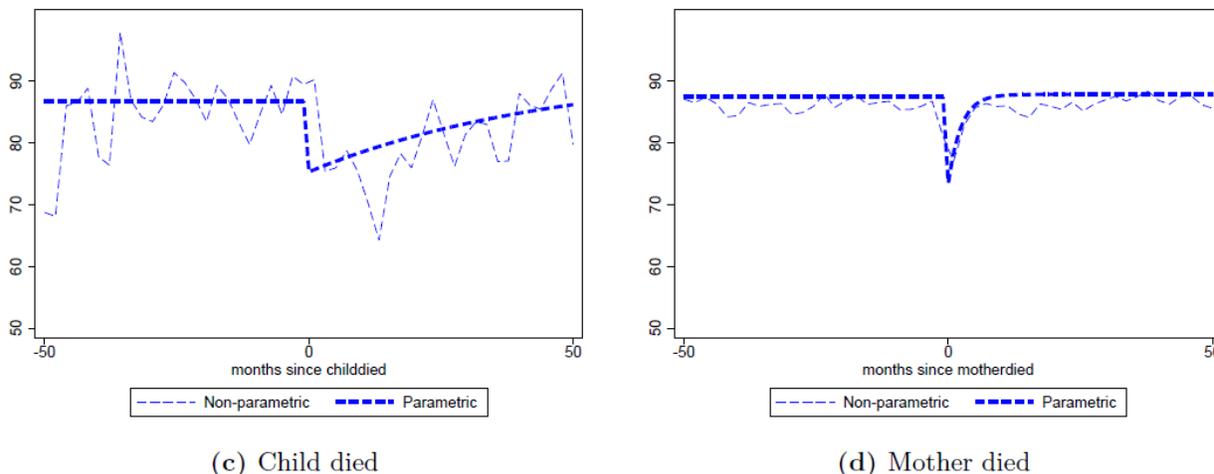


Figure 1, Chapter 5: Part of Figure 1 from Kimball, Nunn, and Silverman (2015)

The early set-point research suggested a lifetime baseline of happiness determined by genetics and one's infancy. However, more recent research by the same researchers suggests the point is quasi-fixed rather than a hard fix: something that can, with effort, be modified. But not necessarily with more goods.

In a seminal 1996 study, [David Lykken](#) and Auke Tellegen, then psychology professors at the U. of Minnesota, interviewed thousands of reared-apart identical twins. They found that 44% to 52% of the variation in reported happiness across individuals could be attributed to genetics. After resampling for years, they concluded that over time, 80% of an individual's stated happiness was inherited. Differences in income, marital status, education, and other observed socioeconomic characteristics could explain only a few percent of the variation across individuals. Put simply, they found that how happy you are, compared to me, is primarily determined by the genes we inherited. Their paper ends with the quote,

It may be trying to be happier is as futile as trying to be taller and therefore counterproductive.

In 2000, after more research, Lykken tempered his conclusion, stating that while happiness is partially determined by genes, environment is also influential. He concludes you can increase your fixed point by changing your attitude and developing solid relationships. Both require conscious effort.

In 2005, the psychologists Frank Fujita and [Edward Diener](#) addressed whether one's life satisfaction is predetermined. By following 2,336 men and 2,873 women over 17 years, they

found that while there is a "soft baseline" for life satisfaction, "some individuals do change significantly", and small fluctuations in happiness are common, some lasting for years.

Personality

One hypothesis to explain a quasi-fixed baseline happiness is through personality: assume personality is a prime determinant of one's WB and that personality is quasi-fixed.

In the [Big Five model of personality](#), the dominant model in psychology, personality has five dimensions: openness to experience, conscientiousness, extroversion, agreeableness, and neuroticism, and one's position on these dimensions can be determined using standard tests.

Numerous studies indicate that neuroticism and extroversion are strong predictors of WB. The more neurotic one is (the extent to which one reacts to setbacks, risks, and ignorance with negative emotions), the less happy. Extroverts tend to spend more time socializing, which is why some researchers hypothesize that extroverts report more life-satisfaction WB. Being an introvert, I don't buy this. The other three dimensions influence whether an individual is likely to put himself in WB-enhancing situations. One study estimated that at least 35% of the variation across individuals in life satisfaction WB is due to personality variation (Wood, Joseph, and Maltby 2008).²⁹

The historical view was that personality is effectively fixed by adulthood. If personality is fixed and affects life satisfaction, it is accounted for in WB studies with *fixed effects*. These individual-specific constant terms capture all the reasons why individuals differ, including, but not limited to, their personality differences. Fixed effects are how individual idiosyncrasies are typically accounted for in statistical models.

But, more recent research shows that personality is malleable, but changes take time. An Australian panel data set follows the same households over time, asking the respondent Big Five personality questions and a life-satisfaction question ("How satisfied are you with your life, all things considered? a 10-pt. scale). The data set also includes household income, job status, and other possible determinants of life satisfaction. Based on this data from over eight-thousand individuals, Christopher Boyce, [Alexa Wood](#), and [Nattavudh Powdthavee](#) find that personality variation across individuals explains more of the variation in life satisfaction than do variations

²⁹ Boyce, Wood, and Powdthavee (2013) survey this literature

in income and job status—confirming previous findings. And, for an individual, more of the changes in life satisfaction over time is explained by personality changes than income and job-status changes.

What are the implications if WB is quasi-fixed for the question: whether consuming more increases WB? While one can tell a story consistent with more consumption more WB, at least in the short-run, if WB is quasi-fixed, increasing consumption will not permanently increase WB—if WB is quasi-fixed, Smith was mostly correct. The quasi-fixed research suggests it isn't more consumption that increases long-term WB, but more and better relationships and conscious efforts to improve how you live and experience your life: actions such as meditation, exercise, therapy, and religion.³⁰ SSRIs, and other depression and anxiety-reducing drugs, can also make life more enjoyable for long periods. At this point, I expect you might conclude, "Morey has lived in Boulder, a rich, liberal enclave, for too long, so he thinks the road to salvation is to give up one's worldly pursuits for a yoga mat." In my defense, I don't do real yoga, only geezer yoga; I don't own a mat.

Many economists implicitly reject the hypothesis that WB is quasi-fixed. One possible reconciliation is that an increase in income beyond what is needed to fulfill basic needs does not increase long-run emotional WB but does increase long-run life satisfaction. Note that economicus can have a quasi-fixed WB: there isn't a thing in the assumptions of NBT that precludes this possibility.

Chapter 6 will review the psychological studies on how and why we adapt to good and bad events. The findings are consistent with a soft happiness baseline. For example, a recent study found that patients on dialysis were no less happy after getting used to dialysis than was a control group.

WB is created by doing, not by acquiring?^{xiii}

Picture a New Yorker cartoon from June 23, 2014: a dog, having climbed the mountain, attentively waits on the words of wisdom from the monk sitting cross-legged in front of his cave.

³⁰ People actively involved with religion are, on average, happier than those who are not; a likely reason is the social involvement and support. Therapies such as cognitive behavioral therapy can be effective in reducing anxiety, so can exercise and meditation.

The caption: *The bone isn't the reward—digging for the bone is the reward.* My first thought was that the monk does not understand dogs, but, on second thought, he does.



Increasing their incomes would have decreased their emotional and life-satisfaction WB if it relieved them from their battles with rabbits and rocks. In contrast to income, health and vigor allowed them to battle longer. N.B. Russell's choice of the word "conquest". Work, not consumption, produced their WB.

Job satisfaction in the U.S. isn't increasing; in 2013, fewer than half of American workers were satisfied with their jobs; in 1987, when the Conference Board's survey started, it was 61%.

Mihály Csíkszentmihályi's Flow^{xiv}

The psychologist Mihály Csíkszentmihályi pioneered the survey technique of contacting people, at random times of the day, to ask what they were doing and how they were feeling (*experience sampling*). Based on his findings, in 1990, he published *Flow: The Psychology of Optimal Experience*. A president of the American Psychological Association, [Martin Seligman](#), has called him the world's leading researcher in "positive psychology".³¹

Flow is an experience, the experience of being so involved in an activity that one loses one's sense of time and oneself. One is motivated and engaged with no desire to be elsewhere—one is "in the zone", focused, and goal-directed. There is neither anxiety nor depression. The experience is enjoyable and rewarding, but unconsciously rather than self-reflectively. While

³¹ Born in 1934, Csíkszentmihályi still actively writes and lectures. See his recent TED lecture, http://www.ted.com/talks/mihaly_csikszentmihalyi_on_flow.html

flowing, you would never say, "Boy, Am I having fun!"; it would not cross your mind; you're too busy doing what you are doing. Russell's gardener and well-digger achieved flow daily.

Csikszentmihályi introduced flow to modern Western science, but he isn't the originator; it, and the benefits of achieving it, is foundational in Taoism, Hinduism, and Buddhism.

Flow is achieved by performing activities with the right amount of challenge: too little, and one becomes bored; too much and frustration results. Appropriate challenges match your intelligence, education, and skill set. Ideally, one's work is a flow experience: one faces problems and tasks that are challenging but doable. Social interactions can, but often don't, produce flow. And, unless one is pretty dull, watching TV will not generate flow, but playing a suitable video game will. Flow requires the perception of control: one must perceive influence and a prospect of success. Individuals who are self-motivated and curious are more inclined to experience flow.

From a Western perspective, flow activities are non-traditional happiness. The experience can be fleeting, but it can also last for extended periods (artists and scientists can spend hours or days engrossed in a project that produces flow). It can be repeated if one chooses suitable activities. Playing games, playing a musical instrument, and doing sports can be flow activities, and so can hobbies. For me, writing this book produces flow. Flow isn't sensual pleasure (like sex, scratching, or a great taste), nor does it necessarily involve doing anything meaningful. One can get flow from doing crossword puzzles but doing them does not give one's life meaning or contribute much to the greater good. While, after the fact, people describe flow experience as enjoyable, the enjoyment isn't conscious during flow.

The pursuit of flow is consistent with the Assumptions of NBT; it only requires that paths that generate flow are high in one's ordering. Flow WB does not depend on one's income if one has enough income to meet one's basic needs; it typically does not require acquiring more possessions. Flow WB is inconsistent with a job numbing in its boredom or far beyond one's ability. Flow outside of work requires the time to peruse flow activities. Interestingly, many do not pursue flow experiences but seek distractions like watching TV.

Buddhist Flow: it's the job, not what the wages allow you to consume^{xv}

Economists typically consider work a necessary evil in the pursuit of consumables. The Buddhist view is that an ideal job, *Right Livelihood*, not consumption, makes us better off (ends our

suffering and awakens us). Right Livelihood is part of the [Buddhist Eight-fold path](#). E.F. Schumacher, in his often reprinted "Buddhist Economics," referring to the modern economist, says,

His fundamental criteria for success is simply the total quantity of goods produced during a given period of time.... From a Buddhist point of view, this is standing the truth on its head by considering goods more important than people and consumption as important as creative activity....For the modern economist, this is difficult to understand. He is used to measuring the 'standard of living' by the amount of annual consumption, assuming ... a man who consumes more is 'better off'..

Quoting earlier in the essay,

The Buddhist point of view takes the function of work to be at least three-fold: to give man a chance to utilise and develop his faculties; to enable him to overcome his ego-centredness by joining with other people in a common task; and to bring forth goods and services needed for a becoming existence. To organise work in such a manner that it becomes meaningless, boring, stultifying, or nerve-racking for the worker would be little short of criminal;...Equally, to strive for leisure as an alternative to work would be considered a complete misunderstanding of one of the basic truths of human existence, namely that work and leisure are complementary parts of the same living process and cannot be separated without destroying the joy of work and the bliss of leisure.

Consider Adam Smith's famous pin factory (it produces straight pins). Smith compares the number manufactured if each pin is manufactured by one worker (he cuts the wire, sharpens one end, etc.) to specialization where each worker does only one of eighteen steps. Smith convincingly shows that specialization will lead to more pins per day (it's more efficient—increased output with no input increase). For a Buddhist, this conclusion, while correct, misses the point. With specialization, the work is tedious and stultifying, so, in the words of Schumacher, “little short of criminal”.³² At this point, an economist would say that specialization

³² “To take an example, therefore, from a very trifling manufacture; but one in which the division of labour has been very often taken notice of, the trade of the pin-maker; a workman not educated to this business (which the division of labour has rendered a distinct trade), nor acquainted with the use of the machinery employed in it (to the invention of which the same division of labour has probably given occasion), could scarce, perhaps, with his utmost industry, make one pin in a day, and certainly could not make twenty. But in the way in which this business is now carried on, not only the whole work is a peculiar trade, but it is divided into a number of branches, of which the greater part are likewise peculiar trades. One man draws out the wire, another straightens it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on, is a peculiar business, to whiten the pins is another; it is even a trade by itself to put them into the paper; and the important business of making a pin is, in this manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands, though in others the same man will sometimes perform two or three of them. I have seen a small manufactory of this kind where ten men only were employed, and where some of them consequently performed two or three distinct operations. But though they were very poor, and therefore but indifferently accommodated with the necessary machinery, they could, when they exerted themselves, make among them about twelve pounds of pins in a day. There are in a pound upwards of four thousand pins of a middling size. Those ten persons, therefore, could make among them upwards of forty-eight thousand pins in a day. Each person, therefore, making a tenth part of forty-eight thousand pins, might be considered as making four thousand eight

will lead to much higher wages, which will compensate for the tedium. They would be correct if unfulfilling work and goods are WB-commensurable. Buddhists reject this.

Buddhism rejects the notion that increased income leads to increase happiness. Economicus could have an ordering based on Buddhist principles: there isn't anything in my NBT that implies that more WB requires more consumption or possessions.

One critical factor for work satisfaction is workplace trust (trust in management and co-workers), and work satisfaction is strongly and positively linked to life satisfaction WB. A 2009 study finds that moving up one place on a 10-point scale of workplace trust has the same effect on life satisfaction WB as a forty-percent bump in income—Wow. Unfortunately, there is little research on how WB varies by job and profession. The evidence on whether the self-employed are happier is mixed. One study finds a self-employment happiness bump but only for the self-employed that are rich.

Decreasing your ill-being by being less reactive^{xvi}

In this Buddhist perspective, a typical life includes a lot of ill-being (unhappiness). And much of it is unnecessary but difficult to avoid because evolution has predisposed us to react in ways that cause ill-being—*dukkha* (Pāli).³³ *Dukkha* is often translated as “suffering” but is not just physical suffering. It includes existential despair and unsatisfactoriness. So, the way to increase your WB is to reduce unnecessary ill-being.

I discuss this Buddhist perspective because while it rejects increased wealth and status as ways to greater WB, an economicus could be on a Buddhist path. And, much of what modern psychology and behavioral economics say about behavioral quirks can be found in historical

hundred pins in a day. But if they had all wrought separately and independently, and without any of them having been educated to this peculiar business, they certainly could not each of them have made twenty, perhaps not one pin in a day; that is, certainly, not the two hundred and fortieth, perhaps not the four thousand eight hundredth part of what they are at present capable of performing, in consequence of a proper division and combination of their different operations.” (Smith, 1776: Book I, Chapter I)

³³ This is one view, my view, of how Buddhism would describe the Buddhist process of decreasing your personal ill-being. [Buddhism as an ethic is considered in Chapter 11.] While I am not a Buddhist, I have interest in it from ethics, WB, and choice perspectives. Some Buddhists would assess my view a wrong view. There are different schools with different Canons, and statements within them can conflict. There is debate as to which are the Buddha's original words. Nothing was written down until hundreds of years after he died (c.480-c.400 BCE), and his belief system was first written in Pāli, now extinct—it is unlikely that he even spoke it. In summary, there are many current views of historical and modern Buddhism—like with Christianity.

Buddhism—see Chapters 6 and 7. [And, looking ahead, WC and Buddhist ethics have a lot in common.]

Dukkha comes with the bad and the good; it is the way of life:

Birth is duhhka, aging is duhhka, sickness is duhhka, death is duhhka, encountering what is not dear is duhhka, separation from what is dear is duhhka, not getting what one wants is duhhka (Bodhi 2000, p. 1844).

While dukkha has multiple sources, one cause is not accepting that life and its components are impermanent. But impermanence is what makes a life a life. All relationships will end; all possessions will be lost. Pleasures, such as sex, drugs, and rock-n-roll are tainted: one knows the sensations will soon be over, one knows one won't always have the ability, the act can be illicit, and it is often regretted in the morning. [Giacomo, my dog, unlike humans, does not suffer from knowing all is impermanent—lucky him, but he suffers in other ways]

The first step in reducing dukkha is to understand that, typically, life is dukkha (putting it too negatively—life is shit, and then you die.)

A cause is *taṇhā* (Pāli: craving, thirsting, desiring, reacting). Buddhism teaches that humans naturally order paths based on wants and desires (Assumption 9b), and this is misguided: it is negatively correlated with an ordering based on WB (Assumption 9a). As I noted in Chapter 1, “craving”, is a synonym for “wanting and desiring” but with a negative edge, chosen to draw the critical Buddhist distinction between want fulfillment and WB.

If we understand dukkha as shorthand for ‘life’, we can think of taṇhā as shorthand for the myriad reactions that life provokes in us (Stephen Batchelor 2015).

Through years of meditation and study, a select few can reach a state where they no longer crave. But most of us will never attain that state: we will continue to want and desire. But with effort, we can become less responsive/reactive to our wants and desires—the itch does not always need to be scratched.

Being sentient and the product of evolution, I interpret what is happening as “it is happening to me”—there is a me being bombarded by sensations. This me interprets them from an “I” perspective: “I am hungry”, “I thought of you.” My arm hurt after I got vaccinated.” “He insulted me.” “I will go to Italy next year.” This “I” becomes attached to wants and desires.

When I say “I”, I am referring to my *self*; when I say “you”, I am referring to your *self*.

Buddhism teaches there is no self in the person I think of as me: *no-self* (*anata* in Pāli). It is only a convenient fiction: a way to identify the object you refer to, a “convenient designator”.

And my false belief that there is a me causes unnecessary ill-being: with “me” comes “mine” (no me implies no mine). This leads to attachments “in the form of desires and aversions, and the strengthening of ignorance concerning the true nature of sentient existence.” ([Mark Siderits](#) 2019). Saying “mine” comes early in life.

You are reasonably asking, “If there is no me, who is experiencing this ill-being?” Consider your cat and dog. While they experience pain, excitement, and hunger, they do not associate the feeling with their self—they do not have one. Likewise, spiders feel pain, but no one argues that spiders have a self. A Buddhist would say that part of the difficulty of understanding no-self is the limitations of language and the multiple meaning of pronouns and proper names.

“The Buddha’s ‘middle path’ strategy first argues there is nothing that the word ‘I’ genuinely denotes and then explains that our erroneous sense of an ‘I’ stems from our employment of the useful fiction represented by the concept of the person.”

Searching for what makes me me and not finding it is not unique to Buddhism. In lockstep with Buddhism, Hume could not find the self.³⁴ [The Berkeley psychologist [Alison Gopnik speculates that Hume might have known of Buddhism](#)]. See also the section in Chapter 6 on the future-self empathy-gap.] A person is a collection of components that are changing over time, and no one component is that person’s self: not their thoughts nor sensations, not their heart, left arm, nor temporal lobe: are all impermanent.³⁵ One sense of self is your one-and-only lifelong control center (the enduring executive in charge). Buddhism rejects this assertion that various and varying processes take charge; there is no one controller, nor even one for a particular function. The processes that caused you, one day, to buy donuts and marry Mary are

³⁴ Descartes believed that one’s self existed and was one’s enduring mind: “Cogito ergo sum” (I think therefore I am). In explanation, the mind is the source of one’s thoughts which it produces in a continuous stream. [Recollect that Descartes believed the mind transcends the brain: he was a dualist.] Hume and Buddhism reject this, arguing that one’s thoughts are brief, distinct, and impermanent, so they do not imply the existence of a mind. One’s thoughts could be the product of an impermanent brain. But, since thought are so ubiquitous, we have the illusion of a mind, and then wrongly define this illusion as the me in me.

³⁵ Echoing this view, the philosopher [Norman O. Brown](#) (1966) defined a person as a corporation of persons, “a corporate body”. “The unity of a person is as real, or unreal, as the unity of a corporation.” Like, for example, Apple: it is a collection of ever-changing people.

not necessarily the same processes that drove you yesterday to go skiing, start with a Green Run, and have wine with lunch. In defense of this view, neurobiologists can't find your executive.

I sense that somewhere inside my container, something constitutes the real me. But since the container is the result of a process to enhance the passing along its genes, one might expect me to have a sense of self even if it's just an illusion—the *illusion-of-self*—it motivates me to pass along my genes.³⁶ To foster the fantasy, one “learns to think of life as a kind of narrative” (Siderits 2007) with the story revolving around you, so strive, too much, for importance and significance.

My point is not to convince you that there is no you in you. Instead, it's to point out that the Buddhist view is that most of us suffer from an illusion-of-self, which causes unnecessary ill-being—you are free to disagree. That said, thinking about no-self makes me think about self as a continuum: at one extreme is no-self; at the other extreme, there is a definite me in me—what most of us believe. Somewhere in the middle is a fluid self. Thinking about whether one has a self makes one think about their desires and cravings more nuancedly.

So, the Buddhist view is that most people don't understand that life is suffering, all is impermanent, and there is no me, so no mine. If we accept this, our ill-being will lessen—cease—objects and experiences will have less “mine” attached to them. We will no longer expect permanence, so react more appropriately to the sensations (good and bad) that life throws at us—our attachments to these sensations will have waned.

In lockstep, trying to fulfill your wants and desires will be counterproductive if the goal is to reduce your ill-being. As noted above, the Buddhist view is that our default ordering of paths is based on cravings (wants and desires) rather than WB, and the two orderings are negatively correlated. Fulfilling desires brings pleasures, but, like everything else, these pleasures are impermanent and simply cause more craving: giving Ralph, the rat, a chunk of cheddar for pushing the bar just causes him to crave and push more.

You need enough resources to fulfill your and your family's need for food, shelter, and safety, but more could be counter-productive if the goal is less ill-being (more WB). I recently read, “Here is how the bored rich are spending their extra cash” (03/20/2021NYT). The Pandemic has been great for many rich white guys who spend their quarantine hours online perusing and buying vintage Rolexes, Porsches, and Pokémon cards.

³⁶ My sense of self motivates me to keep living and to have kids. Chapters 9 and 10 consider the *illusion of choice*.

“All I do is go through watch porn,” he said. “I’m selling watches. I’m buying watches. It’s crazy. I have no reason right now to buy a watch. I’m at home all day at a computer. Time is staring me right in the face. What reason do I have to look at my wrist? But I want a tangible sign of something, so I’m looking at watches.”

[Prices are rising, so it's “investing”.]

How to get liberated from suffering is beyond my scope, but, in summary, the Buddhist view is it is possible but difficult, and one gets gradually closer by adhering to the eight-fold path:

And this is the path: the path with eight branches: complete view, complete thought, complete speech, complete action, complete livelihood, complete effort, complete mindfulness, complete concentration (Bodhi 2000, p. 1844).

Meditating is necessary because achieving liberation from one’s thoughts and sensations is, otherwise, unattainable. It teaches you to look at your thoughts and sensations in new ways, encouraging you to understand and accept no-self and impermanence.

Earlier, I suggested that an economicus could be a Buddhist, a *Buddhist-economicus*. Examples of economici that are not Buddhists prove that they don’t have to be.³⁷ An economicus could be at a point in life where their ordering of paths is based on understanding suffering, impermanence, and no-self. Their wants and desires line up with the Buddhist conception of ill-being. [Assumptions 9a](#) and [b](#) both hold. And the individual is liberated. [The goal, in words, is minimizing ill-being (paths ordered from most to least ill-being), not maximizing WB. The difference is the perspective: Buddhists think dukkha is pervasive, and neoclassical economists have a more optimistic view of the human condition. For Buddhists, framing the goal as maximizing WB is folly. Minimizing suffering has, at least for me, more appeal as an ethic than does maximizing aggregate WB.]

But can a Buddhist really be an economicus? Buddhists believe that most humans order paths based on their wants and desires, and this ordering does not line up with the individual’s WB, which is all consistent with NBT. However, they also believe that humans, with difficulty, can choose a path that is not one of their highest-ranked available paths, which violates [Assumption 7: At every point in time, economicus takes one of their HRAPs](#). So, as envisioned by Buddhism, a human cannot be an economicus. But what about enlightened individuals? [Assumption 9a](#) and [b](#) hold, and they experience their HRAP, but they will have chosen to: they were not required to, so while [Assumption 7](#) is violated, they behave as if it is not.

³⁷ My Roomba vacuum cleaner being one example, worms, probably, another.

The next chapter turns to the NBT assumption that one has a stable ordering of paths (Assumption 4) rather than one that varies with emotional state or changes as a function of the consumption paths one has consumed.

ⁱ Sources:

1. Andrew Clark, Paul Frijters and Michael Shields (2008) Relative income, happiness and utility: an explanation for the Easterlin paradox and other puzzles, *J. of Econ. Literature* 46(1): 95-144.

ⁱⁱ Sources:

2. David Blanchflower and Andrew Oswald (2004) Well-being over time in Britain and the USA. *J. of Pub, Econ.* 88:1359-1386.
3. Edward Diener and Shigehiro Oishi (2000) Money and happiness. In *Culture and Subjective Well-Being* (eds. Edward Diener and E. Suh), Cambridge, MIT Press
4. Paul Dolan, Tessa Peasgood and Mathew White (2008) Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *J. of Econ. Psych.* 29: 94-122.
5. David Dorn, Justina Fischer, Gebhard Kirchgassner and Alfonso Sousa-Poza (2007) Is it culture or democracy? The impact of democracy, and culture on happiness. *Soc. Indic. Res.* 82(3): 505-526.
6. Richard Easterlin (1974) Does economic growth improve the human lot? Some empirical evidence. In *Nations and Households in Economic Growth: Essays in Honour of Moses Abramowitz* (ed. P. A. David and M. W. Reder). N.Y. and London: Academic Press.
7. Richard Easterlin (1995) Will raising the incomes of all increase the happiness of all? *J. of Econ. Behav. and Organ.* 27: 35-48.
8. Ada Ferrer-i-Carbonell (2005) Income and well-being: An empirical analysis of the comparison income effect, *J. of Pub. Econ.* 89: 997-1019.
9. John Helliwell (2003) How's life? Combining individual and national variables to explain subjective well-being. *Econ. Mod.* 20: 331-360.
10. Ronald Inglehart and Hans-Dieter Klingemann (2000) Genes, culture, democracy and happiness, In *Culture and Subjective Well-Being* (eds. Ed Diener and E. Suh), Cambridge, MIT Press.
11. Erzo Luttmer (2005) Neighbors as negatives: relative earnings and well-being, *The Quart. J. of Econ.* 120(3): 963-1002.
12. Andrew Oswald (1997) Happiness and economic performance. *The Econ. J.* 107: 1815-1831.
13. R. Veenhoven (1993) *Happiness in Nations: Subjective appreciation of life in 56 nations, 1946-1992*, Rotterdam, Risbo Press.

ⁱⁱⁱ Sources:

1. Daniel Benjamin, Ori Heffetz, Miles Kimball and Nichole Szembrot (2014) Beyond happiness and satisfaction: toward well-being indices based on stated preferences, *Amer. Econ. Rev.* 104(9): 2698-2735.
2. David Blanchflower and Andrew Oswald (2008) Hypertension and happiness across nations, *J. of Health Econ.* 27: 218-233.
3. Edward Diener (2000) Subjective Well-being: The science of happiness and a proposal for a national index, *Amer. Psych.* 55: 34-43.
4. L.W. Sumner (1996) *Welfare Happiness and Ethics*, Oxford, Oxford University Press.

^{iv} Sources:

1. Andrew Oswald (1997) *ibid*

^v Sources:

1. John Helliwell (2003) *ibid*.
2. Andrew Oswald (1997) *ibid*

^{vi} Sources:

1. Rafael Di Tella, John Haisken-DeNew and Robert MacCulloch, (2010) Happiness Adaptation to Income and to Status in an Individual Panel, *J. of Econ. Behavior and Organization* 76(3): 834-852.

^{vii} Sources:

-
1. Daniel Kahneman and Angus Deaton (2010) High income improves evaluation of life but not emotional well-being, *Proc. of the National Academy of Sci.* 107(38): 16489-93.

viii Sources:

1. Jan-Emmanuel De Neve, George Ward, Femke De Keulenaer, Bert Van Landeghem, Georgios Kavetsos and Michael Norton (2015) The asymmetric experience of positive and negative economic growth: global evidence using subjective well-being data, Centre for Economic Performance Discussion Paper 1304, Said Business School, University of Oxford. Forthcoming *The Rev. of Econ. and Stat.*
<http://www.sbs.ox.ac.uk/community/people/jan-emmanuel-de-neve>

ix Sources:

1. Chicago Project on Security and Threats; web page on domestic extremism (accessed April 8, 2021)
https://post.uchicago.edu/research/domestic_extremism/
2. David Brooks (04/22 2021) The G.O.P, is getting even worse. NYT.
3. Leo Damrosch (2005) *Jean-Jacques Rousseau: Restless Genius*, Boston, Houghton Mifflin Company.
4. Shai Davidai and Thomas Gilovich (2015) Building a more mobile America—one income quintile at a time. *Persp. on Psych. Sci.* 10(1): 60-71.
5. Alan Feuer (April 4, 2021) Fears of white people losing out permeate capital rioters' town, study finds. NYT.
6. Daniel Goleman (2013) Rich people just care less, *N.Y. Times*, October 5. This article describes the research of Michael Kraus, Dacher Keltner and others. <http://opinionator.blogs.nytimes.com/2013/10/05/rich-people-just-care-less/>
7. Raymond Klibansky and Ernest Mossner eds. (1954) *New letters of David Hume*, Oxford, Oxford University Press. <https://books.google.com/books?isbn=0199693234>
8. Michael Kraus, Shai Davidai and David Nussbaum (2015) American Dream? Or Mirage *N.Y. Times*, May 1. <http://www.nytimes.com/2015/05/03/opinion/sunday/american-dream-or-mirage.html>
9. Michael Kraus and Dacher Keltner (2009) Signs of socioeconomic status, *Psych. Sci.* 20(1): 99-106.
10. Michael Kraus and Dacher Keltner (2013) Social class rank, essentialism and punitive judgment, *J. of Person. and Soc. Psych.* 105(2) 247-261.
11. Michael Kraus and Jacinth Tan (2015) Americans overestimate social class mobility. *J. of Exp. Soc. Psych.* 58:101-111.
12. Melissa Lane (2008) Melissa Lane on Rousseau on Modern Society, Philosophy Bites. Leon Festinger (1954) A theory of social comparison processes. *Human Relations* 7(5): 117-114. At www.humanscience.org/docs/Festinger%20%281954%29%20A%20Theory%20of%20Social%20Comparison%20Processes.pdf
13. Melissa Lane (2008) Melissa Lane on Rousseau on Modern Society, Philosophy Bites. <http://philosophybites.com/2008/07/melissa-lane-on.html>
14. Erzo Luttmer (2005) Neighbors as negatives: Relative earnings and well-being, *The Quart. J. of Econ.* 120(3): 963-1002.
15. Edward Morey and David Kritzberg (2012) It's not where you do it, it's who you do it with? *The J. of Choice Mod.* 5: 176-191. <http://www.colorado.edu/economics/morey/papers/MoreyKritzbergJOCM2012.pdf>
16. William Noyes (1890) The insanity of Jean Jacque Rousseau, *The Amer. J. of Psych.* 3(3): 406-29. http://www.jstor.org/stable/1411748?seq=1#page_scan_tab_contents
17. *Organizational Behavior and Human Decision Processes* (2007) Special issue on social comparison processes.
18. Robert Pape (April 4, 2021) Opinion: What an analysis of 377 Americans arrested or charged in the Capital insurrection tells us. *Washington Post*.
19. Paul Piff, Michael Kraus, Stephane Cote, Bonnie Hayden Chen and Dacher Keltner (2010) Having less, giving more: the influence of social class on prosocial behavior. *J. of Person. and Soc. Psych.* 99(5): 771-784.
20. Stephen Reicher and Yasemin Ulusahin (2020) Resentment and Redemption: On the Mobilisation of Dominant Group Victimhood. Chapter 13 in *The Social Psychology of Collective Victimhood* (ed. Johanna Vollhardt). Oxford: Oxford University Press.
21. Thorstein Veblen (1899) *The Theory of the Leisure Class*. <http://www.gutenberg.org/etext/833>
22. Tim Wadsworth (2014) Sex and the pursuit of happiness: how other people's sex lives are related to our sense of being, *Soc. Indic. Res.* 116(1): 115-135.

x Sources:

1. Kaare Christensen, Anne Maria Herskind and James Vaupel (2006) Why Danes are smug: comparative study of life satisfaction in the European Union, *British Med. J.* 333: 1289-1291.
2. Paul Dolan, Tessa Peasgood and Mathew White (2008) *ibid*
3. Alois Stutzer (2004) The role of income aspirations in individual happiness, *J. of Econ. Behav. and Org.* 54: 89-109.

^{xi} Sources:

1. Alberto Alesina, Edward Glaeser and Bruce Sacerdote (2005) Work and leisure in the United States and Europe: why so different? in NBER Macroeconomics Annual 2005 (Vol. 20), edited by Mark Gertler and Kenneth Rogoff. <http://www.nber.org/chapters/c0073.pdf>
2. Robert Frank (1997) The frame of reference as a public good, *The Econ. J.* 107:1832-47. At robert-h-frank.com
3. Paul Krugman (2005) French family values, N.Y. Times, July 25. http://www.nytimes.com/2005/07/29/opinion/29krugman.html?_r=0
4. Annie Lowrey (2012) For two economists, the Buffett rule is just a start, *The N.Y. Times*, April 16. <http://www.nytimes.com/2012/04/17/business/for-economists-saez-and-piketty-the-buffett-rule-is-just-a-start.html>
5. Annie Lowrey (2013) The rich get richer through the recovery, *The N.Y. Times*, September 10. At <http://economix.blogs.nytimes.com/2013/09/10/the-rich-get-richer-through-the-recovery/>
6. Emmanuel Saez (2013) Income inequality: evidence and policy implications, *The Arrow Lecture*, Stanford. http://eml.berkeley.edu/~saez/lecture_saez_arrow.pdf

^{xii} Sources:

1. Christopher Boyce, Alex Wood and Nattavudh Powdthavee (2013) Is personality fixed? Personality changes as much as 'variable' economic factors and more strongly predicts changes in life satisfaction. *Soc. Indic. Res.* 111(1): 287-305. <https://link.springer.com/article/10.1007/s11205-012-0006-z>
2. Philip Brickman, Dan Coates, and Ronnie Janoff-Bulman (1978) Lottery winners and accident victims: Is happiness relative? *J. of Person. and Soc. Psych.* 36(8): 917-27.
3. Frank Fujita and Edward Diener (2005) Life satisfaction set point: stability and change, *J. of Person. and Soc. Psych.* 88(1): 158-164. <http://academic.udayton.edu/jackbauer/Readings%20595/Fujita%2005%20happy%20set%20point%20copy.pdf>
4. Miles Kimball, Ryan Nunn and Dan Silverman (2015) Accounting for adaptation in the economics of happiness, NBER working paper series, National Bureau of Economic Research, Cambridge MA. <http://www.nber.org/papers/w21365.pdf>
5. David Lykken (2000) *The Nature and Nurture of Joy and Contentment*, N.Y.: St. Martin's Press.
6. David Lykken, T.J. Bouchard, K.J. Wilcox, N.L. Segal and S. Rich (1988) Personality similarity in twins reared apart and together. *J. of Person. and Soc. Psych.* 54: 1031-39.
7. David Lykken and Auke Tellegen (1996) Happiness is a stochastic phenomenon, *Psych. Sci.* 7(3): 186-189.
8. Jennifer Senior (Nov. 24, 2020) Opinion: Happiness won't save you, *New York Times*.
9. Adam Smith (1759) *The Theory of Moral Sentiments* http://books.google.com/books/about/The_Theory_of_Moral_Sentiments.html?id=xVkOAAAAQAAJ
10. Alex Wood, Stephen Joseph and John Maltby (2008) Gratitude uniquely predicts satisfaction with life: incremental validity above the domains and facets of the five factor model. *Person. and Individ. Diff.* 45: 49-54. <https://www.sciencedirect.com/science/article/pii/S0191886908000767>

^{xiii} Sources:

1. The Conference Board (2014) *Job Satisfaction: 2014 Edition*. <https://www.conference-board.org/publications/publicationdetail.cfm?publicationid=2785>
2. Bertrand Russell (1930) *The Conquest of Happiness*, London: Allen and Unwin, Partially available online at http://books.google.com/books/about/The_Conquest_of_Happiness.html?id=-lefVQ_MYYC
3. Lauren Weber (2014) U.S. workers can't get no (job) satisfaction, *Wall Street J.*, June 18. <http://blogs.wsj.com/atwork/2014/06/18/u-s-workers-cant-get-no-job-satisfaction/>

^{xiv} Sources are:

1. Mihály Csíkszentmihályi (1990) *Flow: The Psychology of Optimal Experience*, N.Y., Harper and Row.
2. Christopher Peterson, Nansook Park and Martin Seligman (2005) Orientation to happiness and life satisfaction: the full versus the empty life, *J. of Happiness Studies* 6:25-41.

^{xv} Sources:

-
1. Alberto Alesina, A., Rafael Di Tella and Robert MacCulloch, (2004) Inequality and happiness: Are Europeans and Americans different? *J. of Public Econ.* 88, 2009-2042.
 2. David Blanchflower and Andrew Oswald (2008) *ibid.*
 3. Paula Dolan, Tessa Peasgood and Mathew White (2008) *ibid*
 4. John Helliwell, Haifang Huang and Robert Putnam (2009) How's the job? Are trust and social capital neglected workplace investments? In. *Social Capital Reaching Out, Reaching In* (eds. Viv Ona and James Davis), Edward Edgar.
 5. E.F. Schumacher (1966) Buddhist economics, *Resurgence* 1(1)
 6. Adam Smith (1776) An Inquiry into the Nature and Causes of the. *Wealth of Nations*. Edited by S. M. Soares. MetaLibri Digital. Library, 29th May 2007. https://www.ibiblio.org/ml/libri/s/SmithA_WealthNations_p.pdf
- ^{xvi} Sources:
1. Stephen Batchelor (2015) *After Buddhism, rethinking the dharma for a secular age*, Yale University Press. <http://yalebooks.yale.edu/book/9780300205183/after-buddhism>
 2. Bhikkhu Bodhi 2000. *The Connected Discourses of the Buddha: A New Translation of the Samyutta Nikaya*. Somerville, MA.: Wisdom Publications. [ISBN 9780861713318](https://www.wisdompubs.org/product/ISBN-9780861713318).
 3. Clair Brown (2017) *Buddhist economics: an enlightened approach to the dismal science*. New York, Bloomsbury Press.
 4. Norman Brown (1966) *In Love's Body*, Random House.
 5. Rupert Gethin (1998) *The foundation of Buddhism*, Oxford, Oxford University Press.
 6. Rupert Gethin (2008) *Sayings of the Buddha*, Oxford, Oxford University Press.
 7. Alison Gopnik (2009) Could David Hume have known about Buddhism? Charles Francois Dolu, the Royal College of La Flehe, and the global Jesuit intellectual network, *Hume Studies* 35 (1 and 2): 6-28. [She is the sister of the author Adam Gopnik.]
 8. Alison Gopnik (Oct. 2015) How an 18th-century philosopher helped solve my midlife crisis: David Hume, the Buddha, and the search for Eastern roots of the Western Enlightenment, *The Atlantic*.
 9. Joel Magnuson (2022) *The Dharma and the Socially Engaged Buddhist Economic*. Palgrave Macmillan.
 10. Walpola Rahula (1959) *What the Buddha taught*, N.Y., Grove Press.
 11. Mark Siderits (2007) *Buddhism as philosophy: an introduction*. Indianapolis/Cambridge, Hackett.
 12. Mark Siderits (2019) *Buddha*, SEP
 13. Robert Wright (2017) *Why Buddhism is true: the science and philosophy of meditation and enlightenment*, Simon and Shuster,

Chapter 6: Common quirks, incorrect beliefs, and flawed choosingⁱ

08172022

Psychologists, neuroscientists, and behavioral economists study behavioral regularities. These include *duration bias*, emotional empathy-gaps, endowment effects, incentive salience, and lack of *empathy for one's future self*. Refer to these behavioral regularities as *common quirks*. A *quirk* is a peculiar cognitive habit, “peculiar” as in “strange and surprising”. The adjective “common” because all humans, or at least most of them, share these cognitive habits—they are part of our built-in cognitive makeup.¹ We share many of them with non-human primates. Common quirks need to be contrasted with quirks shared by some, but not most humans, including your unique quirks. Examples of individualistic quirks include an aversion to purple foods and not stepping on sidewalk cracks.

In contrast, the common quirk duration bias is the cognitive habit of overestimating the duration of WB shifts: both up and downshifts. Again, most of us do this. Many psychologists and behavioral economists have spent their careers identifying and studying common quirks—there are thousands of published articles—I list a few.

Common quirks influence everyone's behavior. To flesh out their influence on choice and NBT—the goal of this chapter—it is essential to clarify how they differ from other influences on behavior and choice.

Distinguish between NBT violations, common quirks, incorrect beliefs, flawed choosing, and a wrong choice

While a few common quirks violate at least one NBT assumption, not all do. A common quirk might, or might not, imply that I fail to experience my highest-ranked available path: “common quirk” isn't a synonym for violating an NBT assumption.

And while the common quirk duration bias is a type of *incorrect belief*, not all common quirks are incorrect beliefs, and not all incorrect beliefs are common quirks.

Individuals hold *beliefs*: how they think the world works and what they think will happen if event *A* happens. Beliefs include the subjective probabilities associated with different possible

¹ Terms in the literature include “biases” and “anomalies”. But they are not all “biases”, and an “anomaly” is a deviation from what is common, but they are common.

outcomes. For example, I believe there is a 40% chance it will be sunny tomorrow, and it is unlikely (maybe only a 20% chance) that humans cause global warming.² These probabilities are personal beliefs, a function of how I believe the world works. Another set of beliefs is one's subjective beliefs about how the occurrence of an event will affect their WB. My belief before getting tenure was that denial of tenure would make me miserable for years.

I define an *incorrect belief* as a divergence between what the individual believes is true and the best estimate of truth based on science and the available data. Suppose on the next flip of a coin, I believe with certainty that it will be heads, but it's a new penny, and there isn't anything unusual about the flipper, so the best estimate is 50%, even though it was tails on the first three flips. That I believe my depression will last forever is duration bias, an incorrect belief, because all the data and science indicate a mean duration of three months for people of my ilk.³

NBT does not rule out incorrect beliefs; specifically, both,

Assumption 9a: An economicus's ordering of paths is based on its WB (well-being), WB, accounting for the uncertainties and

Assumption 9b: An economicus's ordering of paths is based on its wants and desires.

are consistent with incorrect beliefs. Assumption 9b puts no restrictions on the origins of wants and desires, and Assumption 9a admits beliefs and subjective probabilities that are not based on the best science and data.

Consider my most recent disaster: I Gorilla-glued duck feathers to my arms, and, flapping frantically, I leapt off my roof. Before the leap, I told my friend Neil that I would fly to France and alight atop the Arc de Triomphe. Neither before nor after my crash could Neil conclude I had violated an assumption of NBT. Maybe my subjective probability of success was 97%, and I knew if I made Paris, I would be joyous. And I have always wanted to fly like a bird. I chose my most-WB available path in terms of my subjective probabilistic WB. But most people, including

² Many people agree—Actually, I don't.

³ As someone who developed surveys to estimate how much an impacted group is damaged by an environmental injury, I am acutely aware that the beliefs respondents hold affect how they answer questions designed to estimate damages. People typically come to the interview with varying beliefs, often incorrect. (Consider an oil spill: people hold incorrect beliefs on how much oil was spilled, the physics and chemistry of how oil is dispersed in water, and the biological mechanisms by which oil affects plants and animals.) My hope was that after the respondents have read the description of the injury, they had a common and correct understanding (held correct beliefs). If not, I was estimating the damages for something other than the actual injury.

Neil and my therapist, would say I made a wrong choice in an abnormal-psychology sense. Psychologists assess selections not by what was selected but by whether the selection process was flawed. In my defense, as an economist, my choosing was flawed but consistent with my being an economicus. There is always a story about a strange ordering, awful constraints, or weird beliefs that would make almost any behavior consistent with NBT. So, I won't try to identify a choice inconsistent with NBT. Instead, I ask whether the behavioral quirks discussed are inconsistent with what NBT assumes about *the process of choosing*.

Individuals can, all else constant, rank a path where their beliefs are mostly incorrect higher than one where their beliefs are mostly correct. Beliefs directly affect WB. I prefer to believe that I will live to 120. Believing that I am going to heaven would also be comforting.

Flawed choosing is choosing in a faulty manner: using a process or set of beliefs that are not expected—based on best estimates—to achieve the maximum available WB. Flawed choosing is an abbreviation for *flawed choosing in terms of WB*. Distinguish between the alternative selected (the choice) and the process of choosing, noting that even if your choosing is flawed, you might still end up experiencing the available path that produces the most WB. Some, but not all, common quirks cause flawed choosing. For example, quirks based on incorrect beliefs cause flawed choosing. Duration bias is an example. My confident expectation that Trump would make America great again was incorrect. If your subjective probabilities differ from best estimates, your choosing is flawed. Incorrect beliefs are often, but not always, the cause of flawed choosing.

While flawed choosing does not imply that every specific choice reduces WB, I suspect it reduces aggregate WB (over time and summed over individuals); there would be more overall WB in the absence of flawed choosing. But I am not sure; it is an untested hypothesis that would be difficult to test. I am not alone in believing that flawed choosing reduces overall WB.

Like common quirks, flawed choosing might result from violating an NBT assumption, but flawed choosing does not imply a violation: NBT admits flawed choosing.

Take note that I have not defined a *wrong choice*. It is difficult to define “a wrong choice” (wrong relative to what?). One might define a choice as “wrong” if the choosing was flawed, but I reject that definition: maybe what I flawfully chose greatly increased my WB.

Three + common quirks (duration bias, our emotional empathy-gap, salience effects, and maybe a 4th, lack of empathy for one's future self). What does each mean for NBT?

The endowment effect gets its own chapter, the next one. First, I discuss *duration bias* because it's straightforward, and understanding it will help in understanding the other common quirks. Duration bias does not violate any of the Assumptions of NBT: it is only an incorrect belief, a systematic and persistent bias in our subjective beliefs about the duration of our own WB shifts.

Table 1 summarizes the relationships between NBT and duration bias. Duration bias violates none of the NBT assumptions but causes flawed choosing in terms of WB.

Table 1 includes the relationships between NBT and five quirks (the emotional empathy-gap, the future-self empathy-gap, incentive salience, and the endowment effect). For now, consider only Column 2.

	<i>Duration bias</i>	<i>Emotional empathy-gap</i>	<i>Endowment effect</i>	<i>Future-self empathy-gap</i>	<i>Incentive Saliency</i>
	No assumptions violated			No assumptions violated	
Assum1: at every pt. in time only one path					
Assum2: at every pt. in time, only 1 ordering					
Assum3: ranking of two Paths can't depend on what else is imaginable					
Assum4: order. does not change in SR		Violated	Violated		Violated
Assum5 must be on an avail. path					
Assum6a: ordering can't depend on which Paths are avail.			Violated		
Assum6b: which Paths are avail. can't depend on ordering					
Assum7 chooses HRAP					
Assum 8 sensations					
Assum9a: ordering WB					Violated
Or					
Assum9b: ordering on wants/desires					Specifically assumed
<i>Incorrect belief</i>	Yes	Yes	It depends	Yes	No
<i>Flawed choosing (wrt WB)</i>	Yes	Yes	No, if WB drop realized. Yes, if not realized	Yes	Yes, but...

Table 1, Chapter 6: Which NBT assumptions are violated by each common quirk

Looking ahead, the emotional empathy-gap and incentive saliency violate only NBT Assumption 4; the endowment effect violates 4 and 6a. [Table 1 has a lot of blank space.] In contrast, three quirks cause flawed choosing. Whether the endowment effect causes it depends on whether the

anticipated WB drop is realized. Flawed choosing is an incorrect belief only if the anticipated WB drop is unrealized. Incentive salience assumes Assumption 9b, so it can't be faulted for causing flawed choosing in terms of WB.

Duration bias:ⁱⁱ

It is the habit of systematically over-predicting the duration of WB shifts. Things you imagine will enhance your emotional WB (new car, new spouse, polar bears saved) will, but not for as long as you imagine. And symmetrically, things you believe will make you unhappy (getting dumped, cancer) will, but not for as long as you imagine.⁴ Gilbert et al. (1998) find that we suffer from it but are unaware that we do.

People are generally unaware of the operation of the system of cognitive mechanisms that ameliorate their experience of negative affect (the psychological immune system), and thus they tend to overestimate the duration of their affective reactions to negative events. This tendency was demonstrated in 6 studies in which participants overestimated the duration of their affective reactions to the dissolution of a romantic relationship, the failure to achieve tenure, an electoral defeat, negative personality feedback, an account of a child's death, and rejection by a prospective employer. Participants failed to distinguish between situations in which their psychological immune systems would and would not be likely to operate and mistakenly predicted overly and equally enduring affective reactions in both instances. The present experiments suggest that people neglect the psychological immune system when making affective forecasts.

While duration bias is an incorrect belief, it isn't a bias in the subjective probability that event *A* will occur. Instead, it is an overestimate of the duration of the WB shift that *A* would cause. Biases in anticipated WB shifts could cause Path *j* to be ranked higher than Path *d* even though Path *d* would be higher if the durations were more accurately anticipated.

Duration bias is one source of flawed choosing, and flawed choosing will eventually cause realized WB to be lower than it could have been: our ordering is based on an incorrect belief. Because of duration bias, we spend too much time and money acquiring new items and situations that will produce a constant stream of a positive attribute (e.g., clearer TV picture, Modigliani over the mantle, saved rain forest). And we spend too much money and time trying to avoid or eliminate items and situations that will produce a constant negative attribute. We spend too much effort trying to stay healthy, too much on eliminating ugly stuff from our new house,

⁴ While the observed bias, on average, is people believe the WB shift will last longer than it will, some people believe the WB shift will not last as long as it will.

Related to duration bias is magnitude bias. It is a bias in the magnitude of the immediate shift in emotional WB. Magnitude bias is less studied than duration bias but can also cause one's ordering of paths to be out of sync with realized WB.

and too much time complaining to the HOA about our neighbor's new landscaping and house paint. Why? Because we overestimate the long-term WB implications of these long-lasting changes. Imagine you become incontinent or, worse, paralyzed from the waist down. I believe paralysis would guarantee misery until the day I die. But I am incorrect. I don't understand that I will adapt, maybe not entirely but way more than I could now imagine. Duration bias is an assault on the economic assertion that we typically choose, given our constraints, our highest-ranked path. Still, it does not violate any of the Assumptions of NBT.

More examples of duration bias:

Thinking back to before I got tenure (all universities make mistakes), I thought being denied would ruin my life; that getting it would permanently increase my WB (both emotional and life-satisfaction). Unfortunately, tenure didn't permanently increase my emotional WB—I am still the same-old anxious pessimist. I don't know whether a denial would have ruined my life but probably not. A survey (mentioned earlier) of academics a year after their tenure decision shows no significant difference in average happiness between those that did and did not get tenure.

I spent the better part of my young adulthood working night and day to get it because I thought I would be devastated if I were denied tenure—was it all a big mistake? Assistant professors correctly predict how they will feel when they learn they didn't get tenure—that emotional and life-satisfaction WB plunges. But are incorrect about how long their WB will be diminished. I am unsure whether getting tenure increased my long-term life-satisfaction WB, but I would like to think it did.

We believe that if we become disabled or are diagnosed with a long-term illness, our WB will decrease, so we get physicals and avoid bacon. But bacon tastes great. But are we trying too much if our beliefs about the extent of the decrease in WB are incorrect? Quoting from a University of Michigan press release about a 2005 paper by [Jason Riis](#) and others:

A Study of dialysis patients yields surprising findings. Despite what able-bodied healthy people might think, people with severe illnesses and disabilities don't wallow in misery and self-pity all the time. In fact, a new study finds, such patients, on the whole, may be just as happy as those without major medical conditions. The finding adds to the growing body of evidence that ill and disabled people adapt to their condition and show a resilience of spirit that many healthy people can't imagine.

This isn't to say people prefer being sick or that you will return to your state of WB before the illness.⁵ While adaptation is typical, it can be incomplete, even after years. E.g., Oswald and Powdthavee present evidence of partial adaptation to a disability three years after onset.

While most of the research on duration bias has been for private goods, duration bias also applies to public goods and bads. As President of the United States, Barak Obama was a public commodity. Once he became the President of the U.S. for me, he became that for everyone, including French people and Afghan terrorists. We were all forced to consume the results of his actions; for some, the Obama Presidency was a public good, and for others, a public bad. Before he was first elected President, many people thought his election would bring them tremendous and long-lasting joy; others expected tremendous and long-lasting misery. Both groups were mostly correct in the short run. But now that Obama's Presidency is long over, we realize we weren't as affected by it as we imagined we would be—the good and bad did not last as long as we imagined they would. The same might be true of our love, or hatred, of Trump and Biden

Reducing the rate of global warming is, for most, a public good (if it does not get as hot for me, it does not get as hot for you). However, if we underestimate our ability to mentally adjust to global warming, we could allocate too many resources to reducing it.⁶ A trite example: as an avid skier, I now dread a future with little snow. But as temperatures increase, I will find myself biking more and skiing less—my interest in skiing will decay, but now I can't imagine it. God forbid, I might even start playing golf—again.

Less trite, global warming is, and will, disrupt cultivation zones causing large-scale poverty and migration. E.g., coffee will no longer come from Central America; coffee will cost more, and less will be drunk. As a result, poverty in Central America will increase, causing even more northward migration pressure. Those affected will suffer, many catastrophically, but they won't suffer as much as they imagine if they suffer from duration bias.

I will get a flash of pleasure when I hear the news that wolves have been reintroduced into Colorado, my adopted State.⁷ But after reading the literature on my psychological immune

⁵ Duration bias does not imply you can't shift to a new quasi-fixed WB.

⁶ That said, we are currently underallocating resources to fight global warming. My point is that the how much should be allocated should depend on how much our WB will be affected in the longrun, not on an overestimate.

⁷ In November 2020, Colorado passed a statewide amendment mandating the reintroduction of wolves. This, by itself, gave me a flash of pleasure, but it waned.

system and my tendency to suffer from duration bias, I now suspect my happiness would adjust to the news in a matter of hours. Liberal environmentalists over-predict the long-term pleasure produced by Obamas and wolves, and conservatives and ranchers over-predict the duration of their displeasure. **NYT article about Montana.**

I almost made a mistake. A few years ago, the State of Colorado passed a law that expanded where you could carry a concealed weapon. Based on the law, in 2012, the [Colorado Supreme Court](#) struck down the University of Colorado's campus ban on guns. This ruling meant that any of my 500 Principles students could start packing. I got stressed (an emotional state), worrying that I could get shot if I failed or offended the wrong student. Retiring crossed my mind. If I had retired because I was worried about guns on campus, it would have been flawed choosing. I got used to the idea of guns in class and stopped worrying about it. Put simply, I misinterpreted how long my anxiety would last and came close to making a flawed choice. [Anxious people typically over-predict how long a bout of anxiety will last, especially those who medicate as soon as they become anxious. By circumventing the anxiety with a drug, they never experience the anxiety naturally subsiding, so they do not realize it will subside.]

Duration bias should be suspected whenever a choice involves long-term changes. For example, with environmental policies, there are typically long-run implications. Consider the extinction of a species or the PCB contamination of a water body. When an environmental injury occurs, people get upset (are damaged), but duration bias implies that even if the injury remains, the upset will not last as long as they imagine. This raises questions about estimating how much society was damaged, in dollars, by an oil spill—one of my sidelines for many years. Suppose that soon after the injury, while you were bummed, you were asked whether you would be willing to have your taxes increased by \$25 a year to ensure there wouldn't be another injury of the same type and magnitude (extinction, spill, pollution event). Suppose further that you said yes because (1) you felt it is worth \$25 to not experience, for a year, the bad feeling you are experiencing, and (2) you believe the bad feeling you were experiencing would persist year after year. If the bad-feeling duration turns out to be shorter than you are predicting, the value you place on prevention is less than your answer indicates.⁸

⁸ Some economists argue that answers to hypothetical questions about willingness-to-pay are biased upwards because you answer knowing you will not have to pay. Whether this is true or false is a different from people having an actual WTP that is more than it would be if they did not suffer from duration bias.

So, why do we suffer from duration bias?

Is there an advantage? In his 2002 book, *Strangers to ourselves: discovering the adaptive unconscious*, Timothy Wilson, a social psychologist at the U. of Virginia, discusses reasons why this quirk is common (this is his list with my explanations):

- *Focalism*: When something great happens, or the shit hits the fan, it captures all our attention (conscious and unconscious)—it is all we can think about—we focus—the happening has salience. When we think about how it will make us feel in the future, we don't realize that it will no longer be the be-all-and-end-all, so we overestimate its long-run impact on our WB. It is good that our focus wains: we would get in big trouble if our attention was wholly occupied by past events: survival requires paying attention to the present. If B.B. Wolf is at the door and you're still cranked about something that happened last year, you might get eaten.
- We fail to adequately foresee how we can and will adjust. Economics preaches that when a negative occurs (price increase, loss of job, global warming), we, over time, minimize the cost to us (substitute to cheaper substitute, find a new job, etc.) Duration bias occurs because we don't realize how much we can and will adjust to the new reality. Two friends recently lost their older dog. They were sad, maybe overly so, because they did not foresee her death bringing them closer to the other dog. When your house floods, you, at first, fail to realize that the hated purple-paisley drapes are history, and you no longer need to fret about what basement junk to keep
- There is the afore-mentioned *quasi-fixed WB* that we forget we have (Chapter 5).
- *Misconstrual* (perception errors): We don't know how we will react to a change because it is difficult to predict the WB impact of events we haven't previously experienced. And, even if the experience isn't new, we still have trouble predicting how it will affect our WB: memories of past emotions, distress, and pain are often flawed. When we remember pain, we don't re-experience the pain, so mothers forget how much childbirth hurt. And marathon runners forget how they felt at the eighteen-mile point, making it easy to make duration errors.
- Overestimating how long the happiness will last increases current happiness and contentment. Or, you overestimate the duration of your unhappiness because doing so makes you feel better prepared. This is *motivated distortion*: thinking your stomach pain is cancer makes you think that you will be better able to handle the diagnosis when it comes. Or you convince yourself you did poorly on the final exam, so you will be better prepared for the long-term WB drop when you get your grade.
- *Spin*: Failure to realize how adept our minds are at spinning things to our emotional advantage. We spin/interpret things to uncover the silver lining: when the shit hits the fan, we figure out a way to spin it so it does not seem as shitty. Duration bias occurs partly because we fail to realize that we spin. When my students fail an exam, they don't feel as bad after they spin to blaming someone else, me. But initially, they don't realize they will blame me.

There are a lot of unanswered questions about the duration of WB shifts. We adapt to different things at different rates, and some events are difficult to adapt to. There is only limited research on how long various WB shifts last as a function of the event that caused the shift, which emotions were affected (love, fear, hate, anxiety), and the characteristics of the individual affected (age, gender, personality type).⁹ To estimate the extent of the bias, one first needs the best estimation of the shift's duration based on the best science and data. Then there is the issue of how WB is defined: a shift in emotional WB or a shift in life-satisfaction WB: the durations might differ.¹⁰

I am unaware of research on whether the degree of bias varies with the actual length of the WB shift—it is complicated. I can't imagine that the bias is always a fixed length of time (e.g., one month or one year) or a constant multiple of the actual length, but I don't know this. Other unanswered questions include the length of the duration bias for emotional WB vs. life-satisfaction WB. And, if emotional WB, does the bias depend on the specific emotion (e.g., fear vs. anger)?

Duration bias shows that WB forecasts are based on common and systematic incorrect beliefs. Choices with long-term implications are made based on our perception of how they will make us feel in the long run, but with duration bias, beliefs are systematically biased. One can still use NBT to model behavior (as one can with any incorrect belief); the challenge is to model

⁹ Di Tella, Haisken-DeNew, and MacCulloch (2010) find that duration of a WB shift resulting from an income shift is less than the duration of a WB shift caused by a change in status. Kimball, Nunn, and Silverman (2015) estimate that for an adult the WB drop associated with the death of a parent lasts, on average, 6 months; whereas the WB drop associated with the loss of a child lasts five years. See Figure 1, Chapter 4, the section on a happiness set point)

Wilson and Gilbert (2008) survey the literature on what we do, and don't, adapt to. One study finds that "people adapt more easily to incarceration than to unpleasant noises."—Really! We don't adapt to unpredictable noise (consider traffic noises and overheard cell-phone conversations) but do adapt to clock ticking. And, as noted in Chapters 5, while we adapt to consuming more and better things, the happiness lift from being rich relative to others is, arguably, more enduring. Losing your job "is particularly difficult to adapt to and has a long-term impact on life satisfaction."

It seems that the WB loss associated with commuting endures. "You can't adapt to commuting because it's entirely unpredictable. Driving in traffic is a different kind of hell every day." (*Daniel Gilbert as quoted in [Traffic: Why we drive the way we do](#) by Tom Vanderbilt*). The happiness researchers Bruno Frey and Alois Stutzer have identified the *paradox of commuting*: you go for the bigger house and yard in exchange for a longer commute, not realizing that you will adapt to the house and yard, but not to the commute. Put simply, if you commute your house is too big.

¹⁰ Kahneman and Deaton (2010) find that an income has an enduring effect on life-satisfaction WB but not emotional WB.

it and the extent to which it affects behavior—including how much behavior is based on this incorrect belief.

Maybe you and I can increase our WB simply by being aware that we suffer from duration bias. When a choice has long-run implications, consider the possibility that you are overestimating the durations of the different WB effects. I learned about duration bias when guns were first allowed in my classroom.

The emotional empathy-gapⁱⁱⁱ

Think about jealousy and revenge as possible emotions to discuss. When we select a path, we do not fully appreciate that our selection is based on our current emotional state and the ordering associated with that emotional state. We are limited in our ability to empathize with our other emotional selves. This common quirk is what [George Loewenstein](#), the Herbert A. Simon, Professor of Economics and Psychology at Carnegie Mellon University, calls the *empathy gap*. People have emotion-specific orderings. My ordering of paths depends on my emotional/chemical state. The emotional empathy-gap is that we are limited in our ability to realize this, so do things now, when we are in one emotional state, that we will have to live after that emotion subsides and is replaced with another one. You cut me off in traffic, my anger flairs, and I shoot you with the Glock I keep under the seat. My anger is quickly replaced by fear, anxiety, and guilt since I am now the deserving subject of the all-points bulletin, “Armed and dangerous old white guy driving south on I-25 in a red Buick.” My empathy gap is that while livid, I chose my highest-ranked available path, HRAP, incorrectly, imagining that I would have the same ordering tomorrow.

Separate in your mind the issue of whether you have multiple ordering from whether you fail to empathize with your other emotional selves, thinking about what each implies for the Assumptions of NBT.

Multiple orderings?^{iv}

A hot temper leaps o'er a cold decree. - The Merchant of Venice, CCT 1, scene 2

Behavioral research, primarily by psychologists, demonstrates that we have different orderings of paths for different emotional states—you are not of one mind and don't have one single ordering. In summarizing, in my words, not theirs, which ordering you use depends on your emotional state of mind. Assuming S possible states, in *state-of-mind s*, you use the ordering

coughed-up for State s . These S different orderings determine choice in that in State s , you will select the highest s -ranked path from those available. For example, Jordan, sexually aroused, does not have the same ordering as the Jordan who wakes up with a stranger; so, the selection that increased Saturday-night Jordan's WB might make Sunday-morning Jordan miserable or even happier.

*As she melted small and wonderful in his arms, she became infinitely desirable to him; all his blood vessels seemed to scald with intense yet tender desire, for her; for her softness, for the penetrating beauty of her in his arms, passing into his blood. And softly, with that marvelous swoon-like caress of his hand in pure soft desire, softly he stroked the silky slope of her loins... And she felt him like a flame of desire (*Lady Chatterley's Lover*, D.H. Lawrence, quoted and discussed by Robert Sapolsky in *Why Zebras Don't Get Ulcers*)*

Depending on your literary and erotic tastes, this quote may be causing your pancreas to secrete hormones, changing your visceral state. Economists don't talk or think like D.H. Lawrence. While emotions affect choice, emotions have been purged from NBT and welfare economics. Economists should be pitied: ordinary people have passions, lusts, and rages, while modern economists have only an ordering of paths.

There are states induced by psychoactive drugs such as tobacco, alcohol, and caffeine—most of us start our day by altering our mind with caffeine, blocking the receptors for the neurotransmitter adenosine.¹¹ Even the bacteria in your gut affect your emotional state.¹² Ordering-altering emotions and moods include anxious, depressed, scared, exhausted, sleep-deprived, bored, sad, grieving, curious, joyful, feeling alienated, embarrassed, and my favorite mood, grumpy.¹³ We experience these states in ourselves and others, affecting how we behave and experience the world. But we are not always aware of the state we are in.

¹¹ Caffeine also affects dopamine, acetylcholine, and serotonin.

¹² Your body contains six to ten pounds of bacteria, much of it in your gut. They produce dozens of chemicals, including many that in the brain act as neurotransmitters (e.g., dopamine, serotonin, and GABA). Recent research indicates that their production in the gut stimulate nerves in the gut which synapse with neurons in the brain, influencing emotional state (see [Peter Smith](#) 2015 and the research of [Mark Lyteo](#)). In a path-breaking study ([Javier Bravo](#) et al. 2011), mice were divided into two groups: a control group and a group that were fed a broth containing the common bacteria *Lactobacillus*. Human babies ingest it as they pass through birth canal. It, in the gut, produces GABA, and GABA inhibits nervous activity. Both sets of mice were subject to the same stressful environment (a forced swim where the rat can neither touch bottom nor climb out). The mice fed the broth exhibited less stress. For more on bacteria and emotions see the 2014 TEDMED talk by [John Cryan](#).

¹³ Examples include Loewenstein (1996, 99, 2000), Loewenstein, [Prelec](#), and Shatto (1996), Loewenstein, Nagin, and Paternoster (1997), Lerner, Small, and Loewenstein (2004), Ariely and Loewenstein (2006), Kosfeld et al. (2005), Ditto et al. (2006), Lerner, Li, and [Weber](#) (2012), and [Garg](#) and Lerner (2013). For a neurological discussion of the effects of emotions on choice see [Damasio](#) (1994). See also [Elster](#) (1998).

Loewenstein and coauthors provide examples of how choice differs across emotional states; for example, the turn-back-time for a Himalayan mountain ascent chosen in the planning phase is ignored during the stress and exhaustion of the climb. The result can be tomorrow's self is dead, the ultimate *negative external effect*.

Sexual arousal, induced by erotic pictures, films, and self-stimulation, changes your propensity to take risks, be aggressive on a date, and your sense of appropriate behavior.

Loewenstein, [Daniel Nagin](#), and [Raymond Paternoster](#) had

Sexually aroused and non-aroused males predict their own behavior in a date-rape scenario. Aroused and non-aroused participants were asked a battery of questions designed to measure their perceptions of the costs and benefits of acting in a sexually aggressive manner, their level of arousal, and a probabilistic prediction as to how aggressively they would act in the conditions described in the scenario... Sexual arousal increases the "subjects' expectations of their own sexual aggressiveness and that this impact isn't mediated by perceptions of the costs or benefits of such aggression.

[Peter Ditto](#) and colleagues have researched the effects on choice of arousal. Males not in a relationship and who typically use condoms, either watched a video or read a description of the following situation.

The written text: *Imagine that you are single and that you run into an attractive acquaintance, Rebecca, at a bar. You continue to spend time together throughout the night and end up at her apartment at the end of the evening. After talking and kissing on the couch, it is clear that you are both very interested in having sex with each other. However, neither of you has a condom. You discuss the possibility of going to a store, but there isn't one nearby. She tells you that she is on the pill and has not slept with anyone since her last boyfriend.*

The video: *In the video, two attractive students, Pierre and Rebecca, talk after class and arrange to meet at a bar later that night. After an evening of dancing, they go to Rebecca's apartment where they kiss on her couch. Eventually, Rebecca takes off her shirt and indicates that she is interested in having sex (alleviating the threat of a potential date-rape situation). Pierre then discloses that he did not bring any condoms. Rebecca tells him that she also does not have any condoms but is on the pill (alleviating the threat of pregnancy). The two discuss the idea of obtaining condoms from a local store but conclude that it isn't available. They discuss their sexual history (e.g., Pierre states that he is "clean," and Rebecca states that she does not "sleep around"). The video ends with Pierre asking Rebecca, "What do you want to do?" and Rebecca replying, "I don't know. What do you want to do?"*

The video watchers were more inclined to say they would take the risk and have unprotected sex.

Switching to other emotions, when depressed, you order paths differently, including paths with suicide. Invoking curiosity changes behavior—it killed the cat. When curious, you are more inclined toward a reward now v. later. [Curiosity can be invoked by asking an interesting question to which the respondent does not know the answer.]

Sadness does not have the effect you might expect. Some negative emotions (e.g., disgust) cause you to value everything less. While sadness causes us to value what we have less, it causes us to value new things more. In a 2004 experiment by Loewenstein and [Deborah Small](#) (Professor of Marketing and Psychology at the U. of Penn.), disgust was invoked with a video clip from the movie [Trainspotting](#) and sadness with one from [The Champ](#). They used a neutral clip from *National Geographic* as the control. Subjects were divided into two groups: potential buyers and potential sellers. The potential sellers were given a set of highlighters. Potential sellers subject to the disgust and sad treatments were willing to sell them for less than those in the control group. Potential buyers in the disgust treatment were willing to pay less than those in the control group, but potential buyers in the sad treatment were willing to pay more. [Jennifer Lerner](#) (Director, Harvard Laboratory for Decision Sciences) interprets this to mean that when you are sad, you will pay more to get new stuff, hoping it will perk you up.

In a 2012 Lerner study, subjects were offered either \$85 paid in three months or a lesser amount immediately. The median respondent in the control group needed \$56 now to pass up the \$85, but the median for the sad respondents was only \$37: sadness invokes impatience. However, when both payments were in the future, the difference between the sad treatment and the control group disappeared, suggesting sadness causes us to want the reward now, not simply sooner. [Lerner has also found that subjects who get the sad treatment eat more free M&Ms.]

Boredom makes most other situations look good—my students long to be anywhere else. Boredom turns teenage boys into vandals—it provokes wild-and-crazy behavior. For a review of philosophical takes on boredom, see *A Philosophy of Boredom* by the Norwegian philosopher Lars Svendsen. For extreme reactions to boredom, he considers the 1996 [David Cronenberg movie, Crash](#) and the book and movie [American Psycho](#) (nothing un-bored Patrick Bateman, a rich 80's investment banker and party boy, not even torture and murder).

Market transactions require trust. The team of [Michael Kosfeld](#), [Paul Zax](#), [Urs Fischbacher](#), [Ernst Fehr](#) (economists), and [Markus Heinrichs](#) (a psychologist) have shown that nasally administering the neurotransmitter oxytocin increases the probability you will enter into a financial contract requiring trust—it makes us more trusting. (For mammals, oxytocin helps mom and baby bond, also dogs and their owners).

Happiness is an emotional state and, like other emotional states, affects your ordering of paths.¹⁴ This is crazy-making for proponents of NBT: you choose the available path that will bring you the most WB but being happy (high emotional WB) affects the path you choose, a confounding circularity.

An aside: multiple orderings share many properties with dissociative personality disorders (Dr. Jekyll and Mr. Hyde and the actress Joanne Woodward in the movie, [The Three Faces of Eve](#)).¹⁵

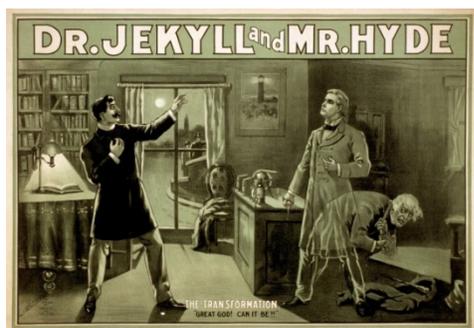


Figure 2: Chapter 6

Consider: as Edwina, I wear dresses, heels, and bar flirt. When Edward, I sit in my basement office doing research, unable to walk in heels. Edward would be shocked to learn Edwina hit the bars last night. My third persona, Edalandra the ballerina, finds Edward boring and hates that Edwina is always borrowing and ruining her tutus. Each personality does things that affect the other two (how could they not, they live in the same body), but none consider these external effects when deciding how to live their lives. Edward and Edalandra hate cigarettes and cigarette smoke, but Edwina has smoked Camels for years—unfortunately, when Edwina dies of lung cancer, Edward and Edalandra will pass as well. People with dissociative personality disorder have multiple orderings; they consist of more than one self. The same holds for all of us living with varying emotions.

The emotional empathy-gap:

You suffer from an empathy gap (little empathy for your other emotional selves) if you select your HRAP based on the WB ordering associated with your current emotional state. And do this at each point in time. While multiple orderings are necessary for the empathy gap, it isn't sufficient. You also need to be unaware that your ordering will be different after you morph into

¹⁴ See Mogilner, Aaker, and Kamvar (2016) and Labroo and Patrick (2009); the latter find that happiness affects choice because when one is happy one's perspective is broader, more future-looking, and more analytical.

¹⁵ DPP was previously called a multiple-personality disorder. While the disorder fascinates, it isn't clear to what extent it exists: one hypothesis is it's a product of suggestions by the therapist.

another emotional state. [Or, aware that it will change but are unable to act contrary to your current ordering.]

Why is there an emotional empathy-gap?

It isn't because we have no experience with different states of mind; we have all been angry, scared, hungry, tired, sexually aroused, curious, and bored. Is there an evolutionary reason we do not empathize with our other selves, or is it simply a kluge?

Loewenstein and [David Schkade](#) point out that emotional memories are “qualitatively different” from other types of memories, making it difficult to project how we will feel and assess in a different emotional state. We are adept at recovering visual images, also the sum of 2+2, but not adept at recovering feelings. When you picture mom, you experience/see an image of your mother; when you remember a song, you hear the song. But when you remember being mad, you don't experience/feel mad, and when you remember a time when you were hungry, you don't feel hungry. The forward reflection of this inability is that it is difficult to correctly comprehend the emotional salience of future hunger unless you are hungry. [Remembering the circumstances that made you mad can make you mad, but that's different.] When you remember being in a visceral state, you recover the circumstances but don't experience the emotion. It is as if you store the events around the emotion, but not the emotion, so don't experience the compulsions it would engender. When I decide in the morning to go to the grocery store after work, I am aware that I will be hungry when I get to the store. But I cannot experience the buying compulsions that that hunger will engender. When my neighbor puts a gun in her glove compartment to feel safer during her night commute in rural Colorado, she can't appreciate the compulsion that her anger will engender when another driver cuts her off and then flips her off. We underestimate because we can't retrieve the compulsive feeling from memory.

An inability to recover old feelings has its evolutionary advantages—it is evolutionary rational. For example, if remembering a good time recovered the good feeling (made you feel great), you could feel great simply by lying around and remembering pleasurable times. But it also would increase the probability of you becoming lion food. And it would decrease your incentive to eat and have sex; you could feel satiated by simply remembering past liaisons and meals. Maybe the inability to experience now a future emotion is merely the flip side of our

inability to recover old feelings, so while not necessarily evolutionary adaptive, it might be a side effect of a trait that is.

The gap might also exist because empathizing with a different self diminishes your current sense of self and the legitimacy of your current feelings.

What does the emotional empathy-gap imply about behavior and choice?

Are the Assumptions of NBT violated? First, recollect that the gap rests on two components: having emotion-specific orderings and behaving as if you don't.

Assumption 2: At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, ...

Emotion-specific orderings and varying emotions will lead to behaviors that seem to violate Assumption 2: at one moment, you rank Path *m* above Path *k*; later in the day, the ranking is flipped.¹⁶ But, emotion-specific orderings and varying emotions don't strictly violate Assumption 2. They are consistent with, at each moment, a unique, effective ordering: there are multiple orderings in your brain, but at this moment, only one is in charge.

Consider *Assumption 4: An economicus's ordering does not change in the time-span behavior is modeled.* It is violated if emotions vary in the period for which behavior is modeled. Since most people have varying emotions, this is problematic.

Like duration bias, the emotional empathy-gap does not violate

Assumption 9a: An economicus's ordering of paths is based on its WB (well-being), WB, accounting for uncertainties. The more WB economicus associates with a path, the higher its rank. And WB is determined, in part, by sensations.

Because higher-ranked paths are better given the incorrect belief that one's ordering will not change. Besides 4, no Assumptions are violated.

The emotional empathy-gap is having emotion-specific orderings and, at that moment, the incorrect belief that you don't. This leads to flawed choosing. Given our state of mind, we won't momentarily reduce our WB even though it would increase our overall WB. When [Homer Simpson](#) is depressed and can't imagine the depression lifting (a characteristic of depression), his

¹⁶ Early in the day you are angry so rank path *m* higher than path *k*, but later when your anger turns to anxiety, path *k* is ranked higher than path *m*.

WB is maximized by not going to work, not socializing, and spending night and day on the couch watching Simpson reruns, given his beliefs. But, if he knew that on anti-depressants, he would be able to empathize with his future self, he would have instead stopped at the pharmacy, picked up an SSRI, and continued on his way to work.

The emotional empathy-gap, like duration bias, is an assault on the common economic assertion that we typically choose, given our constraints, our highest-ranked WB-path in terms of realized WB.¹⁷ But, strictly speaking, it is consistent with the Assumptions of NBT.

How might knowing you have an emotional empathy-gap make you better off?

First, understand your natural inclination to not empathize with your other emotional selves. Try to empathize with them. Learn to recognize and acknowledge your emotions and what triggers them. And finally, think about how you can influence the occurrence of different emotional states. All are tall tasks. Simply understanding that you have an empathy gap is the place to start.

If you can speculate on how your ordering of paths might differ across emotional states, you might choose now not to keep a loaded gun under your bed.¹⁸ And, you might take an Uber to the bar rather than driving—and not bring the gun. How should I react when a guy at the bar successfully hits on my wife? If I can empathize with my future selves, I am less likely to start shooting or hitting the guy—or my wife. Most males regret hitting people, but an estimated [22% of women](#) have been physically assaulted by an intimate partner.

The emotional empathy-gap suggests that if an objective is to train an individual to act, or not act, in a certain way when they are mad, scared, or aroused, the training will be more effective if it occurs when they are in that state. Train soldiers and police officers when they are afraid and angry how to avoid shooting civilians when they are afraid and angry. Training males to practice safe, consensual sex when they are sexually aroused will be more effective if the training occurs when they are aroused.

¹⁷ The emotional empathy-gap will not generate flawed choosing if the individual cares not at all about the future: if you live in the moment, you care nothing about your future WB. If I shoot you in anger after you insult my dog, it's not flawed choosing if I don't care about the future. Most of us discount the future but not fully, so the emotional empathy-gap typically leads to flawed choosing.

¹⁸ [Oscar Pistorius](#) (the double amputee, “Blade Runner”), shot his girlfriend four times through the bathroom door, either in anger, or, as his lawyer argued, because he was in a vulnerable and fearful state. He is in prison until at least 2023.

Complicating even further: your assessment of your emotions can be incorrect, an incorrect belief^v

The psychologists [Don Dutton](#) and Arthur Aron had an attractive female approach males in a Vancouver, B.C. Park and ask them to fill out a short survey about the park. At the survey's end, she provided the respondents with her phone number "in case they had further comments or questions". Half the males were approached on a scary foot-path suspension bridge over a deep ravine, half after they had crossed the bridge. Dutton and Aron speculated that a few subjects would call and ask her out. They wanted to determine if the ask-rate varied by where she interviewed them. She got many date requests from those interviewed on the bridge. Their propensity to call was greater because they mistook fear for arousal, an incorrect belief, resulting in flawed choosing. Even an unconscious emotional state can affect behavior.¹⁹

Incentive-salience effects^{vi}

Recollect the incentive-salience hypothesis of Chapter 4. This common quirk is most associated with the neuroscientist Kent Berridge and his lab at the U. of Michigan.²⁰ Summarizing the discussion of *Incentive salience* in Chapter 4, it is a fancy term for serious wanting (a strong urge—conscious or not) compared to run-of-the-mill wanting. The *incentive* is to choose the most *salient* (striking, attention-grabbing). Incentive salience assumes you choose the alternative you most want/desire. Often, it is the alternative you will like the most (generate the most WB). But wanting can deviate from liking. This happens when you have a physiological or emotional need (you are hungry, aroused, tired, etc.), more than one alternative fulfills that need, and one alternative is cued. You are thirsty and notice an attractive person enjoying a Pepsi (a cue), so you choose Pepsi over Coke even though you would have enjoyed the Coke more. The root of this common quirk is evolution's separation of wanting from liking.

¹⁹ In a lovely study, the psychologists [Lawrence Williams](#) and [John Bargh](#) had each subject met in the lobby by a greeter who accompanied them to the lab. The greeter was juggling a pile of papers and a drink. She asked the subject to hold the drink; sometimes it was hot, sometimes cold. When they got to the lab, the greeter left, taking her drink. In the study interview, the researcher described a person and then the subject was asked to describe that person in their own words. The subjects who carried a hot drink were more likely to judge the person as having a warm personality. (I always carry around a Diet Coke with lots of ice; maybe this explains why people find me cold and unapproachable.)

²⁰ See Berridge and [Robinson](#) (1998), Wyvell and Berridge (2000 and 2001), Kringelbach and Berridge (2012), and Berridge and Kringelbach (2013), also the textbooks: Bear, Connors, and Paradiso (2007) and Gazziniga, Ivry, and Mangun (2014).

So, how should one think about incentive-salience in terms of NBT? It is inconsistent with [Assumption 9a](#) but consistent with [Assumption 9b: *Paths that are more desired/wanted are ranked higher.*](#) Physiological states are like emotional states. While incentive-salience affects behavior, they don't produce incorrect beliefs—you really desire the item more. The incentive-salience quirk can momentarily change an ordering of paths based on desires/wants, so they can violate [Assumption 4: *An economicus's ordering does not change in the time-span behavior is modeled.*](#) They make it challenging to model and predict behaviors: behavior depends on one's need state, along with any cues. If the objective is desire fulfillment, the incentive-salience quirk does not cause flawed choosing—you still end up experiencing the most desired available path. Incentive salience does cause flawed WB-choosing, but *So What?* if NBT adopts Assumption 9b but not Assumption 9a.

Imagine a top-level ordering unaffected by cues and a ground-level ordering that is. Your ground-level ordering deviates from your top-level ordering when you are in a need state, and one of the alternatives is cued. If the 9a and 9b ordering are identical when there are no incentive-salience effects, their presence will cause them to diverge, and there will likely be flawed WB-choosing.

The future-self empathy-gap^{vii}

The *future-self empathy-gap* is a lack of empathy for one's future selves. It violates no NBT Assumptions. Does it cause flawed choosing? The word *gap* suggests maybe.

Animals discount the future, including humans.²¹ When ordering paths, everything else constant, humans order paths lower (higher) the farther in the future each WB bump (dip) occurs.

²¹ Non-primates will only wait a few seconds for a larger reward: birds and rodents will wait a few seconds more for a reward three-times larger than the immediate alternative, but if the wait is around 10 seconds, they go with the immediate but smaller reward (Santos and Rosati 2015). Non-human primates are more patient, but not as much as humans. Lesser primates will wait longer than birds and rodents, but still less than a minute. However, great apes will wait minutes for a larger reward. Macaques become indifferent between the immediate smaller reward and the three-times larger reward at approximately 40 seconds (Tobin et al. 1996). Tamarins and marmosets (new-world monkeys) hold out for 10-20 seconds (Stevens, Hallinan, and Hauser 2005). [Michael Beran](#) and his colleagues (Beran, [Savage-Rumbaugh](#), Pate, and [Rumbaugh](#) 1999, Beran 2002, Beran and [Evans](#) 2006) conducted experiments with chimps where the size of the reward visually grew over time if the chimp did not touch the reward. The chimps would hold off for about 10 minutes, about the same as a small child, using the same sort of distracting tactics as a small child (e.g., not looking at the reward and playing with toys as a distraction (Evans and Beran 2007)). Quoting Santos and Rosati, "Further converging evidence that apes are capable of thinking about their future selves comes from planning studies in which apes must anticipate that saving a tool now will allow them to use it in the future. Indeed, some apes successfully plan to use a tool as long as 14 hours in advance ([Mucalhy](#) and [Call](#) 2006; see also

I will rank a path higher if its bump comes earlier. And if three paths only differ by when the same WB dip occurs, I will rank a path higher if its dip is farther in the future. Put simply, humans would rather experience the good stuff sooner and the bad stuff later and order paths accordingly. Researchers and policy makers who consider discounting a human flaw describe discounting as *impatience*.

There isn't anything about discounting the future inherently inconsistent with NBT. And—there are many good reasons we discount the future; these include the future is uncertain, everyone has a limited life span, and most animals lack the cognitive ability to imagine their future selves. In our world of scarcity, predators, and prey, discounting the future increases the probability that your genes will survive until tomorrow. Worrying about old age when one is young is a trait unique to humans. Evolutionary biology also suggests that, within a species, there will be genetic variation in the degree of discounting.

The degree to which a human discounts the future depends on how much she empathizes with her future selves (cares about them). And this depends on the extent to which you will remain the same person (identity as the same person). Your identity depends on *connectedness*, and your empathy for your future self increases with your connectedness to them (Parfit 1984, Bartels and Urminsky 2011, and Tappolet 2010).²² My current choices and behavior will affect

Osvath and Osvath 2008). Overall, these results suggest that humans and other great apes may share similar capacities to overcome immediate temptation and to act in ways that benefit their future selves as well as similar strategies for doing so.” That animals lack patience should not surprise; most animals have no conception of their future selves. What is remarkable is that some great apes will wait longer than many small children and seem to have longer planning horizons.

²² This issue of what makes a person the same person they were yesterday can quickly become speculative, but you are free to suppose a person is the thing that inhabits the same body over time. But this *materialistic* notion of identity does not solve the problem of how you should behave if your body's ordering of paths changes unexpectedly.

Two bookends in the evolution of identity definitions are John Locke and Derek Parfit. Paraphrasing Gottlieb (2016), what makes someone the same person over time was a hot topic amongst English pastors in Locke's time: on Judgement Day, what part(s) of you will God need to resurrect you? Body parts? Not a problem if you are still buried in the parish cemetery, but what if you were burned at the stake or eaten by cannibals?

Locke concluded that if you now remember doing something, you are now the same person who did it—implying that if I don't remember that last night the drunken me, robbed a 7-11, I am not that guy. Locke's definition implies that the teenage you and young-adult you are the same person only if the latter remembers high school. And the young-adult you and the old you are the same person only if the latter remembers the behavior of the former. But even if both are true, the teenage you and the old you are not the same person if the old you does not remember high school.

Parfit gets around this, arguing that what is important for survival and responsibility is psychological connectedness: remaining the same person is unnecessary. Person *X* at time t_1 is psychologically continuous with person *Y* at t_2 if and only if *X* is psychologically continuous with *Y*, “where psychological continuity consists in overlapping chains

my future selves, generating external effects on other people (future me-s) whose ordering could differ from mine in ways I can't predict.

Daniel Bartels and Oleg Urminsky, both marketing professors at the U. of Chicago, hypothesize that a reason we discount the future is a disconnect between our current and future selves. I don't save for old age, partly because I don't comprehend or relate to the old guy who will be spending my money, my *future-self empathy-gap*.²³ Quoting them,

We will argue that our understanding of what constitutes a 'reasonable' discount rate ... has been limited by the implicit assumption that people should maximize the utility of a constant self over one's lifetime. An alternative position, proposed by the philosopher Derek Parfit (1984), is that a decision about consuming now or later should depend not only on the temporal distance between events but also on the perceived continuity between one's present and future selves. This implies that the degree of concern one has for one's future self should be scaled by the degree of 'psychological connectedness'—overlap in personality, temperament, major likes and dislikes, beliefs, values, ambitions, life goals, ideals, and so on—held between one's current and future self... Economists assume a constant self.

In five separate experiments, Bartels and Urminsky get results consistent with decreased connectedness causing increased impatience. [In the lab, connectedness is manipulated by citing evidence that significant future events (e.g., graduation, marriage) will likely change the subject's core identity. Or citing evidence that core identity is fixed early in life.)

In *Procrastination and Personal Identity*, the U. of Montreal philosopher and ethicist Christine Tappolet attributes procrastination to an inability to connect with your future selves. [In Latin, *pro* means “for” and *crastinus* means “of the morrow”]. I put off what I should be doing because I don't connect and emphasize with the person who will suffer from my procrastinating. She speculates that this is why the Russian nobleman [Itia Ilitch Oblomov](#) (the main character in an [Ivan Goncharov](#) novel) stays in bed rather than doing what has to be done to keep his estate afloat.

Empathy for one's future selves, or its absence, is, conceptually, no different from empathy now, or not, for other human and non-human animals (caring now, or not, about

of strong psychological connections like memories, intentions, beliefs/goals/desires, and similarity characteristics.” ([Shoemaker](#) 2019).

As noted, Buddhists consider the notion of a persisting person a useful fiction, arguing that there is no me in the me (*no-self*). Parfit's conclusion is similar, a conclusion he has described as liberating and consoling.

A common view, not shared by philosophers but preached from pulpits, is your identity is your eternal soul. Then there is the “Beam-me-up Scotty” [transporter](#) where the you in you is converted into an energy pattern—destroying you. But you are resurrected on Vulcan when you are converted back. Sometimes Scotty screws up, and duplicates appear, problematic for personal identity.

²³ Imagine how you would act if you expected your future self to be a [Franz Kafka insect](#) or a [155-pound Phillip Roth breast](#).

others).²⁴ The assumptions of NBT are consistent with any degree of empathy for one's future self or others.

But—a lack of empathy for one's future selves causes flawed choosing if modern humans are programmed by biological or cultural evolution to have less future-self empathy than is warranted, given the recent and dramatic, jump in human life spans. We are quirked to not save for our sixties and beyond. Why? Because until the last few centuries, few lived that long. If you fail to emphasize with your future self because you believe, incorrectly, that you are unlikely to live that long, the gap is caused by an incorrect belief.²⁵ Believing now that you will not care about your future self when you are him is also an incorrect belief but of a different sort.

But not all life spans have jumped dramatically: most animals have not experienced a jump in longevity, and neither have the many millions of humans that live in violence and poverty. So, while the future-self empathy-gap causes rich Western me to suffer flawed choosing, it does not cause everyone to.

ⁱ Sources:

1. Coren Apicella, Eduardo Azevedo, Nicholas Christakis, and James Fowler (2014) Evolutionary Origins of the Endowment Effect: Evidence from Hunter-Gatherers, *Am. Econ. Rev.* 104(6): 1793-1905. <http://www.jstor.org/stable/42920866>
1. David Cesarini, Magnus Johannesson, Patrik Magnusson, and Bjorn Wallace (2012) The behavioral genetics of behavioral anomalies, *Manage. Sci.* 58(1): 21-34. <http://dx.doi.org/10.1287/mnsc.1110.1329>
2. Timothy Flemming, Owen Jones, Laura Mayo, Tara Stoinski and Sarah Brosnan (2012) The endowment effect in orangutans, *Intern. J. of Comp. Psych.* 25: 285-298 http://www.comparativepsychology.org/ijcp-vol25-4-2012/2.Flemming_etal_Final.pdf
3. Steffen Huck, Georg Kirchsteiger, and Jorg Oechssler (2005) Learning to Like What You Have- Explaining the Endowment Effect, *Econ. J.* 15(505): 689-702. <https://doi.org/10.1111/j.1468-0297.2005.01015.x>
4. Moshe Levy (2015) An evolutionary explanation for risk aversion, *J. of Econ. Psych.* 46: 51-61. <https://doi.org/10.1016/j.joep.2014.12.001>
5. Laurie Santos and Alexandra Rosati (2015) The evolutionary roots of human decision making, *Ann. Rev. of Psych.* 66:321-347 <https://doi.org/10.1146/annurev-psych-010814-015310>
6. Jeffrey Stevens, Elizabeth Hallinan and Mark Hauser (2005) The ecology and evolution of patience in two New World monkeys. *Biol. Lett.* 1: 223–26. [PubMed: 17148172] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1626214/>

²⁴ It is notable that economists do not balk at modeling what is WB-best, overall, for a sequence of individuals who all inhabit the same, but changing, body, but balk at modeling what is WB-best, overall, for a group of individuals who inhabit separate bodies. It is also notable that many psychologists find lack of empathy for one's future self a flaw but not lack of empathy for others.

²⁵ Imagine that you believe you will die soon, and this belief is based on ignorance and a misreading of the longevity statistics for people like you. So, you blow the retirement account driving your new Mercedes to Grateful Dead reunions. You live longer than you expected and end up eating cat food in your twilight years, cheap cat food. You would have been better off in your old age if you had forgone the car and concerts, but you did what was best, given your knowledge and ignorance at the time. But it was flawed choosing.

ii Sources:

1. Dan Gilbert, Elizabeth Pinel, Timothy Wilson and Thalia Wheatley (1998) Immune Neglect: A Source of Durability Bias in Affective Forecasting, *J. of Person. and Soc. Psych.* 75(3): 617-38.
2. Miles Kimball, Ryan Nunn and Dan Silverman (2015) Accounting for adaptation in the economics of happiness. NBER working paper series, National Bureau of Economic Research, Cambridge MA. <http://www.nber.org/papers/w21365.pdf>
3. George Loewenstein and Shane Frederick (1997) Predicting Reactions to Environmental Change, In *Psych. Perspectives on the Environment* (eds. M. Bazerman, D. Messick, A. Tenbrunsel and K. Wade-Benzoni), Russell Sage Foundation Press.
4. Andrew Oswald and Nattavudh (Nick) Powdthavee (2006) Does happiness adapt? A longitudinal study of disability with implications for economists and judges, Institute for the Study of Labour, IZA DP No. 2208.
5. Jason Riis, George Loewenstein, Jonathan Baron, Christopher Jepson, Angela Fagerlin and Peter Abel (2005) Ignorance of Hedonic Adaptation to Hemodialysis: A Study Using Ecological Momentary Assessment, *J. of Exp. Psych.-General* 134(1): 3-9.
6. Alois Stutzer and Bruno Fey (2008) Stress that Does not Pay: The Commuting Paradox. *Scand. J. of Econ.* 110(2):339–366.
7. Tom Vanderbilt (2009) *Traffic: Why we drive the way we do*, N.Y., Vintage Books.
8. Timothy Wilson (2002) *Strangers to ourselves: discovering the adaptive unconscious*, Cambridge Massachusetts, The Belknap Press of Harvard University Press.
9. Timothy Wilson and Daniel Gilbert (2008) Explaining away a model of affective adaptation. *Perspectives on Psych. Sci.* 3(5): 370-386. This article surveys the literature on “To what kind of events do people most quickly adapt.”
10. Timothy Wilson, Thalia Wheatley, Jay Meyers, Dan Gilbert and Danny Axsom (2000) Focalism: A source of durability bias in affective forecasting. *J. of Person. and Soc. Psych.* 78: 821-36.

iii Sources:

1. Timothy Wilson (2002) *Strangers to ourselves: discovering the adaptive unconscious*. Cambridge Massachusetts, The Belknap Press of Harvard University Press.
2. George Loewenstein, Daniel Nagin and Raymond Paternoster (1997) The effects of sexual arousal on expectations of sexual forcefulness, *J. of Res. in Crime and Delinquency* 34(4): 443-473.
3. George Loewenstein and Daniel Schkade (1999) Wouldn't it be nice, In *Being: The foundations of hedonic Psychology* (eds. Daniel Kahneman, Ed Diener and Norbert Schwarz), N.Y., Russell Sage Foundation.

iv Sources:

1. Dan Ariely and George Loewenstein (2006) The heat of the moment: the effect of sexual arousal on sexual decision. *J. of Behav. Decision Making* 19(2): 87-98.
2. Javier Bravo, Paul Forsythe, Marianne Chew, Emily Escaravage, Hélène Savignac, Timothy Dinan, John Bienenstock, and John Cryan (2011) Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *PNAS* 108 (38) 16050-16055. <https://doi.org/10.1073/pnas.1102999108>
3. Antonio Damasio (1994) *Descartes' error: emotion, reason and the human brain*. N.Y., Putnman.
4. Peter Ditto, David Pizarro, Eden Epstein, Jill Jacobson and Tara Macdonald (2006) Visceral influences on risk-taking behavior. *J. of Behav. Dec. Making* 19: 99-113.
5. John Elster (1998) Emotions and economic theory. *J. of Econ. Literature* 36(1): 47-74.
6. Nitika Garg and Jennifer Lerner (2013) Sadness and consumption, *J. of Consumer Psych.* 23(1): 106-113.
7. Michael Kosfeld, Markus Heinrichs, Paul Zak, Urs Fischbacher and Ernst Fehr (2005) Oxytocin increases trust in humans. *Nature* 435(2): 673-676.
8. Aparna Labroo and Vanessa Patrick (2009) Psychological distancing: Why happiness helps you see the big picture, *J. of Consumer Res.* 35: 800-809.
9. Jennifer Lerner, Deborah Small, and George Loewenstein (2004) Heart strings and purse strings: carryover effects of emotions on economic decision, *Psych. Sci.* 15(5): 337-341.
10. Jennifer Lerner, Ye Li and Elke Weber (2012) The financial cost of sadness, *Psych. Sci.* 24(1): 72-79.
11. George Loewenstein (1996) Out of Control: Visceral Influences on Behavior, *Organ. Behav. and Human Dec. Processes* 65(3): 272-292.

12. George Loewenstein (1999) Because it is there: the challenge of mountaineering ...for utility theory, *Kyklos* 52(3): 315-44.
13. George Loewenstein (2000) Emotions in economic theory and economic behavior, *The Amer. Econ. Rev.* 90 (2): 426-432, Papers and Proceeding of the One Hundred Twelfth Annual Meeting of the American Economic Association.
14. George Loewenstein, Daniel Nagin and Raymond Paternoster (1997) The Effect of Sexual Arousal on Predictions of Sexual Forcefulness, *J. of Res. in Crime and Delinquency* 34(4): 443-73.
15. George Loewenstein, G., Prelec, D. and Shatto (1996). Hot/cold intrapersonal empathy gaps and the prediction of curiosity. Working Paper, Carnegie Mellon University.
16. Cassie Mogilner, Jennifer Aaker, and Sepandar Kamvar (2012) How happiness effects choice. *J. of Consumer Res.* 39: S312-326.
17. Robert Sapolsky (1994) *Why Zebras Don't Get Ulcers*, N.Y., Henry Holt and Company.
18. David Shoemaker (2014) SEP on *Personal Identity and Ethics*. <http://plato.stanford.edu/entries/identity-ethics/>
19. Peter Smith (2015) Can the bacteria in your gut explain your mood? New York Times June 28th. <https://www.nytimes.com/2015/06/28/magazine/can-the-bacteria-in-your-gut-explain-your-mood.html>
20. Lars Svendsen (2005) *A Philosophy of Boredom*, London, Reaktion Books. First published in Norwegian in 1999 by Universitets Forlaget.

^vSources:

1. Don Dutton and Arthur Aron (1974) Some evidence for heightened sexual attraction under conditions of high anxiety, *J. of Personality and Soc. Psych.* 30: 510-17.
2. Lawrence Williams and John Bargh (2008) Experiencing Physical Warmth Promotes Interpersonal Warmth, *Sci.*332(5901): 606-607.

^{vi} Sources:

1. Mark Bear, Barry Connors, and Michael Paradiso (2007) *Neuroscience: exploring the brain: third edition*, Lippincott, Williams, and Wilkins.
2. Kent Berridge and Wayne Aldridge (2008) Decision utility, the brain and the pursuit of hedonic goals, *Soc. Cognition* 26(5): 621-646.
3. Kent Berridge and John O'Doherty (2013) Chapter 18: "From experienced utility to decision utility" in Paul Glimcher and Ernst Fehr editors (2013) *Neuroeconomics: 4th ed.: Decision making and the brain*, Boston, Elsevier: Academic Press.
4. Kent Berridge and Morton Kringelbach (2013) Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology* 23: 294-303.
5. Kent Berridge and Terry Robinson (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.*28(3): 309-369.
6. Michael Gazziniga, Richard Ivry, and George Mangun (2014) *Cognitive neuroscience: the biology of the mind 4th ed.*, W.W. Norton.
7. Morton Kringelbach and Kent Berridge (2012) The joyful mind, *Sci. Amer.* 307(2): 40-45.
8. Cindy Wyvell and Kent Berridge (2000) Intra-accumbens amphetamine increases the conditional incentive salience of sucrose reward: enhancement of reward "wanting" with enhanced "liking" or response reinforcement, *J. Neuroscience* 20(21): 8122-30.
9. Cindy Wyvell and Kent Berridge (2001) Incentive sensitization by previous amphetamine exposure: increased cur-triggered "wanting" for sucrose reward, *J. Neuroscience* 21(19): 7831-40.

^{vii} Sources:

1. Daniel M. Bartels and Oleg Urminsky (2011), [On Intertemporal Selfishness: How the Perceived Instability of Identity Underlies Impatient Consumption](#), *J. of Consumer Res.* 38: 182-198.
2. Michael Beran (2002) Maintenance of self-imposed delay of gratification by four chimpanzees (*Pan troglodytes*) and an orangutan (*Pongo pygmaeus*). *J. Gen. Psychol.* 129:49–66. [PubMed: 12038494] [\[PDF from researchgate.net\]](#)
3. Michael Beran and Theodore Evans (2006) Maintenance of delay of gratification by four chimpanzees (*Pan troglodytes*): the effects of delayed reward visibility, experimenter presence, and extended delay intervals, *Behav. Process.* 73: 315–24.
4. Michael Beran, ES Savage-Rumbaugh, JL Pate and DM Rumbaugh (1999) Delay of gratification in chimpanzees (*Pan troglodytes*). *Dev. Psychobiol.* 34:119–27. [PubMed: 10086230]

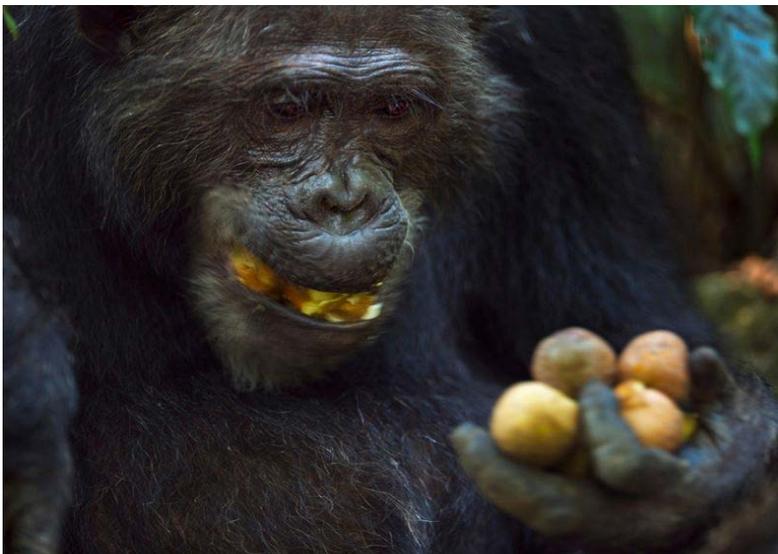
-
5. Theodore Evans and Michael Beran (2007) Delay of gratification and delayed maintenance by rhesus macaques, *J. of Gen. Psych.*, 134(2): 1999-216. DOI:10.3200/GENP.134.2.199-216
 6. Anthony Gottlieb (2016) *The dream of enlightenment: the rise of modern philosophy*, N.Y. and London, W.W. Norton and Company.
 7. Nicholas Mucalhy and Joseph Call (2006) Apes save tools for future use. *Science* 312:1038–40. [PubMed: 16709782]
 8. Mathias Osvath and Helena Osvath (2008) Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: self-control and pre-experience in the face of future tool use. *Anim. Cogn.* 2008; 11:661–74. [PubMed: 18553113]
 9. Laurie Santos and Alexandra Rosati (2015) The evolutionary roots of human decision making. *Ann. Rev. of Psych.* 66:321-347 <https://doi.org/10.1146/annurev-psych-010814-015310>
 10. Jeffrey Stevens, Elizabeth Hallinan and Mark Hauser (2005) The ecology and evolution of patience in two New World monkeys. *Biol. Lett.* 1: 223–26. [PubMed: 17148172] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1626214/>
 11. David Shoemaker (2019) Personal identity and ethics, SEP.
 12. Christine Tappolet (2010) Procrastination and personal identity, in *The Thief of Time: Philosophical Essays on Procrastination* (edited by Chrisoula Andreou and Mark White), Oxford University Press.
 13. H. Tobin, AW Logue, JJ Chelonis and KT Ackerman (1996). Self-control in the monkey *Macaca fascicularis*. *Anim. Learn. Behav.* 24: 168–74.

Chapter 7: The endowment effectⁱ

02172022

It is the most studied common quirk: thousands of articles. The *endowment effect* is the empirical finding that primates, including most humans, value a commodity more after they possess/own it. They value it more in the sense that what they must be paid to relinquish it (sell it) is greater than what they were willingness-to-pay to acquire it. Once you acquire an item that makes you better off, you are reticent to give it up, even when offered something you previously valued more.

Picture a non-human primate given a choice between a grape and a carrot (chimps prefer grapes); he grabs for the grape but is given the carrot. Even though he chose the grape over the carrot, he is now unwilling to give up the carrot to get the grape. But he will give up the carrot to get a fig: figs are much preferred to grapes and carrots.



Picture the classic 1989 [Jack Knetsch](#) experiment with coffee mugs, chocolate bars, and college students. Half were given a mug and half a chocolate bar. They were then allowed to trade, but only 10% did. In contrast, when the students were offered a choice between a mug and a chocolate bar, approximately half chose each, suggesting half preferred the mug and half the bar. But then why did only 10% endowed with a mug or bar trade for the other item? The endowment effect. Other famous empirical studies include Kahneman, Knetsch, and [Richard Thaler](#) (1990) and [Carmon](#) and [Ariely](#) (2000).

The endowment effect has even been observed with humans and other species when the item was acquired minutes earlier, even with M&Ms—people value the bag in their hand more than the bag in the researcher’s hand. If standing in the store, the individual has to be paid more than they just paid to return the item.¹ The endowment effect has been observed with different goods, different populations, children, and different primate species. In econ-speak, what you were willing-to-pay. WTP, to acquire an item (what you were willing to sacrifice to acquire it) is less than what you would have to be paid to voluntarily give it up—your WTA, willingness-to-accept its loss.

As above, the endowment effect is typically described as your WTP to obtain an item is less than your WTA its loss after it is acquired. But NBT is specified in terms of paths, not WTP nor WTA. So, it is crucial to understand it in terms of an ordering of paths.

Imagine two paths: an initial path, Path i , and a proposed path, Path i^{+A} , where A is one unit of good A . Path i has no A ; otherwise, the two paths are identical in terms of what is exogenous to the individual (same income (y), prices, and other exogenous factors). [A could be a durable good or a public good.] Path i^{+A} is ranked higher than Path i because of that unit of A . Now imagine a third path, Path i^{+B} , identical to Path i , except it has one unit of good B . Assume Path i^{+B} is ranked higher than Path i^{+A} . Initially, there are no available paths with A or B .

The individual is initially on Path i , where neither A nor B is sold or provided. The individual is then given the option of acquiring A at her wtp for A , $wtp(A)$. The amount of money would have to be subtracted from her income, y , after A is acquired, so she would rank Path i^{+A} with that subtraction the same as the initial path, Path i . Call this path, Path $i^{+A-wtp(A)}$: by definition of $wtp(A)$, the individual is indifferent between Path $i^{+A-wtp(A)}$ and Path i : they share the same rank. Because acquiring A would increase her WB, $wtp(A)$ is positive but less than $wtp(B)$. The above holds whether there is or is not an endowment effect: an endowment effect would only kick in if A is actually acquired.

¹ If the store offers them another bag plus some change to give up the bag they are holding, many individuals won’t do it. Of course, it depends on how much money is offered; everyone would return their bag of M&Ms for \$10 and a new bag, but that isn’t the issue. And you would not sell it for what you paid for it if re-purchasing required too much time and effort (driving back to the store, etc.).

Now assume the individual pays her wtp^A and acquires A , so she is **now on** Path i^{+A-wtp^A} .² Now consider her $wta(-A)$ the loss of A : it is how much money would have to be added to her current income, $(y-wtp(A))$, after A is lost, so she would rank Path $i^{(wta(-A)-wtp(A))}$ and Path $i^{+A-wtp(A)}$ the same. By definition, $wta(-A) > 0$, and she is indifferent between these two paths.

If there is no endowment effect, $wta(-A) = wtp(A)$, making Path $i^{(wta(-A)-wtp(A))=0}$ identical to Path i . And the individual is indifferent between them and Path $i^{+A-wtp(A)}$: all three have the same rank in her ordering, her ordering before A is acquired and her ordering after it is acquired. This is because the two orderings are the same: the acquisition and loss of A have not affected her ordering: the ordering was unaffected by Path i^{+A-wtp} becoming available and her taking it.

But, if the acquisition of A creates an endowment effect, by definition,

$wta(-A) > wtp(A)$, and Path $i^{wta(-A)-wtp(A)}$ is not identical to Path i : it has $wta(-A)-wtp(A)$ more income.

And because of this endowment effect, there is a different ordering of paths: before A was acquired, Path i and Path $i^{+A-wtp(A)}$ had the same rank. But after A is acquired, Path $i^{+A-wtp(A)}$ is ranked higher than Path i . Why? Because after A is acquired, the individual ranks Path $i^{+A-wtp(A)}$ the same as Path $i^{wta(-A)-wtp(A) > 0}$, but the latter is ranked higher than Path i because it is identical to Path i but with more income.

And recollect that before the individual acquired A , she ranked Path i^{+B} above Path i^{+A} . But their ranking can flip after A is acquired. The probability it flips increases with the strength of the A endowment effect and increases the smaller the initial $wtp(B)-wtp(A)$.

In summary, if acquiring A changes the individual's ordering of paths, A has an endowment effect.³

Which, if any, assumptions of NBT are violated by the endowment effect? It violates *Assumption 6A: An economicus's ordering of paths can't depend on which are currently available*. It also violates the spirit of *Assumption 4: An economicus's ordering does not change in the time-span behavior is modeled*. If the endowment effect kicks in immediately when Good

² Path i^{wtp} is ranked below Path i because it is identical to i but has wtp less income.

³ While acquiring a good with an endowment effect is sufficient to change your ordering of paths, it is not necessary: many other things can change your ordering—for example, a knock on the head.

A is acquired, Assumption 4 is violated. But how long it takes to kick in is an empirical question.⁴

NBT is typically presented as a separation of one's ordering of paths (Assumption 2) from which are available (Assumption 5). However, the endowment effect destroys this separation: the ordering often changes when a path is taken. [In Figure 2 in Chapter 1: if there is an endowment effect, the ordering isn't in set A:6a.]

But pause! While the ordering of paths is changed, does the change persist, or is it only a blip? Imagine you have two orderings: a *top-level ordering*, unaffected by endowment effects, and *ground-level ordering*, which is. Imagine the top-level ordering is how you would order paths if you currently were not on a path, and your ground-level ordering is your ordering given that you are currently experiencing a specific path. Maybe after you relinquish a good that generated an endowment effect and before you replace it with a new good, your ordering reverts to your top-level ordering. I don't know.

Does the endowment effect cause flawed choosing?ⁱⁱ

Maybe Yes. Maybe No. Yes—but only if the effect is based on an incorrect belief. You incorrectly believe that losing the item will decrease WB by more than the WB loss associated with the item's services.

No! It does not cause flawed choosing if you would experience a WB decrease above and beyond the decrease associated with losing the item's services. I will call the former an *anticipated but unrealized endowment effect*, the latter an *anticipated and realized endowment effect*. Often, the distinction isn't made.

⁴ The endowment effect leads to behavior that appears to violate [Assumption 2: At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, ...](#) But Assumption 2 isn't violated: at both points in time there is one, and only one, ordering, the new one is just different.

Table 2 outlines four possibilities.

	Realized WB drops more than acquisition increased WB	Realized WB drop equals the acquisition gain
Anticipated WB drops more than the acquisition increased WB	1. anticipated and realized endowment effect: belief correct, so choosing not flawed	2. anticipated but unrealized endowment effect: incorrect belief, so flawed choosing
Anticipated WB drop equals the acquisition gain	3. incorrect belief, so flawed choosing	4.No endowment effect: no incorrect belief, so choosing not flawed

Table 2: Realized vs. unrealized endowment effects.

Boxes 1 and 2 are endowment effects. Box3 is a realized endowment effect that was not anticipated. My sense is many economists define the endowment effect as Box 2, but I might be wrong. And whether the WB drop is realized and for how long are empirical questions.

If the anticipated loss is experienced, it explains why an increase over the long run in average National income might not cause an increase in emotional WB: economic contractions decrease emotional WB more than expansions increase it. To re-acquire the same emotional WB you had before the last recession, the subsequent expansion will need to raise your income substantially above its pre-recession level.]

Is the anticipated loss realized?

...it is unclear whether losses are actually experienced more intensely than comparable gains, or whether people simply behave as if they were. Some have argued that loss aversion isn't anything more than an affective forecasting error, while others have argued that there are many situations in which losses are actually more impactful than comparable gains" (Rick 2011)

Boyce et al. (2013), looking at income gains and losses, conclude the anticipated WB drop is realized and enduring. They estimated that an income decrease of 10K euros will decrease realized WB twice as much as a 10K increase. They analyzed a German data set of 28K households. For each household, there is data for multiple years, including yearly income, yearly answers to "How satisfied are you with your life, all things considered?" (0=totally unhappy, 10=totally happy), plus age, gender, etc. There are 63K household/year observations. The authors estimated each respondent's WB in year t as a function of their reported WB in the previous year (t-1) and as a function of whether income increased or decreased from the last year.⁵ WB increased (decreased) when income increased (decreased).

⁵ The log of the ratio of current income divided by last year's income. This transformation adjusts for the fact that a 10,000 euro increase in income has a smaller effect on WB the higher one's initial income.

They then expanded their regression, allowing the effect on WB of an income increase to be different from an income decrease. They found that realized losses affected WB more than realized gains. A second British data set of 120K observations supports their finding.⁶

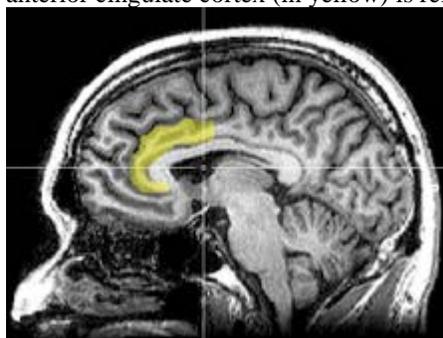
[Gehring](#) and [Willoughby](#) (2002) provide additional support that the anticipated WB drop is realized. They recorded neural activity while subjects gambled. The gambles were designed such that the financial gain (or loss) varied independently from whether the subject won the gamble. E.g., on some, you could win the bet and lose financially—you would have lost more if you had bet on the other alternative.⁷ They found that whether the individual won the gamble is recorded in a different brain region than financial gains and losses. Financial gains and losses are recorded in the anterior cingulate cortex. And, realized financial losses generate more activity than realized financial gains.⁸

In a fMRI study, [Rutledge](#) and [Glimcher](#) (2009) compared realized dollar gains and losses of equal magnitude. The losses generated more neural activity.

⁶ Instead of asking directly about WB, it asked standard psychological questions (extent respondent feels worthless, unhappy, and depressed).

⁷ Neural activity was recorded while the subjects gambled. Each subject was repeatedly presented with two squares, each contained a number; they choose a square. Then each square turned green or red (green for a gain, red for a loss) The number indicated gain or loss in cents (e.g., red 25 would indicate a loss of 25 cents if you had chosen that square). The colors and amounts appearing in each square were determined randomly, so both squares could be green or red. E.g., if the two numbers were 5 and 25, and the subject picked 5, and 5 turned green and 25 red, they won 5 cents and found out they would have lost 25 cents if they had chosen the the other square. The beauty of this design that it separates financial losses and gains from correct and incorrect choices. E.g., if both the 5 and 25 boxes turn red and you choose the 5, you experience a financial loss, but you chose correctly in that you would have lost more if you had chosen the 25 square. And, if they both turned green, you experience a financial gain but chose incorrectly in that you would have gained more if you had chosen the 25.

⁸ Approximately 1/5 of a second after the colors appears, neural activity indicates, in terms of both location and intensity, whether the subject gained or lost financially. Specifically, there was more activity for financial losses than gains at the electrode located at the medial frontal scalp (activity in the anterior cingulate cortex). And, this activity was independent of whether the individual chose correctly or incorrectly, so the increased activity in the anterior cingulate cortex (in yellow) is reflecting the loss, not an incorrect choice.



Complicating, Kermer et al. (2016) find, also with a gambling experiment, that subjects overestimate the WB impacts of losses (the authors call it an *affective forecasting error*). The authors note that their finding is consistent with a WB loss greater than the initial gain. It only implies that subjects overestimate the WB loss.⁹

In a second experiment, they give each subject \$5 and then tell them they will flip a coin. And if it's heads, the subject will get an additional \$5, but if tails, she will lose \$3 of the initial \$5, so end up with \$2. Subjects anticipated that the \$3 loss would hurt more than the \$5 gain would feel good. After the flip, they rated the two outcomes as having the same effect on their WB. As Rick (2011) points out, since a \$3 loss is less than a \$5 loss, this result is consistent with a realized loss of \$5 affecting WB more than a realized gain of \$5. “The [Kermer] data provide only equivocal support for their title, “Loss aversion is an affective forecasting error” (Rich 2011). I agree.

For the three studies just discussed, keep in mind that they investigated the WB effects of monetary gains and losses, not the gain and loss of a good.

One unanswered empirical question is, “When do you first anticipate that selling something will decrease your WB more than its acquisition increased it? Is it immediately when you acquire it? Or only after you have possessed it for a while? Or only when you first consider relinquishing it? Then there is the issue of the magnitude of the endowment effect. Is it an enduring constant, or does it wax or wane? If it wanes, the new ordering of paths will revert to the initial ordering, even though there is still an endowment effect in terms of WB

Another complication of the endowment effect is that if you sell Good *A* and use the money to buy Good *B* (e.g., sell the ski house to buy the beach house), Good *B* is now subject to the endowment effect. Is this anticipated? I sense that the literature assumes no.

⁹ [For example, my WTP for good *G* is \$*X*, my WTA its loss is \$(*X*+2), but it is an overestimate of \$*I*.

Explanations for the endowment effectⁱⁱⁱ

There are at least two competing explanations for the endowment effect: *loss aversion* and *ownership/self*.¹⁰ Loss aversion was the textbook explanation until recently, but it is being supplanted, at least among psychologists, by the ownership/self-explanation. Recent research supporting the ownership/self-explanation includes [Gawronski](#), [Bodenhausen](#), and [Becker](#) (2007), [Morewedge](#) et al. (2009), and [Maddux](#) et al. (2010)

A third possible explanation is that the behavior called the endowment effect is caused by the fact an individual does not have a complete ordering of paths. Instead, she has an incomplete ordering caused by WB-incomparabilities (some pairs of paths are not comparable in the amount of WB they generate). WB-comparability was briefly discussed in Chapter 1. Chapter 8 considers WB-incomparability and WB-incommensurability in detail. There, I consider whether an incomplete ordering can generate behaviors that look like endowment effects.

¹⁰ Two also rans are the *wealth-effect explanation* and the *cognitive-dissonance explanation*. The wealth-effect explanation ([Hanemann](#) 1991): if your real wealth is increased when you acquire something, the now wealthier you would likely have to get more to give it up than you would have originally paid for it. An example would be the Monet the previous owner forgot in the basement when he sold you your house. Another is your WTP for a protocol that would drive your nasty brain cancer in remission vs. what you would have to be paid to voluntarily abandon the protocol and have the cancer return. Your WTP was limited by your income, your WTA isn't. The wealth-effect explanation violates no NBT assumptions, but it does not explain most examples of the endowment effect: most don't include major wealth shifts.

Cognitive dissonance ([Festinger](#) 1957 and 64) is the negative feeling (stress) you experience when you hold conflicting beliefs and attitudes. A typical reaction to cognitive dissonance is to change your beliefs (reject some, modify some, or add some new ones). You chose *A* over *B*; but now that you have *A*, the fact that on some dimensions *B* is better than *A*, causes dissonance. You eliminate the dissonance by suppressing your positive thoughts about *B* and finding more reasons for why *A* is great. And, the result is your WTA the loss of *A* is now greater than what you were willing to pay for it ([Jack Brehm](#) 1956 and [Dieter Frey](#) 1986). This reaction to cognitive dissonance is “the spreading of alternatives effect”, as in, you reassess to make the difference between the chosen alternative and the rejected alternatives greater. For example, you buy the Fiat 500 rather than the Dodge Ram Truck. But now that the Cinque Cento is in your driveway, you start stressing your choice. The Fiat is fun to drive around town, easy to park, and you get lots of compliments. BUT. It is lousy in the snow and there isn't enough room for both your dog and your toys. Everything would fit in the Ram truck, and it would blast through the snow. While you chose the Fiat over the truck, some characteristics of the truck are better. You are conflicted and second-guessing your choice (experiencing cognitive dissonance). You reduce your stress by suppressing your positive thoughts about the Ram Truck and finding more reasons for why the Fiat 500 is great. You justify your choice by changing your assessments, which increases what you would have to be paid to give up the Fiat.

But—If the endowment effect is solely caused by cognitive dissonance there wouldn't be one of the item was not chosen. For example, if you inherited the Fiat from your aunt—rather than choosing it over the truck—acquiring it would not cause cognitive dissonance, and there would be no endowment effect. But the effect is observed when *A* is endowed rather than chosen, and cognitive dissonance can't explain this. And, if the endowment effect is solely caused by cognitive dissonance it's hard to imagine why some non-human animals experience it.

Loss aversion:

Loss aversion is that you are averse to relinquishing items you possess—the emphasis is on the adjective *loss*. Losing causes unpleasant emotions, or you anticipate it will—losing what you own is painful, even if the relinquishing is voluntary. Sometimes, I spend hours looking for junk: I look in the drawer for stamps and notice the worthless watch I have not worn for years isn't there. I then search for it. If I find it, I put it back in the drawer. I spent hours looking for something I have no use for. Why? I do not like to lose stuff. You might not waste your time this way, but psychologists say I have company.

There is a growing empirical and theoretical literature on whether loss aversion is genetic Huck, Kirchsteiger, and Oechssler (2005) argue that if the potential seller in a one-on-one barter is loss averse, he will drive a harder bargain. Therefore, potential sellers with loss aversion make better deals (end up better off). Thus increasing the genetic fitness (survival and reproduction) of loss-averse people.

In support, Levy (2015) argues that loss aversion minimizes the probability that your line of descendants will go extinct (maximizes the likelihood of having descendants forever). He provides evidence that for thousands of years, most exchanges were one-on-one barter (first between groups and then between individuals within a group). [His critical assumption is that bartered exchanges are *Nash-bargaining solutions*: both parties get their best outcome, given the other's outcome. So, how well a seller does depends on their bargaining skills relative to other potential sellers.]

Studying 11K Swedish twins, Cesarini et al. (2012) find that loss aversion is “moderately heritable”—with identical twins, if one twin is loss averse, the other twin is more likely to be loss averse.

But in opposition to loss aversion being genetic, Apicella et al. (2014) found a group of humans who do not experience the endowment effect, suggesting the effect is cultural rather than genetic.

We experimentally test for the endowment effect in an isolated and evolutionarily relevant population of hunter-gatherers, the Hadza Bushmen of Northern Tanzania. We find that Hadza living in isolated regions do not display the endowment effect, while Hadza living in a geographic region with increased exposure to modern society and markets do display the endowment effect. (Apicella et al. 2014).

The Hadza bushman (those Hadza living in isolation) is one of the few remaining hunter-gatherer populations. “All food brought to camp is equally shared, among all camp members, regardless

of kinship.” And “While the Hadza do own a few things, such as knives, bows and arrows, and animal skins, ownership is limited to what can be carried.” Unlike the rest of us, they lack a culture of property rights and ownership. Their lack of the endowment effect suggests that the endowment effect is a cultural phenomenon, a widespread one, and one, by evolutionary standards, recent but universal in cultures with ownership.

Amos Tversky, Kahneman, Knetsch, and Thaler are big-name supporters of the loss aversion explanation because it does not depend on how you came to own the item (a gift, a find, or you choose it over other alternatives). And according to them, sellers in research studies typically don’t rate the item higher than the potential buyers—suggesting that owning it does not increase an item’s value.¹¹

The loss aversion literature is typically vague on whether or not the endowment effect is realized. But much of this literature must implicitly believe it isn’t because it suggests that loss aversion causes flawed choosing.¹²

The ownership/self-explanation of the endowment effect

It is based on the following assumptions: (1) Everyone implicitly views themselves somewhere on a continuum that goes from an extremely negative assessment to extremely positive. E.g., considering yourself a terrible person and a failure is an extreme negative self-assessment (self-image). (2) When you acquire an item, you form an implicit association between yourself and the item: it becomes part of who you are. You transfer your self-assessment to it: it gets assessed, in part, as you assess yourself. If I imagine myself as worthless, acquiring new skis will decrease their value because they now belong to worthless me. If I imagine myself as a great guy, I value my new skis more because they are my skis.

Since most people positively self-assess, people need more money to sell the item than they were willing to pay to acquire it, which is the endowment effect. In contrast, the loss-

¹¹ See Kahneman and Tversky (1979), Kahneman, Knetsch, and Thaler (1990 and 1991) and Tversky and Kahneman (1991). Some psychologists argue that loss aversion, without an underlying theory as to why we are loss averse, isn’t much of an explanation.

¹² This characterization of the endowment effect as flawed choosing is common in the literature on financial investing (e.g. Haigh and List 2005): the argument is that you would make a higher rate-of-return on your stock portfolio if you were not averse to cutting your losses (selling at a loss). But this isn’t flawed choosing if the anticipated loss would be realized. In the financial literature, the anticipated loss isn’t realized.

aversion explanation argues that I don't like my ski more because they are my skis, but because I hate losing stuff, and selling them is a form of losing them.

Evidence for the ownership/self-explanation

Experiments by Gawronski, Bodenhausen, and Becker support the ownership/self-explanation. Their goal was to see how your implicit valuation of an item is affected separately by whether you own it, chose it, and your implicit self-assessment. The ownership/self-explanation predicts that if you negatively self-assess, ownership implies a smaller rather than a larger WTA for the item's loss, so the opposite of what it typically observed.

According to the authors, two processes determine how we value ourselves, another person, or an item. One explicit and one implicit. For example, my explicit valuation of a pair of skis is based on how they look and ski: their actual characteristics and how they work for me. In contrast, my implicit valuation is based on implicit attitudes. For example, Mikaela Shiffrin is famous and a skier beyond compare, and she endorses the ski, so I implicitly associate her greatness with the skis, independent of how they work for me.

In addition, the authors hypothesize an association between one's implicit self-assessment and one's implicit assessment of the owned item.

To test this, they measure three things: implicit self-assessment, the item's implicit value, and the implicit association the subject makes between it and themselves.¹³ Experiment 1 found that choosing which of two photos you will be gifted increases the chosen photo's implicit—the standard result.

¹³ The items were two scenic photographs. After the computer screen was blank for one second, one of the photos appeared for 1/5 a second followed by either a positive or negative word for 1/5 of a second. The subject's task was to identify whether the word was a positive or negative word. Each photo appeared 40 times, 20 times followed by a positive word and 20 times by a negative word. Positive words included *paradise*, *health*, and *vacation*; negative words included *evil*, *vomit*, and *virus*. The more often, and more quickly, the subject answers correctly when the word is positive is a measure of the implicit value of the photo. The opposite for negative words. Corrects and incorrects are tabulated along with time to response. The result is a measure of the implicit value of each photo. This is an *affective priming task*: it is widely used and has many variants. Implicit self-assessment is measured by randomly presenting letters of the alphabet and having the respondent rate the likeability of each letter. How much more the respondent likes letters that are in their initials (my name letters are *e*, *r*, and *m*) is a measure of implicit self-esteem. What! This method is common practice. For a critique of this method [Carl Falk](#) and Steven Heine (2015).

The test for association between the self and an item used a priming task like the item implicit-valuation task, but after a photo was presented, it was followed with either a "self" word (e.g., *me*, *my*, *I*) or an "other" word (e.g., *other*, *them*, *it*), and the individual had to indicate whether the word was a "self" word or an "other" word. This measure of association is increasing in how often, and how quickly, the subject correctly responds "self", and decreasing in how often, and how quickly, they correctly respond "other".

But in Experiment 1, choice is conflated with ownership. So, Experiment 2 took into account the implicit association between the self and a photo. It increased for the chosen photo and decreased for the other photo, supporting the hypotheses that with choice and ownership comes association with the self. In Experiment 3, subjects first did a self-assessment task; they then chose which photo they would be gifted and finished by implicitly valuing them. Consistent with the authors' hypothesis, there was a positive correlation between implicit self-assessment and implicit assessment of the chosen photo: the higher the self-assessment, the higher the chosen photo's implicit assessment.

Their 4th experiment was designed to rule out cognitive dissonance as an explanation for the first three experiments. Experiment 4 was identical to Experiment 3, except which photo was gifted was determined by a roll of a die, so not chosen. With no choice, there should be no cognitive dissonance. However, one still gets the result that implicit valuation of the photo is positively correlated with implicit self-assessment, implying that cognitive dissonance isn't driving the results in the first three experiments.

Adding independent support to the ownership/self-explanation, Morewedge et al. (2009) compare it with the loss-aversion explanation. They argue that most endowment-effect studies conflate loss aversion and the ownership/self effect. This occurs because most sellers own what they are selling, so they lose what they possess when they sell it. They describe the results of their Experiment 1 as, "In short, *ownership without loss aversion* [ital. added] caused the endowment effect, but *loss aversion without ownership* did not."

Explaining: They experimentally created *ownership without loss aversion* and *loss aversion without ownership* by separating buyers into two treatments: buyers who already owned an identical mug to the one they were buying (*buyer-owner*) and buyers who did not (*buyer-nonowner*). All the sellers owned the mug (*seller-owner*) in their first experiment. They find an endowment effect when the buyers are nonowners: their maximum buying price is significantly less than the minimum selling price of the owner-sellers. But find no endowment effect when the buyers are owners (*buyer-owners*). [A subject's minimum selling price is determined by asking them whether they would sell at different prices: minimum selling price is the lowest price for which they said yes. A subject's maximum buying price is the highest price they said they would buy the mug.]

To interpret these two results, the authors implicitly require that buyer-nonowners are *the loss aversion without ownership group* and buyer-owners are the *ownership without loss aversion group*. Their quoted conclusion follows from these two assumptions and their two empirical results.¹⁴

Maddux et al. (2010) provide further evidence supporting the ownership/self-explanation. If the endowment effect is driven by ownership/self, its magnitude depends on how individuals think about themselves (their self-construal). Some assess ourselves by our abilities, values, wants, and likes (independent self-construal). In contrast, one can assess themselves by their relationships with others (interdependent self-construal). The authors reasonably hypothesize that the endowment effect should be weaker for those with an interdependent self-construal. And this is what they found.

Thus, if the endowment effect is at least partly a function of the tendency to value the self—and by extension the tendency to value objects that are owned [by the self] and thus part of the self relative to non-owned objects—then one should expect the effect ...to differ due to self-construals.

Explaining: Westerners lean towards independent self-construal and self-enhancement, and East Asian towards interdependent self-construal and less self-enhancement. This suggests that East

¹⁴ One should ask what additional assumptions, or empirical findings, would justify their assumption that a buyer-nonowner must experience loss aversion. The authors implicitly assume (see their footnote 2) that people are loss averse when it comes to the gains and losses of others. When an individual makes a choice that directly affects only other people (e.g., the selling price of another's mug or the price to buy it for someone else), or which World Bank vaccination project to implement), the individual weighs a potential loss to the others of X dollars or X lives more than a gain of X dollars or X lives. [Taken on its own, it seems reasonable that if I am loss averse, others are as well, so I would take this into account if I were choosing based on what I think it best for them.] In addition, the authors assume that if people are loss averse when it comes to the gains and losses of others, it somehow causes the buyer-nonowners in their experiment to be loss averse.

In their second experiment, there are only seller and buyer agents: sellers do not own the mug they are selling, and buyers are buying for someone else. Sellers typically own what they are selling; in contrast, a seller's agent determines at what price they are willing to sell the mug but do not suffer its loss. Buyer's agents determine at what price they are willing to buy the mug but don't get the mug. There were four treatments: buyer and seller agents who owned an identical mug to the one they were buying or selling and those that did not. The findings: there was no endowment effect when both agents personally owned an identical mug. But, owning an identical mug increased both maximum buying price and minimum selling price; that is, prices were higher when both agents owned their own mug. The authors argue that these results are what one would expect if the endowment effect was driven by ownership/self effects but would not be expected if the endowment effect was driven by loss aversion. They argue that the loss aversion explanation should cause the agent seller's minimum selling price to be higher than agent buyer's maximum purchase price "because sellers' brokers should think of the transaction as a loss and buyers' brokers should think of it as a gain."

Asians will exhibit weaker endowment effects, and this is what the authors observe with coffee mugs and boxes of chocolates.¹⁵

To test whether the observed difference is caused by a difference in self-construal, the authors manipulated self-construal in a sample of 97 undergraduates at Yunnan Normal University in China. Each subject valued a coffee mug as either a buyer or seller after being divided into three treatments. The first treatment wrote a brief essay “about their friendships and camaraderie with other people and how they might foster these relationships.” This treatment enhances interdependent self-construal. In the second treatment, subjects wrote about “their unique character and skills and how they might stand out compared to other people”—thus enhancing independent self-construal. The third treatment (the control group) had no priming task. The observed endowment effect was strongest for the first treatment and weakest for the second, supporting the assumption that the difference in the strength of the endowment effect is driven, in part, by independent vs. interdependent self-construal.

Adding more support, the ownership/self-explanation suggests that we increase our implicit assessments of what we own to protect our implicit self-assessment from a threat to that assessment. For example, I must be ok despite the threat to my self-image because my house and car are valuable. In this case, the ownership/self-explanation suggests that the value of an owned good should increase after the self has been threatened. If there is a positive association between the self and what you own, a threat to your self-image should cause the value of the items you own, which is what [Dommer](#) and [Swaminathan](#) (2013) find.

In summary, for years, loss aversion was the standard explanation of the endowment effect, cognitive dissonance was a distant second, and the ownership/self-explanation was either rejected or not considered. But in the last fifteen years, the ownership/self-explanation has emerged as the forerunner, at least in psychology. I find both loss aversion and ownership/self intriguing explanations, and they can coexist—and I don’t like to lose things. Think of a sense-of-loss as a feeling/emotion, an unpleasant, specific emotion, like sadness and anger. Imagine its intensity depends on how much WB increased when the item was acquired. That said, my new

¹⁵ The study of coffee mugs included 116 Northwestern students (59 identified as White or European U.S. citizens, 57 identified their ethnic background as Asian or East Asian). The chocolate study included 105 students at the University of Waterloo in Canada (34 identified as White or European Canadian citizens, 60 identified as Asian or East Asian).

gravel bike (coming soon) will surely increase in value when it's mine, which will improve my self-image, but first, I need to sell a bike.¹⁶

So, what does the ownership/self-explanation imply about whether the endowment effect loss is or isn't realized?

The loss is realized.

Because my assessment of the item—which was, in part, a reflection of my self-assessment—transfers back to further reinforce my self-assessment. I value my new skis more because I transfer my positive self-image to the skis, which then feeds back, enhancing my self-image. If so, I experience more WB because the skis are mine, so my WB will drop when they are no longer mine, and the endowment-effect loss is realized. In which case, my hesitancy to sell the skis isn't flawed. When I sell my ski, my WB drops more than the drop caused by the loss of the skis' services.

The endowment effect: summing up

It violates *Assumption 6A: An economicus's ordering can't depend on which paths are currently available*, and the spirit of *Assumption 4: An economicus's ordering does not change in the time-span behavior is modeled*. It causes flawed choosing only if the WB drop is unrealized. If the ownership/self-explanation is correct, there is no flawed choosing because the anticipated loss is realized. Whether it is realized is an open empirical question.

The endowment effect is complicated, and unanswered questions abound. It is typically applied to the loss of something that provided WB, but what about the loss of items that are currently reducing your WB: the pain in your knee from an old sports injury or the old toaster that randomly burns the toast? They are also part of your self-assessment. Does the endowment effect imply that tossing the toaster will increase your WB more than you anticipate?

And how does its magnitude vary with what is lost? For example, chimps exhibit the endowment effect with food, but not toys. Little is known about how the endowment effect in

¹⁶ The rule in my house, and in many of my friend's homes is, "You can't buy another bike until you get rid of at least one of your current bikes." While I know of no specific study of the endowment effect and bicycles, I am sure how it would turn out. Bikes have characteristics that make them prime candidates for the endowment effect: bicycles are objects you experience; in action, your body and bike become integrated, they are long lived, they are one of the few objects you also experienced as a child; you compete with yourself and others on a bike; you socialize on a bike, and they are highly customizable, so one of a kind. See [Karly Coleman's](#) thesis (2015) on cyclists' emotional attachment to their bikes.

humans varies by what is lost (a relationship, an activity, a commodity, income) and what type of WB was lost (happiness, love, life satisfaction, etc.).

An aside: Buddhism has, for thousands of years, preached the perils of both loss aversion and ownership/self, arguing that they, combined with the fact that all (possessions, relationships, life) is *impermanent* (nothing lasts), lead to ill-being. The road to a less unsatisfactory life is to avoid the experience of loss by enjoying what you have, limiting one's possessions, and breaking the ownership/self-link between your implicit self-esteem and your possessions. See Batchelor (2015).

Summarizing the effects of the common quirks on NBT: collectively, they cause systematic incorrect beliefs, flawed choosing, and fluctuating orderings. That your ordering changes isn't a new criticism. Incentive salience, the endowment effect, and the emotional empathy-gap are simply reasons why.

A few thoughts on modeling the common quirks:

The common quirks argue for a behavior model involving a sequence of decision points. Either your current path is no longer available, or new and attractive available paths are. Decision points occur when prices change, a pandemic arrives, you lose or gain a job, you get sick, a baby arrives, your partner leaves or a new one arrives, and whenever uncertainties become certain.

I would initially model the sequence by restrictively assuming that there are only a small number of available paths at each decision point, limiting them to feasible paths similar to the one you are leaving.

For a few emotional states (e.g., angry, aroused, tired, hungry, and other), I would suppose a constant but different top-level ordering. For a utility function, picture different top-level parameters for each emotion, e . In the absence of cues or the endowment effect, the ground-level ordering for emotion e is the top-level ordering for emotion e . But the ground-level ordering can temporarily deviate from its top-level ordering due to cues or the endowment effect.

In empirical modeling, it is hard to imagine a researcher observing what cues an individual is experiencing and how they affect that individual's behavior.¹⁷ So, I might model their influence with random parameters. Imagine a sequence of discrete-choice random-utility models with random parameters (a mean and variance parameter) on the different path components. The important thing to capture is that your choices tell us less about your top-level ordering than is typically assumed. There are already many estimated repeated, random-parameter models, and one might argue that the behavioral effects of random cues are one reason there are so many random-parameter models.

The endowment effect can be modeled and estimated. At every choice point, the researcher observes what you own. When the individual considers which path to take next, Path j 's utility parameters momentarily shift on owned items lost if one takes Path j . For example, imagine that Path j involves replacing the current car. At the decision point, the parameter on car ownership shifts down for all paths that include replacing it.¹⁸ Later, the parameter on car ownership shifts back to its top-level value.

Emotions are observable and can be invoked in the lab. As suggested above, imagine a top-level ordering for each emotion, e . If the researcher observes an individual's choice and the emotion they were experiencing when they made that choice, the research can estimate emotion-specific top-level utility parameters. This would be a conditional model, conditional on emotion.¹⁹ A hard part is modeling the sequence of an individual's emotions, which would be required to model their sequential choices.²⁰ One place to start would be modeling the behavior

¹⁷ Cues are often registered unconsciously, and what might be a cue for me, might not be for you. That said, cues could be manipulated in choice experiments. This is probably already happening in choice experiments, but the researcher is not necessarily aware they are doing it. For example, in studies to estimate environmental damages, the description of the injury likely includes phrases or photos that act as cues.

¹⁸ If the current path does not include a car, the parameter on utility from a car does not shift—you can't experience the endowment effect for something you don't own.

¹⁹ One might imagine a discrete-choice latent-class model with class-membership probability a function of emotional state. Morey and Thiene (2017) value choice as a function of personality—the same model would work for emotional states. If a researcher followed individuals over time recording their purchases, they could also ask questions about their emotional state.

²⁰ If emotional state can't be predicted, one could model emotional state as a latent variable, and identify variables that influence the probability that individual i is in emotional state e .

of a group of individuals who are persistently in emotional state, e (e.g., the chronically depressed or chronically anxious).²¹ I defined a persistent emotion as a *temperament*.

Alternatively, in the lab, one could invoke different groups with different emotions, and then ask hypothetical choice questions.²²

Does NBT survive the common quirks?^{iv}

NBT based on *Assumption 9b: Paths that are more desired/wanted are ranked higher* (but not 9a), survives quite well; desires are desires: it does not matter whether those desires were produced by an emotion, a physiological state, a cue, a lack of empathy, or what you will like. Flawed WB-choosing is immaterial: WB is not the objective. Fluctuations in your wants and desires make it more difficult to predict your behavior but don't cause flawed desire-based choosing. In fact, fluctuating desires make it easier to fulfill more desires. If today I want to leave Shirley for Wanda, but tomorrow I want to leave Wanda for Shirley, more of my wants are fulfilled if I am successful on both days.

NBT based on *Assumption 9a: An economicus's ordering of paths is based on WB...* does not survive the common quirks. NBT based on Assumption 9a survives duration bias and lack of empathy for one's future self—and neither is inconsistent with *Assumption 4: An economicus's ordering does not change in the time-span behavior is modeled*.

The emotional empathy-gap, the endowment effect, and incentive saliences each assault NBT: each violates Assumption 4—a questionable and problematic assumption. The common quirks highlight that your ordering fluctuates. It is surprising this isn't more discussed. The common quirks reinforce why there is flawed WB-choosing.

²¹ **Morey and X (2222) estimated the preferences of depressed individuals for different treatment programs. Other examples include**

²² A large part of my life has been estimating recreational choice (how many times you go skiing, fishing, or whatever, and where you go when you go). My ancient dissertation was a model of where and how often individuals ski. Consider how to estimate skiing choices when the WB effects of skiing depend on whether the individual is currently anxious, depressed, or curious. [I imagine that one has little interest in skiing when one is depressed; when anxious one ranks the familiar and safe ski areas higher, but when one is curious one ranks the unknown higher.] Picture a laboratory experiment with skiers where an emotion is invoked (e.g., anxiety), and subjects are then presented with hypothetical ski choices that vary in difficulty and familiarity.

For some commodities and activities, choice is insensitive to emotional state, for others choice is extremely sensitive. For example, I suspect that an individual's demand for air conditioning depends on the outside temperature but does not vary much with their emotional state, while the demands for alcohol and sweets are quite sensitive.

The common quirks, ethics, and welfare economics:

The ethical implications of quirks depend on whether the ethical goal is maximizing WB or fulfilling wants.²³ Welfare (well-faring) economics is built on the former. What is right (morally preferred) for the individual is choosing their highest-ranked available WB-path. For society, it is maximizing societal WB in an aggregate sense.²⁴

In contrast, for many, the ethical goal is want-fulfillment: what is right for the individual is choosing their highest-ranked path in terms of wants fulfilled; for society, it is maximizing, in an aggregate sense, the amount of wants fulfilled. Peter Singer, the most famous living utilitarian, argues for want fulfillment. Economic ethicists who agree with Singer are not welfare economists: their objective isn't to increase WB. If one adopts 9b and the goal of want fulfillment, each individual left to their own devices will do what is right for them, no matter how their behavior is affected by the common quirks. And there is no ethical reason to interfere in anyone's behavior unless it impinges on the ability of others to fulfill their wants and desires.

In contrast, if the goal is WB (the goal of welfare economics), *false choosing* should be avoided. While NBT is consistent with false choosing,²⁵ the common quirks explain false choosing and its direction. Incorrect beliefs are a significant cause of false choosing, duration bias, the emotional empathy-gap, and the future-self empathy-gap. All involve incorrect beliefs.

Incentive salience also causes flawed choosing by pushing desiring further from liking.²⁶

More nuanced, the endowment effect only causes flawed choosing if the anticipated loss is not realized. Flawed choosing caused by the common quirks is flawed in terms of increasing WB but not flawed in terms of want fulfillment. So, for welfare economists, the common quirks motivate paternalism: you are not doing what is best for you. For example, you are not saving enough for your old self. So nudging you to save more is warranted.²⁷

²³ Economists who consider how to parse right from wrong behaviors are typically unclear on whether the objective is fulfilling wants or maximizing WB.

²⁴ If the objective is WB, is it maximizing realized WB, or is it choosing the path associated with more WB in some subjective-probability sense? I will not consider this complication for reasons of taste, clarity, and brevity.

²⁵ As long as full certainty is not assumed.

²⁶ It would be confusing to say that increased desiring is an incorrect belief: its intensity really increases.

²⁷ See, for example, Thaler and Sunstein (2008) and Thaler and Benartzi (2004).

If one wants to argue that less government interference is ethically preferred, claiming so is easier if you believe in NBT based on Assumption 9b and believe the ethical objective is the fulfillment of wants and desires.

The argument for less government is more challenging if you adopt Assumption 9a and believe the ethical objective is more WB. While they violate only a few NBT assumptions, the common quirks pile havoc on the notion that an individual left to his own devices will choose his highest-ranked available path, ordered in terms of realized WB. [Arguing for no governmental controls requires, in addition, that one imagines individual behavior has no effect on others, either their WB or wants. See the discussion of welfare effects in Chapter 12.]

This ends our conversation about the common quirks, making it an excellent time to take another general look at critical terms in Chapters 6 and 7.

Knowledge, beliefs, incorrect beliefs, subjective probabilities, and best estimates

(Review more philosophical lit on beliefs.) These concepts play an important role in behavior, choice, and WB. They are both distinct and overlapping. For example, a subjective probability is a belief based on knowledge or maybe not. And if you believe something, can you know it is an incorrect belief? Probably not—if you knew it was incorrect, you would not believe it.

While our beliefs affect what we buy, who we marry, and who we vote for, they also directly affect our WB. An incorrect belief might enhance WB, while a correct one causes despair. So, an incorrect belief might lead to greater WB even though it causes choices that, assessed alone, reduce WB. For example, my disbelief in global warming might cause me more overall WB, even though it caused me to buy a Swiss ski chalet where snow will soon be a thing of the past.

Alternatively, an incorrect belief can directly increase WB and lead to choices that further increase WB. For example, suppose God and heaven don't exist, but I believe they do. As a result, I spend a lot of time attending church, reading the scriptures, and being nice. All is a waste of time because these activities will not increase my probability of making it to heaven.

But I am a happy camper, happier than if I were an atheist. And studying scripture reinforces my incorrect belief, which leads me to choose paths in the future with even more scripture time.²⁸

Put simply, the WB associated with a path depends on the set of beliefs associated with that path. And, one's ordering of paths depends on one's beliefs. Your beliefs will differ depending on whether you choose to experience path *m* or path *t*. This looks like a circular conundrum: A causes B, and B causes A, so one's beliefs and one's ordering are not separable concepts, a conclusion troubling for NBT and for explaining behaviors. I leave it at that.

ⁱ Sources:

1. Ziv Carmon and Dan Ariely (2000). Focusing on the forgone: How value can appear so different to buyers and sellers, *J. Consum. Res.* 27(3): 360–370. [doi:10.1086/317590](https://doi.org/10.1086/317590)
2. Daniel Kahneman, Jack Knetsch, and Richard Thaler (1990). Experimental tests of the endowment effect and the Coase Theorem, *J. of Polit. Econ.* 98 (6): 1325–1348. [doi:10.1086/261737](https://doi.org/10.1086/261737). [JSTOR 2937761](https://www.jstor.org/stable/2937761)
3. Jack Knetsch (1989) The Endowment effect and evidence of nonreversible indifference curves, *Am. Econ. Rev.* 79(5): 1277-84.
4. William Harbaugh, Kate Krause and Lise Vesterlund (2001) Are adults better behaved than children: Age, experience and the endowment effect, *Econ. Letters* 70: 175-181. At <http://www.pitt.edu/~vester/econletter2001.pdf>

ⁱⁱ Sources:

1. Christopher Boyce, Alex Wood, James Banks, Andrew Clark, and Gordon Brown (2013) Money, well-being, and loss aversion: does an income loss have a greater effect on well-being than an equivalent income gain? *Psych. Sci.* 24(12): 2557-62. DOI: 10.1177/0956797613496436
2. William Gehring and Adrien Willoughby (2002) The medial frontal cortex and the rapid processing of monetary gains and losses, *Sci.* 295(5563): 2279-82. DOI: 10.1126/science.1066893
3. Deborah Kermer, Erin Driver-Linn, Timothy Wilson, and Daniel Gilbert (2006) Loss aversion is an affective forecasting error, *Psych. Sci.* 17(8): 649-53. DOI: [10.1111/j.1467-9280.2006.01760.x](https://doi.org/10.1111/j.1467-9280.2006.01760.x)
4. Scott Rick (2011) Losses, gains, and brains: neuroeconomics can help to answer open questions about loss aversion, *J. of Consumer Psych.* 21: 453-463.
5. Robb Rutledge and Paul Glimcher (2009). Asymmetric BOLD responses to positive and negative outcomes. Poster presented at the annual meeting of the Society for Neuroeconomics

ⁱⁱⁱ Sources:

1. Coren Apicella, Eduardo Azevedo, Nicholas Christakis, and James Fowler (2014) Evolutionary Origins of the Endowment Effect: Evidence from Hunter-Gatherers, *Am. Econ. Rev.* 104(6): 1793-1905. <http://www.jstor.org/stable/42920866>
2. Jack Brehm (1956) Post decision changes in the desirability of alternatives, *J. of Abnorm. and Soc. Psych.* 52: 384–389. <https://pdfs.semanticscholar.org/aea3/2a2d5a76f2f7c9cab957a4ed160bc8648fb8.pdf>
3. Sarah Brosnan (2011) Property in non-human primates, *New Direct. For Child and Ad. Dev.* 132: 9-22. doi:

²⁸ The influence of beliefs is further complicated by the fact that some beliefs are unconscious and these sometimes conflict with your conscious beliefs: you can have dueling beliefs. For example, many white people hold the conscious belief that they are not prejudiced against non-whites, but, in fact, are. To check out your racial prejudices, conscious and unconscious, take the test at [Harvard's Project Implicit](https://www.harvard.edu/implicit/). Quoting from the web page, "It is well known that people don't always 'speak their minds,' and it is suspected that people don't always 'know their minds.' Understanding such divergences is important to scientific psychology. The web site presents a method that demonstrates the conscious-unconscious divergences much more convincingly than has been possible with previous methods."

10.1002/cd.293

4. David Cesarini, Magnus Johannesson, Patrik Magnusson, and Bjorn Wallace (2012) The behavioral genetics of behavioral anomalies, *Manage. Sci.* 58(1): 21-34. <http://dx.doi.org/10.1287/mnsc.1110.1329>
5. Karly Coleman (2015) Bicycles as objects: identity, attachment, and membership categorization devices. M.A. thesis, University of Alberta. <https://era.library.ualberta.ca/items/384f7788-e638-4a01-adf9-2cbe8e7b7e80>
6. Rafael Di Tella, John Haisken-De New and Robert MacCulloch (2010) Happiness adaptation to income and to status in an individual panel, *J. of Econ. Behav. and Org.* 76(3): 834-852.
7. Sarah Dommer and Vanitha Swaminathan (2013) Explaining the Endowment Effect through Ownership: The Role of Identity, Gender, and Self-Threat, *J. of Consumer Res.* 15: 1034-50. DOI: 10.1086/666737
8. Carl Falk and Steven Heine (2015) What is implicit self-esteem, and does it vary across cultures? *Personality and Soc. Psych. Rev.* 19(2): 177-98. DOI:10.1177/1088868314544693
9. Leon Festinger (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
10. Leon Festinger (1964). *Conflict, decision, and dissonance*. Stanford, CA: Stanford University Press.
11. Timothy Flemming, Owen Jones, Laura Mayo, Tara Stoinski and Sarah Brosnan (2012) The endowment effect in orangutans, *Intern. J. of Comp. Psych.* 25: 285-298 http://www.comparativepsychology.org/ijcp-vol25-4-2012/2.Flemming_etal_Final.pdf
12. Dieter Frey (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41–80). New York: Academic Press. <https://www.sciencedirect.com/science/article/pii/S0065260108602129>
13. Bertran Garwronski, Galen Bodenhausen and Andrew Becker (2007) I like it, because I like myself: associated self-anchoring and the post-decisional change of implicit evaluations, *J. of Exp. Soc. Psych.* 43: 221-32.
14. Michael Haigh and John List (2005) Do professional traders exhibit myopic loss aversion? An experimental analysis, *J. of Finance* 55(1): 523-34.
15. Michael Hanemann (1991) *Willingness To pay and willingness to accept: how much can they differ? Reply*, *Am. Econ. Rev.* 81(3): 635–647. doi:10.1257/000282803321455449
16. Steffen Huck, Georg Kirchsteiger, and Jorg Oechssler (2005) Learning to Like What You Have- Explaining the Endowment Effect, *Econ. J.* 15(505): 689-702. <https://doi.org/10.1111/j.1468-0297.2005.01015.x>
17. Daniel Kahneman, Jack Knetsch, and Richard Thaler (1990). (Ibid)
18. Daniel Kahneman, Jack Knetsch and Richard Thaler (1991) The endowment effect, loss aversion, and status quo bias, *J. Econ. Perspec.* 5: 193–206.
19. Moshe Levy (2015) An evolutionary explanation for risk aversion, *J. of Econ. Psych.* 46: 51-61. <https://doi.org/10.1016/j.joep.2014.12.001>
20. William Maddux, Haiyang Yang, Carl Falk, Hajo Adam, Wendi Adair, Yumi Endo, Zvi Carmon and Steven Heine (2010) For whom is parting with possessions more painful? Cultural differences in the endowment effect, *Psych. Sci.* 21(12): 1910-17. DOI: 10.1177/0956797610388818
21. Cary Morewedge, Lisa Shu, Daniel Gilbert and Timothy Wilson (2009). *Bad riddance or good rubbish? Ownership and not loss aversion causes the endowment effect*, *J. of Exp. Soc. Psych.* 45: 947-951.
22. Amos Tversky and Daniel Kahneman (1991) Loss aversion in riskless choice: a reference-dependent model, *Quart. J. of Econ.* 106: 1039-61.

^{iv} Sources:

1. Alan Hajek (2019) Interpretation of Probability. Stanford Encyclopedia of Philosophy.
2. Edward Morey and Mara Thiene (2017) Can personality traits explain where and with whom you recreate? A latent-class site-choice model informed by estimates from mixed-mode LC cluster models with latent personality traits. *Ecological Econ.* 138:223-237.

Conscious and unconscious mental processes: whose ordering of paths is it?ⁱ

08172022

As noted in Chapter 1, while most believe humans have conscious thoughts and consciously choose, Assumptions 1-9 don't imply this: NBT does not require consciousness. Assumption 11 (self-awareness) implies consciousness but not conscious choosing. Assumption 12 (the choosing axiom) adds it; it makes you consciously in charge—you choose, and this determines the path you experience. This chapter starts our discussion of conscious thought and unconscious processes in behavior and choice. Part II continues the conversation in more detail.

In 1782 Rousseau reflected on the unconscious when he realized why he had been avoiding a corner: a boy on that corner would always ask him for money.

That is what I discovered on reflection, for until then none of it was distinctly in my thoughts. This observation has made me recall a multitude of others that convince me that the true and original motives of most of my acts are not as clear to me as I long supposed.

In the path of a bus, you jump back rather than staying put, making a "choice". This happens even before the situation registers as electrical activity in the parts of your brain that embed conscious thought. Alternatively, sometimes you experience cogitating for weeks on whether you should choose option *A* or *B*.

Economists do not distinguish between conscious and unconscious cognition. In psychology and neurobiology, the distinction is central to explaining behavior, beliefs, and how we make our way. Our brain consists of many separate processing centers: some for conscious thought (ruminating about when I should retire) and others unconscious but instrumental in determining what I do, including when I retire.

The unconscious was once viewed as limited to harboring our infantile Freudian traumas and coordinating physical activities (language, not falling over, and seeing—activities the conscious mind did not need to be bothered with). Now, it plays a leading role in most of our actions. For a persuasive and accessible presentation, see Wilson (2002).

Keep in mind that if we consciously consider and decide how to act before we act, this does not mean it causes us to act the way we did. I thought hard and long about buying a vacation home before doing so, but maybe both the thinking and the doing were caused by my

unconscious reaction to a prior event. [Event B precedes C, but B does not necessarily cause C; instead, both are possibly caused by an unobserved event A].

Research on the unconscious indicates that its contents are determined by genetics, culture, and experience. For an example of research on how unconscious motives are formed, see [McClelland](#), Koestner, and Weinberger (1992).

There are two strands of research on the conscious's role in perception formation and choice. One strand argues actions are (all or mostly) initiated by the unconscious, and before the act occurs, and if there is time, the unconscious causes a conscious experience of choosing. This strand, *The Illusion of Choice*, is discussed in Part II. The other strand assumes beliefs are formed and actions are initiated by both the conscious and unconscious. In this strand, decisions made by the unconscious are more instinctive/programmed/quick than those formed consciously:

A hallmark of the adaptive unconscious is automaticity, whereby information is processed in rapid, unconscious, involuntary ways—Wilson.

When the bus is barreling down, when you lock eyes with the attractive passing stranger, and when the pause in the conversation gives you only an instant to say something funny or witty, quick action is required. And the action taken is determined by one of our unconscious processors: cogitation would take too long. According to the second strand, when there is leisure to cogitate, our conscious thoughts play a role in what we will do, but not always. The context also plays a significant role in whether a choice is made by our conscious or unconscious minds. Context includes emotional state, your current self-image, and whether someone attractive asked you to do it.

Survey questions on which option you prefer are typically answered quickly, suggesting the unconscious might play a larger role than the conscious in determining whether you circle option A or B or whether you circle "mostly agree" or "somewhat disagree". The timelines for consumer purchases are more varied: sometimes we agonize over a purchase for months; sometimes, we go to Costco for groceries and come home with a portable power generator.

And, the different processing centers don't always agree on the course of action. Quoting Wilson

*There is little research on the consequences of having disparate conscious and nonconscious 'selves' that are out of synch. An exception is the work of Joachim Brunstein and Oliver Schultheiss... they found little correspondence, on average, between people's nonconscious and conscious motives. **Maybe add some detail about their methodology.***

We are not of one mind, either at a point in time or over time. Many questions arise. Who is in charge? It depends on what is being determined and the context—contextual orderings. Are we conscious a decision was made? Not always. Are we aware/conscious of why we do what we do? Probably not, but our conscious brain makes up a good story to explain what was determined by our unconscious, or so say many psychologists. And, what happens when our conscious and unconscious disagree?

Economists typically imagine the brain as a single black box that embeds one stable ordering of paths. They also don't think about how or why this ordering arises, arguing the mechanics of its creation are unimportant because the only goal is to predict choices. We are not concerned with why you like what you like. [A camp in psychology called *behaviorists* ruled the psychology journals in the 20th Century's first half. Again, quoting Wilson, "The behaviorist onslaught in psychology was fueled by a rejection of mentalism; behaviorists argued there was no need to consider what occurred inside people's heads, consciously or unconsciously." Psychology has largely rejected behaviorism; economics has not.]

The goal of both NBT (and NBT plus) isn't only to predict what is chosen but to deem it the best choice.

If I have an ordering of paths, am I consciously aware of it?

Am I conscious of this ordering and able to describe it, or are only parts of it revealed to me when I observe my choices? I could imagine having an ordering but not a conscious awareness of it. My dog prefers meat to milk, milk to bread, and meat to bread (yes, his ordering is transitive), but he isn't consciously aware of it. He isn't unique in my household; quoting Shakespeare, "I have much ado to know myself." Bertrand Russell agrees with Shakespeare:

The discovery of our own motives can only be made by the same process by which we discover other peoples, namely, the process of observing our actions and inferring the desire which could prompt them—Bertrand Russell, The Analysis of the Mind.

Does it matter whether I am aware of my ordering?

If the researcher's intent is to estimate your ordering (your demand curves), and if she is doing this by observing your real-world choices, it does not matter whether you were, a priori, aware of your ordering. But what if she asked the hypothetical question, "Would you prefer A or B?" Many economists mistrust hypothetical-choice questions asserting the responses don't reflect the respondent's ordering of paths, explaining the disconnect is caused by lying or a lack of

introspection. "Ask a hypothetical question, get a hypothetical answer."¹ But the issue is more profound. If I am not conscious of my ordering, I can't recover it through conscious thought, no matter how hard I introspect. And, my conscious mind will make up a reason for why I said I preferred A. Research shows we rationalize our choice to maintain a positive, or at least consistent, self-image. I would be embarrassed to say, "I am clueless as to why I do what I do." so make something up and believe it? As Wilson and J.I. Stone suggest, another person might be a better predictor than me of why I do what I do.

Maybe, the way to learn about your own ordering is to observe your behavior or ask your friends. And conclude that your willingness-to-pay for a flat-screen TV, or environmental cleanup, is what others would pay, after adjusting for age, gender and income. In summary, introspection might be a misguided tool for determining how you order a set of paths.

ⁱ Sources:

1. J.C. Brunstein, Oliver Schultheiss and R. Grassman (1998) Personal goals and emotional wellbeing: the moderating role of motive dispositions, *J. of Person. and Soc. Psych. Rev.* I: 3-25.
2. Peter Diamond and Jerry Hausman (1994) Contingent valuation: Is some number better than no number" *J. of Econ. Perspect.* 8(4):45-64.
3. Michael S. Gazzaniga and Joseph E. LeDoux (1978) *The integrated mind*, N.Y., Plenum.
4. Dan Gilbert and Timothy Wilson (2000) Miswanting, In J. Forgas, ed., *Thinking and feeling: the role of affect in social cognition*, Cambridge Massachusetts, Cambridge University Press.
5. David McClelland, Richard Koestner and Joel Weinberger (1992) How do self-attributed and implicit motives differ? In C.P. Smith, ed., *Motivation and personality handbook of thematic context analysis*, Cambridge, Cambridge University Press.
6. Jean-Jacques Rousseau (1979) *Penguin Classics: Reveries of the Solitary Walker*, London, Penguin Ltd. (First published in 1782 as *Les Reveries du promeneur Solitaire*)
7. Daniel Wenger (2002) *The Illusion of Conscious Will*, Cambridge MA, MIT Press.
8. Daniel Wegner and Thalia Wheatley (1999) Apparent Mental Causation: Sources of the Experience of Will, *Amer. Psych.* 54(7): 480-92.
9. Timothy Wilson (2002) *Strangers to ourselves: discovering the adaptive unconscious*, Cambridge Massachusetts, The Belknap Press of Harvard University Press.
10. Timothy Wilson and J.I. Stone (1985) Limitations of self-knowledge: more on telling more than we know, In Shaver, ed., *Rev. of personality and Soc. Psych.*, Vol. 6, Beverly Hills, Sage.

¹See, for example, the Diamond and Hausman article. Interestingly, economists who disparage stated-preference data believe we have a stable and unique ordering and observed choices are a manifestation of that ordering.

Chapter 8: Returning to whether all the different kinds of WB are commensurable and all the different bearers-of-WB are comparable

08172022

It is the incommensurability that is the crux of the problem. It is this incommensurability of 'plural; values that provides perhaps the main challenge to Utilitarianism, as well as to any attempts to present a unique system of moral theory (Wilfred Beckerman 2017)

Many believe they're not always. Chapter 1 pointed out that Assumptions 2 and 9a

Assumption 2: At every point in time, an economicus has one, and only one, ordering of all paths—1st, 2nd, An economicus has an ordering if for all paths j and k , either Path j is ranked higher than Path k , Path k is ranked higher than Path j , or paths j and k have the same rank.

Assumption 9a: An economicus's ordering of paths is based on its WB (well-being), WB, accounting for uncertainties. The more WB economicus associates with a path, the higher its rank

imply that an economicus has a complete WB-ordering of all paths. Chapter 1 also demonstrated that Assumption 2 is violated if any paths are WB-incomparable. And, if any kinds of WB are incommensurable, there will be WB-incomparable paths. Recollect from Chapter 1:

*Paths h and k are WB-incomparable if **none** of the following statements are correct: h is ranked higher than k , k is ranked higher than h , or h and k have the same rank. They are WB-comparable if one, and only one, of these statements, is correct.*

And

WB-commensurability: WB kind A and WB kind B are incommensurable if you are incapable of comparing them in terms of overall WB

This chapter discusses the arguments for and against complete WB-comparability and complete WB-commensurability.

Before jumping in, here are a few things to recollect from Chapter 1: Even if paths h and k are incomparable in terms of overall WB, they still could be comparable in terms of a specific kind of WB (e.g., Path k generates more happiness than h , but less life satisfaction). And Paths h and k being comparable does not imply that either is comparable with any other paths.

The existence of bearer-of-WB incomparabilities doesn't mean you can never WB-compare paths. For example, paths with more chocolate cake and fewer rats than the current path, all else the same, are ranked higher even if cake and rats are WB-incomparable. Also, be mindful that

rats and cake might be WB-comparable at a few levels, but not all. Maybe I can conclude that a path with 75% fewer rats and one less cake is WB preferred to my current path, but not be able to compare if the alternative path has only 15% fewer rats.

More on why complete WB-incommensurability is importantⁱ

Complete WB-comparability requires complete WB-commensurability. So, if there is not complete WB-commensurability, the individual's ordering of paths will not be complete. And this has significant implications for behavior: choice theory becomes less specific in its predictions.¹

The impact of incompleteness on behavior and choice

If the ordering is incomplete (Assumption 2 is violated), the theory only predicts the individual will choose a path not ranked lower than any other available path (“maximizing” but not “optimizing”).² And there can be many paths that maximize (Sen 1979, Hsieh 2007 and 2016). In which case, we can’t say which the individual will select— we can only determine which available paths will not be chosen.³ Consider again Figure 4, in Chapter 1, but now suppose the three commodities are market goods, *m*, less global warming, *lgw*, and third, instead of relationships, saved souls, *s*. Suppose *m* and *s* are WB-comparable, but neither is with *lgw*. Finally, assume the origin path is available. For convenience, I repeat the figure here.

¹ Recollect from Chapter 1, that WB-incompatibility is sufficient, but not necessary for incompleteness to occur. In the 1920s and 30s, WB-commensurability was debated in the context of the functionality of socialism (the ability to compare values in the absence of market prices. See Neurath (1928:1973), Von Mises (1920 in Hayek 1935), and Hayek (1935) on the functionality of socialism (the ability to compare values in the absence of prices.

² Maximizing and optimizing are equivalent when the ordering is complete.

³ Maximizing when the ordering is incomplete is akin to *satisficing* (Simon 1982): The individual concluding, “I’m satisfied, and I know I could have done worse.”

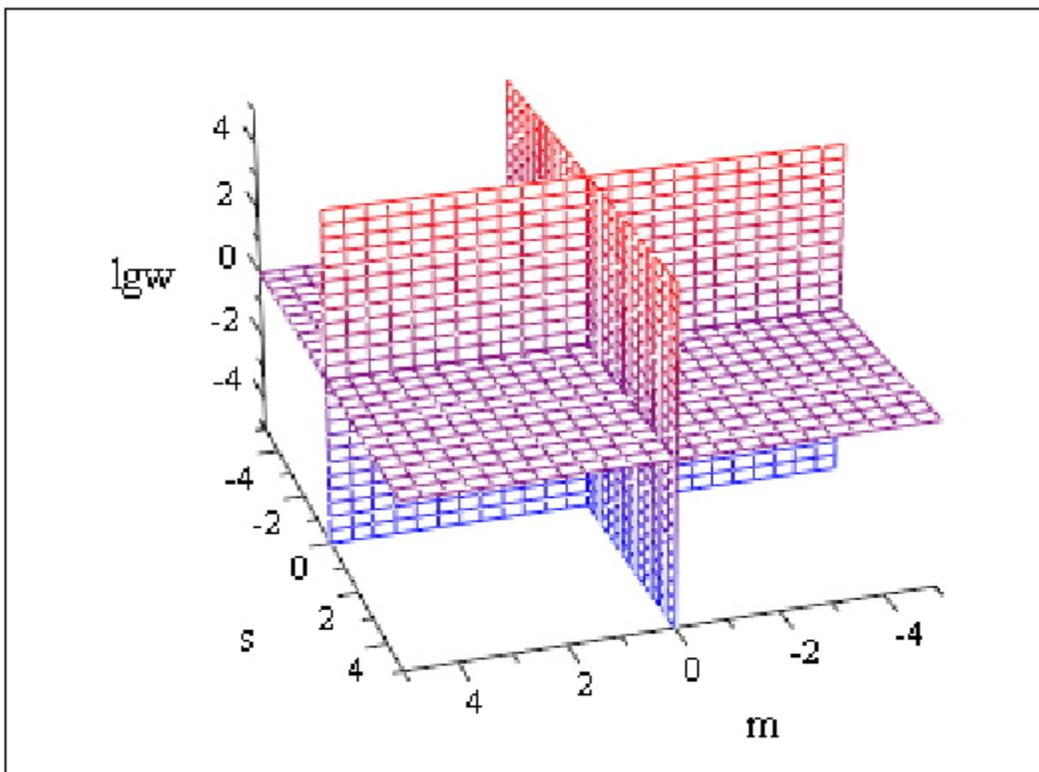


Figure 1, Chapter 8: m - s - lgw paths (Fig. 4 in Chap 1 with r replaced by s)

Repeating a bit from Chapter 1: all paths with the same level of gw as the origin path are ranked relative to each other; Paths in Quadrant A (front top-left) are ranked higher than the origin path; those in G (back bottom-right) are ranked lower. Paths in the interiors of Quadrants B-F and H are not ranked relative to the origin path.

To predict behavior, the budget set must be specified. Suppose the individual has a fixed income, y , and the prices p_m and p_s are exogenous. [The Church posts the cost of saving a soul and accepts PayPal]. The level of gw is exogenous at the level represented by the horizontal plane through the origin path. And the origin path exhausts her budget. The *available set* of paths is the horizontal gw plane on and to the right of the m/s budget line through the origin (budget line not shown). The budget line's slope in the horizontal plan reflects the relative prices of m and s . It goes from right front to back left (as her consumption of market goods increases, the number of souls she saves must decrease).

The individual takes gw as given and selects the highest-ranked path on her budget line (she optimizes): her income will be exhausted, and her chosen path might, or might not, be the

origin path. She has a complete ordering of all available paths, even though she can't order all conceivable paths.

The impact of incompleteness on transitivity

Recollect from Chapter 1 that a complete ordering—meaning all paths are WB-comparable—is by definition transitive. As noted there, transitive means that for all paths i, j, k , and l , if Path i is ranked at least as high as j and j as least as high as k , and k at least as high as l , then i is ranked at least as high as l . Transitivity no longer holds for all sets of paths if the ordering is incomplete.

Intransitivity is worrisome because it can produce incoherent behavior called *money pumps* (Bossert and Suzumura 2010, Hansson 2018). Put simply, you might voluntarily exchange Path i for j , then j for k , then k for l , and then pay to get back to i —you end up back where they started but with less money—so you pumped money down the drain. This led many to infer that say coherent behavior must be transitive. But they are wrong.

In explanation, while with incompleteness there is no longer transitivity, there still can be a restrictive property on the relationships between paths called *Suzumura consistency* (*S-consistency*) (Suzumura 1983, Bossert and Suzumura 2010, Bossert 2018, Sen 2018): if Paths i, j, k , and l have the relationships j is ranked at least as high as i , i is ranked at least as high as k , and k at least as high as l , then S-consistency means l will **not** be ranked strictly higher than j .⁴

S-consistency is both necessary and sufficient to make money pumps impossible. Put simply, incomplete orderings don't necessarily imply incoherent behavior, even though incompleteness negates transitivity.

⁴ An example of S-consistency but not transitivity: Returning to Figure 1 with market goods, m , saved souls, s , and less global warming, lgw , consider four paths, i, j, k , and l . Let i be the path at the origin. Define Path j to have the same amounts of m and s as Path i but less gw . So, in terms of Figure 1, it is vertically right above Path i . Therefore, Path j is ranked strictly higher than i (so ranked at least as high). Let Path k be on the horizontal indifference curve that passes through i but with more s and less m than Path i . The individual is indifferent between Paths i and k , so i is ranked at least as high as k . Finally, let Paths l and k have identical amounts of m and s , but Path l has more gw , so Path l is vertically right below Path k . So, Path k is ranked strictly higher than l (so ranked at least as high as l). Summarizing: j is ranked at least as high as i , i at least as high as k , and k at least as high as l . But now examine paths j and l ; j has less gw than l but also less s . Since in this example, lgw and s are WB-incomparable, Paths j and l are not comparable, so not ranked relative to each other, and the ordering is incomplete. Transitivity is violated: Path j is not ranked at least as high as l , but S-consistency holds: l is not ranked strictly higher than j .

Transitivity implies S-consistency, but S-consistency does not imply transitivity. If the ordering is complete, transitivity and S-consistency amount to the same.⁵

Individuals sometimes behave as if their ordering of paths is intransitive. Incompleteness would explain. If your ordering is incomplete, you will sometimes be intransitive. But if you are S-consistent, you are coherent and won't money pump.

The impact of incompleteness on valuing, in money a shift from Path i to Path t

Consider the implications of a complete ordering vs. an incomplete one on my life's work: defining and estimating how an individual values, in money, a shift from Path i to t . For example, the enactment of a government policy that will reduce the rate of global warming.

The most common money measure is the *compensating variation*, CV, for a shift from Path i to t , where paths are expressed in terms of what is exogenous to the individual on that path. Typically, these exogenous variables are in three categories: the individual's income on that path, the prices of market commodities (quantities consumed endogenous), and fixed components (those whose quantities/levels are exogenous).

If the individual has a complete ordering—the neoclassical assumption—the CV for a change from the origin path to any other path exists. A CV exists if either (I) or (II) holds.

(I): There exists a finite amount of money (positive or negative) that, when subtracted from income on the proposed path, causes the individual, after the subtraction, to be indifferent between the two paths. With the compensation, the two paths are tied at the same rank. If such an amount exists, it is a finite amount of money and is the CV. If the shift from i to t improves WB, this finite CV is a positive number and willingness-to-pay, WTP, to enact the shift. If the shift from i to t would decrease WB, this finite CV is a negative number and, in absolute terms, willingness-to accept, WTA, the shift.

Or (II): No matter how much is subtracted or added to income on the proposed path, the initially ranked higher path remains higher (the same rank can't be achieved). For Case II, if the shift from Path i to t would improve WB, the CV is positive infinity and indicates that the individual

⁵ Note that S-consistency is different from quasi-transitivity, a different weakening of transitivity that was defined in Chapter 1. S-consistency is also different from another weakening of transitivity called *acyclivity*: if Paths i , j , and k have the relationships j is ranked higher than i , and i is ranked higher than k , then acyclivity means j is ranked at least as high as k . Acyclivity is less restrictive than S-consistency. Neither S-consistency nor quasi-transitivity imply the other.

would be WTP all the money they have, no matter how much they have, to enact the shift. Conversely, suppose the shift from Path i to t would decrease WB. In that case, the CV is negative infinity indicating that for no amount of money would the individual voluntarily accept Path t . [N.B, II implies the ordering, while complete, has a lexicographic property. Lexicographic was defined in Chapter 1. Unfortunately, many economists who value amenities in terms of money do not understand the distinction between a complete but lexicographic ordering and an incomplete ordering.]

For both Cases I and II, money is a meaningful comparison metric. It says a lot about you and me if you would pay \$1000 to half the rate of gw , and I would only pay \$10. It also says a lot if you would pay your entire income, no matter its magnitude, to half the rate.

If the individual has a complete ordering of paths, the CV for every path shift is defined. But if the ordering of paths is incomplete, the CV is undefined for some shifts. In my example, commodities m and s are WB-comparable, but neither is WB-comparable with gw , so the ordering of paths is incomplete, as noted above.

Consider a shift from Path i to t , where i is the origin path in Figure 1. If paths t and i are not WB-comparable, a CV for the shift is meaningless because there is no answer to whether paths i or t would produce more WB. For example, if Path t has more m or s than Path i and more gw , there is no CV . There is also no CV if Path t has less m or s and less gw .

If the ordering is incomplete, the CV can be meaningless even when Paths i and t are WB-comparable. Continuing with our example, assume Path t is identical to Path i (the origin path) except that it has less gw (in Figure 1, Path t is now directly above i , like Path j). In this case, Path t is WB-preferred to i , but a CV remains meaningless because the kinds of WB produced by less global warming are not comparable to those produced by m and s .

Think about this in terms of a budget set where y , p_m , p_s , and gw are exogenous to the individual. A CV for a shift from Path i to t exists only if it does not change the exogenous level of gw . Otherwise, the paths can't be compared in terms of money, making it meaningless to talk about this individual's WTP to reduce global warming or her WTA an increase in global warming. [One could still collect data, estimate a model, and incorrectly call the result an estimate of WTP to reduce global warming: it is impossible to produce an estimate of something

that does not exist.]. The CV for a change in p_m or p_s is also meaningless if the price change causes g_w to change, which it likely will.⁶

WB-comparability and commensurability are also critical to tort law (e.g., can you be compensated with money for the loss of an arm, or loss of a cognitive ability, or loss of a loved one?).⁷ WB-comparability is also critical to assessing whether a system or economy is WB-sustainable, and it is critical to hedonic wage and price studies.⁸

Modeling behavior when the ordering is incomplete

Whether you believe an incomplete ordering of paths makes for a more realistic and better behavior theory is your call. Sen (2018) thinks it would: quoting him, “I explore the need for allowing—and incorporating—incompleteness of preference in choice theory...”⁹ I agree. Above I noted incompleteness as an explanation for coherent intransitivities.

An incomplete ordering can exhibit as an endowment effect (Chapter 7) and, more generally, as a *status-quo effect* (staying on the current path even though additional paths have

⁶ Market goods and saving souls are not likely to be carbon equivalent.

⁷ See, for example, the *University of Pennsylvania Law Review* symposium on commensurability ([Matthew Adler](#) 1998), and Sunstein (1994).

⁸ If the objective is the sustainability of overall WB, the objective is only meaningful if it is possible to convert all packets of WB into overall WB. Hedonic wage and property-value studies are a method of estimating the CV for amenities such as cleaner air, lower crime rates, better schools, and distance to parks. They are dependent on the assumption that in equilibrium everyone is indifferent to where they work and live. The existence of this equilibrium requires complete WB-commensurability.

⁹ Interestingly, Sen also asserts that WB-incomparabilities are a “much-hyped issue” and a “mundane occurrence”. I have stated that WB-incomparabilities cause the ordering of paths to be incomplete making complete WB-comparability a necessary condition for a complete ordering. Sen questions this: his argument is that one can, often, order paths even if they contain components that are WB-incomparable and neither path Pareto dominates the other. He finds incomparability ubiquitous, but that it, in his, view does not stop us from ranking most pairs of paths. Our difference comes down to how one defines “incomparable”. Consider two of his examples. “If I love banana much more than apples, I would not be deterred from going for a banana by the peculiar worry that apples cannot be measured in the same unit, which is what commensurability is concerned with.” I’m confused, “loving bananas much more” seems to say they are comparable, at least in terms of love. And what is the same unit in “same unit”. Then he considers the choice between a fine mango that provides nutrition, “as well as some palatal or olfactory pleasure, and a vinal record of a song that would “offer a different reward (not immediately reducible in the dimensions of the other). But, given a budget constraint, we could quite possibly face the choice of having one or the other. This involves choosing between noncommensurable results. And yet we might have no great difficulty in opting for the mango when hungry and starved and going for the song when well endowed with tasty food but short of melodious entertainment. The choice need not be hard to make in many situations, despite the noncommensurability involved. The distinct dimensions of value might not be reducible into one another and yet there might be no problem whatsoever in deciding what one **should** do ...” Maybe our different takes come down to commensurability vs. WB-commensurability.

become available). For example, why switch if another path becomes available but is not WB-comparable with the current path? Only switch if it is ranked strictly higher than the current path—incompleteness due to WB-incompatibilities encourages remaining on your current path ([Mandler 2004](#)).

Recollect the famous 1989 [Jack Knetsch](#) experiment with coffee mugs, chocolate bars, and college students that I discussed in Chapter 7. Half were given a mug, half a chocolate bar, and then could trade. Only 10% in each group traded. In contrast, when students were offered a choice between a mug or a chocolate bar, approximately half chose each, “suggesting half preferred the mug and half the bar”. But then why did only 10% endowed with a mug or bar trade for the other item? Knetsch, and most everyone else, reports this as evidence of an endowment effect. But it could, instead, be an example of WB-incomparability. The student knows that getting the chocolate or the mug is WB-wise better than getting neither. Still, if the student cannot WB-compare the chocolate with the mug, there is no reason to trade for the other once you acquire either. This is a status-quo effect caused by an incomplete ordering.¹⁰

So, precisely what is required for WB-commensurability?

WB-commensurability requires that all the kinds of WB generated by a bearer can be separated from the bearerⁱⁱ

The pleasure from eating chocolate must be separable from the eating, and your pride in your children must be separable from what they did to make you proud. If not, different kinds of WB couldn't be aggregated independently of their bearers. This necessary condition is called *WB/bearer separability*—think of it as a type of consequentialism (only the act's consequences matter, not the bearer of those consequences).

But complete WB/bearer separability is not sufficient for WB-commensurability—an example demonstrates: the separation of happiness from the bearers of happiness and the

¹⁰ Half chose the chocolate and half the mug when given a choice between the two. But this does not imply half like the chocolate more and half like the mug more. An inability to compare them would generate the same result: you select by mentally flipping a coin if you can't compare chocolate and mugs. I am not claiming that chocolate and mugs are WB-incomparable; instead, I am pointing out that an incomplete ordering can produce behavior that looks like what an endowment effect would cause.

separation of life-satisfaction from its bearers doesn't imply happiness and life-satisfaction are WB-commensurable.

The rejection of WB/bearer separability goes back to Aristotle. Nussbaum (2012):

Throughout his [Aristotle's] work, he insists on the tremendous importance of qualitative distinctions among the diverse constituent parts of human life;... pleasure is something that comes along with, supervenes on, activity, "like the bloom on the cheek of a young person." it is so closely linked to the relevant activities that it cannot be pursued on its own, any more than bloom can be adequately cultivated by cosmetics... what Aristotle has in mind is that pleasure is a kind of awareness of one's own activity,...

She argues that J.S. Mill rejected WB/bearer separability, and so do contemporary philosophers (2012, p 338).

Consider your ability to separate a sensation such as pain from the activity or circumstance that produced it. For example, the same chest pain could be because you just ran your best marathon, found your partner in bed with someone else, or you are having a heart attack. Whether a sensation increases or decreases different kinds of WB depends on whether you chose it, whether you think it will be gone in the morning, what you imagine is causing it, and whether experiencing it helped you achieve an important goal. Humans think about the causes of their sensations: this suggests difficulty in separating the feeling from its cause.

The WB/bearer separability issue is whether everyone can always separate every kind of WB from its bearer. Maybe you can, but Sunstein (1994) says most of us can't. According to him, the *awe* (an emotion) produced by viewing a mountain differs from the awe produced by viewing a skyscraper, which differs from the awe produced by a remarkable musical or athletic performance. Since these awes can't, in his view, be separated from their bearers, they are WB-incommensurable. And besides, they are incommensurable with the kinds of WB produced by consuming conventional goods and services. Keep in mind that he is asserting WB/bearer inseparability, not proving it.

Many bearers of WB are public goods in that everyone experiences their existence (e.aking., if gw is reduced for you, it's reduced for everyone, same for saving a species from extinction). An issue is whether it is easier or harder to separate the kinds of WB produced by a bearer from the bearer if the bearer is a public good. I don't have an answer.

I have sympathy for WB/bearer inseparability. Still, I suspect it depends a lot on the specific bearer and the specific kind of WB, and the extent to which that kind of WB has a significant cognitive component.

Non-humans vs. humans

I imagine that the extent to which kinds of WB can be separated from their bearers varies across species: becoming more difficult the greater the species' ability to cogitate. All animals experience sensations, but only humans consciously think about them in light of their past experiences, knowledge, and the big picture. And the world and your reactions to it produce new sensations, perceptions, thoughts, and emotions. Contrast this with animals who experience positive and negative sensations but are more limited than us in their cognitive capacity to assess and evaluate them. I conjecture that it is easier for my dog than for me to separate pleasure from the activity that produced it and easier for a worm than for my dog.

Related to the above conjecture is my conjecture that there are fewer kinds of dog WB and ill-being than human kinds, and there are still fewer for worms. Why? Humans experience more thoughts and emotions than other primates, primates more than dogs, and worms none. And, the more thoughts and emotions one can experience, the more kinds of WB and ill-being there are.

For humans, the ultimate existentialist fear is fear of their death, but most other living things are incapable of experiencing it. So we don't need to be told, "You and I were born to die" ([Lana Del Ray](#), the sadcore album *Born to Die*, 2012). As the philosopher [Arthur Schopenhauer](#) (1788-1860) put it, "The brute first knows death when it dies, but man draws consciously nearer to it every hour he lives;" (Schopenhauer 1818/2011, p. 68).

A flawed argument for complete WB-comparability: flawed because it is circular

In common economic vernacular, the argument goes as follows: everyone makes choices involving trade-offs over different kinds of WB, so everyone must be able to compare them. What makes this logic circular is saying everyone makes trade-offs over different kinds of WB requires that, for everyone, all paths are WB-comparable, which requires that, for everyone, there is complete WB-commensurability. Put another way, the argument starts by assuming what it wants to demonstrate. [*circulus in probando*: quoting Wikipedia, "the reasoner begins with what they are trying to end up with. Circular reasoning is not a formal logical fallacy but a pragmatic

defect in an argument where the premise (in this case, everyone makes tradeoffs over different kinds of WB) is just as much in need of proof or evidence...”]

One way to prove the premise that all paths are WB comparable would be to prove what? That everyone constantly is experiencing their highest WB-ranked available path? If so, no one has proven it, and many psychologists believe they have disproven it. See the studies cited and discussed in Chapters 6 and 7.

At this point, you might find it helpful to consider, again, the distinction between *behaviors* and *chosen behaviors*—we all do things (behave in specific ways), but this does not imply that all of our behaviors are chosen behaviors.¹¹ I hope everyone would agree that if there are N alternatives on the table and the individual must experience only one, the individual will experience one. Behavior is required. So, what is the best word to neutrally describe the alternative experienced? I am inclined to go with the *alternative experienced*; the term *alternative chosen* implies what one is trying to demonstrate.

A weak defense of the premise that all paths are WB-comparable is that people will tell you that they chose the alternative they are experiencing by consciously trading off all the kinds of WB that would have been generated by the different alternatives. And We believe we do this. This is obviously more convincing than if one said, “I have no clue as to why I do what I do.” But saying something does not make it true. Hundreds of psychological studies indicate that our reasons for doing what we do are often made up, post hoc, to justify what we are doing. See Chapter 10.

You are wrong if you think you can WB-compare apples and oranges simply because you ate the orange rather than the apple. Eating the orange doesn’t imply you “chose” the orange: you had to eat something.¹² More generally, just because you believe you can compare two bearers-of-WB or two kinds of WB does not mean you can.

Don’t use circular reasoning to defend WB-comparability and commensurability.

¹¹ Chapter 9 is about the distinction between behaviors and choices (chosen behaviors)

¹² Sen (1979) calls this *decision inescapability*: you have to start another path no matter what.

Two kinds of WB are WB-incommensurable if an ability to compare them is incompatible with experiencing themⁱⁱⁱ

Experiencing some kinds of WB is inconsistent with comparing them with other kinds. Quoting [Tetlock](#) (2003), "...incommensurability arises when values are treated as commensurable subverts one of the values in the trade-off calculus." Consider, for example, the kinds of WB produced by a meaningful and symmetric relationship ([Raz](#) 1986). He argues you can't compare a loving relationship with market goods because it isn't a loving relationship if you or your partner can. [People do compare, but they are not in love, or so the argument goes.] Feeling love and friendship, feeling trusted, and the safety in knowing trusted others have your back are all produced by personal relationships. It is difficult to argue that these kinds of WB are comparable to those produced by market goods.¹³ Consider also one's relationship with God. You can't experience the grace of God if you can compare his grace with the taste of chocolate cake.

What about personal relationships between members of different species, such as Giacomo and me? Or if I feel a personal bond with nature. Starting with the dog, I think it is fair to say that "I love him", that he is bonded to me, and that we enjoy most of our interactions. In the future, I will likely have to put him to "sleep" or delay his death with vet bills, but that does not imply I can compare my love for him with the kinds of WB produced by market goods. If I could, I question whether I am capable of loving him. Could the same question of capability be raised about a love of nature?¹⁴

Consider the inconsistency argument for resources that you believe to be sacred: when one is in the presence of a feature that they believe to be sacred (e.g., a geographical feature, historical site, or religious shrine), spiritual and religious thoughts and emotions are produced. There are corresponding negative thoughts and emotions if something sacred is lost or denigrated. Tetlock (2003) defines "sacred values as those values a moral community treats as

¹³ Sunstein (1994): "Suppose that [Adam] Smith has arranged to have lunch with a friend today, but that he has become very busy and perhaps would like to cancel. Suppose Smith thinks in this way: if he cancels, his friend will be disappointed, because he would like Smith's company, and also a bit insulted, because it is cavalier and disrespectful to cancel lunch at the last minute. Maybe Smith should make it up to him, or provide compensation, by offering a nontrivial cash payment...[but] A cash payment would be inconsistent with *the way* that someone values a friend. Even if the friend would prefer \$1, or \$10, or \$100, or \$100 to lunch with Smith—the offer of cash would be perceived as an insult rather than a compensation. In this context, the difference in kinds of valuation means that a financial exchange would be inappropriate."

¹⁴ For many of us, our relationship with living things is fundamentally different from our relationship with goods and services. Unlike golf balls, animals suffer when they are struck, and much animal suffering is caused by humans.

possessing a transcendental significance that precludes comparisons, trade-offs, or any mingling with secular values.” A synonym for “sacred” is “inviolable”—“secure from assault or trespass”—Merriam-Webster. Sacred is a belief; if one believes a feature is sacred, its existence evokes reverential awe, a kind of WB. For indigenous populations, natural environments generate the kinds of WB associated with culture and religion, including the experience of sacredness. But such experiences are not limited to indigenous populations.¹⁵

Experiencing sacredness is inconsistent with comparing it with the kinds of WB that money can buy, so a sacred resource is incompatible with it being WB-comparable with market goods. Remember that observing a profane individual consuming market goods does not imply honoring the sacred is WB-comparable with market goods. The individual had to do something. [See the flawed—because it is circular—argument outlined above.]

If an economist wants every kind of WB to be commensurable with kinds produced by market goods, she will have to argue that nothing is sacred. Maybe nothing is—I don’t know. [We do know that for some individuals, nothing is sacred. I won’t mention names.]

Most of us prosper from having a cultural identity: belonging to a group, being supported by that group, and sharing its values—belonging to the tribe. And many of us would feel a significant loss if we were exiled. Or if the group and its values were threatened. Related to cultural identity is one’s way of life: what one does, friends and family, and occupation. It, like one’s cultural identity, bears kinds of WB. These bearers are WB-incompatible with market goods if an ability to compare them is inconsistent with a capacity to experience the kinds of WB unique to one’s cultural identity and way of life. Maybe culture and goods are WB-comparable for people who believe they are, but not for people who don’t? For many people, shopping and consumption is their culture.

I think about whether experiencing certain emotions is possible if their effects on your WB are commensurable with the kinds produced by market goods. For example, consider your ability to experience grief due to losing a loved one. It is a painful emotion, but you would want

¹⁵ There is a long, and continuing, history of humans believing some environmental resources are sacred (e.g., specific species and certain geographical features and places). Many secular Westerners poo-poo sacredness, but such beliefs are common elsewhere and in other groups. Even among Westerners, there are many individuals who believe nature has sacred components.

to feel grief given your loss. Can one experience grief if one can compare it with the pleasures of market goods? Can I grieve for the dead wife I bumped off for the insurance money? I might miss her chocolate cake, but grieve? Could I have really loved her?

Consider depression (a kind of ill-being). It affects most people and severely affects the WB of many. [Remember that a primary way to increase WB is to reduce ill-being.] A characteristic of severe depression is that you are incapable of experiencing many kinds of WB—the pleasures of an ice-cream cone on a hot day can't be experienced—often, you can't even get out of bed. Given the nature of depression, how could you compare the ill-being of depression with the kinds of WB produced by market goods?¹⁶

Consider *inclinations*: on your current path, you are, in your view, a woman of integrity, whereas, on my current path, I'm selfish and lacking in empathy (a bit of a sociopath). Consider paths that vary by one's inclinations and consumption levels. A complete ranking of paths in terms of WB requires that you can WB-compare the kinds of WB associated with different inclinations with the kinds produced by market goods. For example, can you experience honesty if you can WB-compare it with a path that keeps your kid out of jail, but you are dishonest? Can I WB-compare two paths if I am a sociopath with lots of toys on the current path, and on the other, I am empathetic but toyless? Can I know what empathy would feel like if I have never experienced it?

The kinds of WB produced as you acquire knowledge about a process are WB-incommensurable with the kinds of WB produced by a change in its level.^{iv}

Consider the WB-effects of acquiring knowledge about the process of global warming (its causes, its effects, man's role, etc.) compared to the WB-effects of a change in its rate. I chose gw for this example because it's a global public process that will cause a lot of ill-being, and many are gw ignorant and seem to want to stay ignorant.

¹⁶ In some of my past research I have assumed you can. In hypothetical choice experiment I have asked depressed subjects to choose between treatment programs, programs that varied by dollar cost and effectiveness. See Morey, [Jennifer Thacher](#) and [Edward Craighead](#) (2007a and b). But maybe some of my subjects were incapable of such WB-comparisons.

Education leads to knowledge acquisition, and the acquiring produces different kinds of WB.¹⁷ Education (reading, studying, and listening) is the bearer. When unsuccessful, it generates ill-being (confusion, frustration, a sense of failure). Then there is the issue of how you are affected by the new knowledge. New knowledge about processes directly affects WB. For example, acquiring knowledge about gw and cancers can affect many different kinds of WB in many different ways.

The question at hand is whether knowledge WB is WB-commensurable with the kinds of WB that would be produced by a change in the level of what you have acquired knowledge of. Knowledge, once acquired, can't be traded away. You can sell your house and car, but you can't unlearn things, even if what you learned decreases your WB.¹⁸ Another aspect of knowledge about a process is that before you acquire it, you are ignorant. You don't always have a good prior as to whether its acquisition will increase or decrease your overall WB. These two aspects of knowledge suggest that it is difficult to compare education with other bearers of WB.

Consider my knowledge of gw. For simplicity, suppose paths vary only in my understanding of gw and the expected rate of gw. Can I compare these two bearers-of-WB? I'm not sure; there are inherent difficulties. To keep things simple, assume everyone agrees that a lower expected rate would increase their WB, but I don't know how much my gw angst will decrease because I'm uneducated on the topic.

Further, imagine that my uneducated prior is that while global warming will be bad, it will not be terrible.¹⁹ How would I determine whether I would have more WB in a world where I'm more educated about the effects of global warming, and there is a different expected rate? I don't know that I could make all such comparisons.

Now flip the example, and imagine I'm already educated about global warming (so know how it works and its effects). Then I'm asked to decide whether I would experience more WB if

¹⁷ One subset in the list of emotions are *epistemic emotions*: emotions that have a knowledge component (Scarantino and de Sousa 2018). Examples include curiosity, surprise, validation, feelings of doubt or certainty, the feeling of knowing, and the bliss of ignorance. So, learning for learning's sake can directly increase or decrease WB.

¹⁸ In a recent study, [Ananda Ganguly](#) and [Joshua Tasoff](#) (2016) found that 16% of the university students sampled were willing to forego a payment of \$10 to avoid learning whether they had genital herpes. David Eli and [Justin Rao](#) (2011) have found that people will pay to not learn about their IQ or how attractive others find them.

¹⁹ Also think through the example imagining that in my ignorance I imagine global warming is worse than it actually is.

I knew less, and there was a different expected rate-of-global-warming. How do I assess not knowing what I already know?

In summary, many people are ignorant about physics, chemistry, and biology, in particular the environmental sciences, ecosystem dynamics, and the health effects of chemicals in the environment. And the kinds of WB we experience when our environment changes (species go extinct, the expected rate-of-global-warming changes, there is less (or more) PCB contamination) are a function of our level of knowledge—ignorance can be bliss. But once knowledge is acquired, it is not freely disposable. Comparing environmental kinds of WB with knowledge about the environment is difficult, causing me to suspect I'm incapable of ordering paths in terms of their environmental effects and my understanding of those effects.

While my example has been global warming, this comparability conundrum arises whenever paths vary by the levels of different processes and the individual's understanding of those processes. For example, compare a medical test that will inform you whether the black mole on your toe is cancer with being treated for that cancer. Compare educating yourself on the research on the benefits and costs of immunizing your children vs. immunization rates. Compare learning the actual probability of being killed in a terrorist attack vs. how much is spent on fighting terrorism.

You cannot compare certain kinds of WB because it is morally unacceptable to even imagine such comparisons.^v

How much money would you need to push the button that would extinct elephants or dump a gigaton of PCBs in Lake Michigan? How much would you need to eliminate equal rights? Many regular people, but not neoclassical economists, would find it unimaginable to compare these bearers with market goods or even WB—compare the existence of elephants with equal rights. Suppose you can't WB-compare the extinction of elephants with ending Biden's Presidency. That means you can't WB-rank elephants and no Biden vs. Biden and no elephants. Or, if you like both elephants and Biden, it means you can't WB-rank elephants and no Biden vs. Biden and no elephants. [Remember my caution: an inability to compare for moral reasons is different from comparing and concluding there is no amount of money you could be paid to be willing to give up your morals.]

Inability to compare because of moral or cultural sanctions is likely to vary drastically by culture and religion. Research indicates that secular Westerners with high socioeconomic status (Europeans and North Americans) are willing to make comparisons that non-Westerners and Western religious conservatives will not make ([Jonathan Haidt](#) and Jesse Graham (2007) and Graham, Haidt and [Brian Nosek](#) (2009)). The two latter groups, but not the first group, morally require in-group/loyalty, authority/respect, and purity/sanctity. This makes it difficult for them to compare sanction violations with money. [Loyalty, respect, and purity are not something one compares with chocolate cake.] For many American First Peoples, the kinds of WB born by their culture and natural surrounding are likely WB-incommensurable, on moral grounds, with the kinds of WB resulting from market goods and services.²⁰

But it is not only comparisons with cake. For Western religious conservatives, WB-comparability requires that the kinds of WB reduced by disloyalty are commensurable with those reduced by impure and sacrilegious acts. Complete WB-comparability requires that one be able to WB-compare, for example, desecration of a cross, disrespect for the President, religious freedom, a ban on abortion, income (in)equality, global warming, animal extinctions, and market goods.

Many secular Westerners believe that we have a moral obligation to the environment, a preservation ethic. Consider the famous quote by [Aldo Leopold](#) in his 1940s essay “The Land Ethic”,²¹

A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise.

Summarizing his ethic, humans don't stand alone. They are part of the community of plants, animals, lands, and waters (collectively, the “land”), and, as members of that community, we

²⁰ Tribes are often trustees in NRDA cases. To cite an extreme example, consider the “water pollution, soil contamination, deforestation, and cultural upheaval” (Wikipedia: Lago Agrio oil field) resulting from oil exploration and extraction in the Sucumbíos Province of Ecuador caused by Texaco (now owned by Chevron) and its effect on the indigenous tribes who, arguably, lost their culture and way of life. Chevron and Ecuador battled over damages in both the U.S. and Ecuadorian courts. If their culture and money are, in fact, WB-incomparable, monetary damages are undefined.

²¹ Leopold's *Sand County Almanac* was published by his son in 1949 (shortly after Leopold's death). It is a collection of essays: essays about the land in Wisconsin where he had a farm (the part of Wisconsin with sandy soil), essays about other regions, essays about his personal experiences in nature, and essays that sketch an environmental ethic. The “Land Ethic” is the last essay in the book. The book was unknown until the 1970s. It, along with [Rachel Carson](#)'s *Silent Spring* (1962) are foundations of the environmental movement in the U.S.

must respect its other members. They have the “right” to exist and prosper (e.g., animal rights), and it is our moral responsibility to see that they flourish—we play the role of parents, and parental responsibility is inconsistent with sacrificing one’s children for money.²² Accepting the right of others to exist and prosper is, arguably, inconsistent with their demise being WB-comparable with market goods.

An environmental aside: ecological economists accept as gospel that complete WB-commensurability does not exist^{vi}

This is what distinguishes ecological economists from environmental economists: environmental economists take complete WB-commensurability as gospel. To start:

Ecological economists believe that there are kinds of WB that only the environment can produce

Are there? I’m not sure, but many people believe it. If there are, more wilderness, less gw, and saving a species from extinction are bearers.

Kinds of WB unique to natural environments include the freedom and self-reliance that can only be experienced in the wilderness. They include experiencing the joys of in-situ learning about plants and animals. They include experiencing the awe that only can be produced by the sight of majestic mountains and rivers. They might include the kinds of WB produced by nature-based recreational activities (hiking, skiing, biking, swimming, diving, fishing, hunting, and camping).²³ Also, the aforementioned kinds of WB natural environments produce for indigenous populations. And, the argument that an understanding of the significant relationships that natural environments create (e.g., human to human, human to animal, animal to human, animal to animal, animal to plants, and plants to animal) contributes, in a unique way, to a “worthwhile life” (see [Alan Holland](#) 2006 and Dan Firth 2008). Many of us aspire to a worthwhile life.²⁴

²² There are differing views on how Leopold’s might have applied his land ethic in specific situations, so I suspect there would be disagreement amongst environmental ethicists as to whether Leopold believed the WB humans get from the land are incommensurable with other kinds of WB. That said, many environmentalists say they are incommensurable and view Leopold as a source.

²³ In contrast, there are kinds of WB that are experienced in natural settings such as the thrill of going fast on skis, a bike, or a snowmobile but that are, arguably, not environmental-specific kinds of WB. For example, when I was young, I got similar thrills from going fast in cars.

²⁴ Quoting Firth, “A meaningful relationship occurs when the interactions between two entities have significance in their past history and its anticipated continuation.” Examples include our historical relationships, as humans, with nature, animal parent/child relationships, historical relationships between land and human communities, the relationships between humans and the plants and animals they eat, the relationship between a rancher and his herd,

[You might find nothing unique in the types of sensations, perceptions, thoughts, and emotions nature can produce, but that does not prove others don't.]

Ecological economists value ecological systems (ES), and ES are prime suspects when it comes to incommensurabilities

Ecological systems (ES) such as wetlands, forests, coral reefs, estuaries, bays, and rivers integrate biological, hydrological, physical, and human processes providing services to plants, animals, and humans. An ES can be viewed as a bearer of WB or as a vector of bearers of WB. These bearers are inputs into the production of recreation, animal and plant preservation, ways-of-life, water purification, and flood control. And they produce overlapping vectors of different kinds of WB. For example, a reef fishery might maintain for an indigenous population the WB associated with the continuation of one's way-of-life, recreational benefits for the non-indigenous, and knowledge of preservation for distant populations, all while making coastline residents worry less about gw sea rise.²⁵

One has to question whether these different kinds of WB are commensurable and with the kinds of WB produced by other goods and activities. There are many articles about ES in the ecological literature, and many of these include variations on the word "incommensurable". They take incommensurability as a given. Paraphrasing [Martinez-Alier](#), Munda, and [O'Neill](#) (1998), WB-incomparabilities and WB-incommensurabilities are the "foundation of ecological

the relationship between two wolves in a pack, and the relationship between a wolf pack and its potential prey. This last example is discussed in detail by Aldo Leopold (1949/87). In summary, an understanding and appreciation of nature-based relationships is an important component of a worthwhile life, according this perspective.

²⁵ These kinds of WB effects vary drastically in their magnitudes: the loss of a sport-fishing opportunity has only a small effect on our WB, whereas the loss of one's culture, community, and way-of-life are life shattering. Cynthia Burke (2010) reports on how the loss of subsistence fishing activities has affected the way-of-life of the Nuxall First Nation of British Columbia. Quoting from [Chan](#), [Satterfield](#) and [Goldstein](#) (2014), "A Kyuquot–Checleset elder (of the northwest coast of Vancouver Island, B.C.), described (pers. comm.) the loss of fishing opportunities as causing a loss of knowledge and cultural identity in the community's youth, which she seemed to attribute to a lack of transformative experiences, all of which were entangled with both self- and other-oriented, group and individual values... a Kyuquot–Checleset fisherman (pers. comm.) suggested the decline of local Chinook ... as triggering loss of inspiration and spiritual benefits..."

economics”.²⁶²⁷ Typically what ecological economists try to value is has more interconnected bearers than what environmental economists value: most of my (neoclassical) valuation studies have been limited to a component of an ES such as a catch rate, a fish-consumption advisory, or a ski or mountain-bike trail.

While this aside has been about the environment, my applied field in economics, similar arguments could be made by those who study bearers of WB such as families, public institutions, and political processes. For example, childcare is produced by families, firms, and the government, and each produces different kinds of WB and ill-being. Parents worry about whether the kinds produced inside and outside the family are child WB-commensurable. Many kinds of WB produced by the criminal/justice system are also produced by the market (e.g., a sense of safety from a lock or alarm system), but many aren't, such as the sense that justice was served and your rights are protected.

If you have no control over how much of a bearer you experience, you don't compare it with other bearers^{vii}

²⁶ This article has been cited a thousand times by ecologists and ecological economists— environmental economists don't cite it. The article starts from the premise that environmental amenities are WB-incomparable with market goods—the authors don't provide arguments for their premise. The literature citing this article typically cites it as proof of the premise. Unlike the philosophical literature on incommensurability, this literature does not adopt the distinction between kinds of WB and bearers of WB, which is unfortunate.

Martinez-Alier is a past president of the [International Society of Ecological Economics](#). I find “foundation of ecological economics” telling: the title is asserting that all that separates ecological economist from neoclassical environmental economists is ecological economists believe the kinds of WB produced by environmental resources are incommensurable with the kinds produced by market goods. [K. William Kapp](#) (1983) might be the first modern economist to state this position. He said, referring to the kinds of WB produced by environmental resources, “they are heterogenous and cannot be compared quantitatively among themselves and with each other, not even in principle.” Three publications that cite Martinez-Alier, Munda and O'Neill and assume incomplete WB-comparability are Spash (2008), O'Neill, Holland and [Light](#) (2008) and Munda (2016). Many citing articles are in *Ecological Economics* or *Environmental Valuation*.

²⁷ Because of incommensurabilities, ecological economists choose environmental projects based on multiple criteria: “Incommensurability means that there is not a common unit of measurement, but it does *not* mean that we cannot compare alternative decisions on a rational basis, on *different* scales of values, as in multi-criteria evaluation.” (Martinez-Alier (1995). Another example is [Martin-Lopez](#), [Gomez-Baggethun](#), [Garcia-Llorent](#), and [Montes](#) (2014). If for an individual WB-comparability is not complete, his ordering of paths will be incomplete, so, for example, he might rank environmental projects *A* and *B* both higher than the status-quo but *A* and *B* are not ranked relative to each other, so additional criteria are needed for him to decide between them. That is, multiple-criteria are often needed to decide between environmental projects when the kinds of WB produced by each is incommensurable with the kinds produced by the other and incommensurable with the kinds produced by market goods. For example, one might select *B* over *A* if *B* helps poor people more. There is much debate as to what the additional criteria(s) should be. Besides Martinez-Alier (1995) and Martinez-Alier, Munda, and O'Neill (1998), see, for example, Munda (2004), and [Stabell](#) (forthcoming)

Put simply, don't sweat what you cannot influence. [Elizabeth Anderson](#) (1997) articulates this argument. If you cannot control a bearer's level, there is no reason to think about how to compare it with another bearer. And besides it being a waste of your time, it would, for many, be something to avoid. Economists, in contrast, assume that when you wake up in the morning, you have a complete ordering over all conceivable paths, so they reject that the two paths are first compared only when the individual has to choose between them. Economists, like me, who value in dollars environmental resources, assume you have a well-defined *CV* for less global warming, even if you cannot affect its rate. You have it even before the surveyor asks you what it is. Economists of my ilk would say that a contingent-valuation survey is simply a way for the researcher to find out your *CV* to reduce global warming (less global warming is a bearer).

In contrast, Anderson would say that you don't show up at the survey center with a *CV* for panda preservation because it never crossed your mind that you would ever be comparing panda preservation with beer and chocolate cake. And, you won't during the survey unless you are convinced by the survey that panda preservation is something you can influence. Those of us in the actual business of creating and using contingent-valuation surveys to value environmental resources have sympathy for her point, but not enough to reject the WB-comparability of market goods with environmental commodities whose levels you have no control over.

Historically, most of us could not control pollution levels or amounts of environmental preservation. But this is starting to change, at least probabilistically, particularly if one is extremely rich and/or politically active. [Jeff Bezos recently pledged \$10 billion to address climate change.]

Evolution and WB comparison

The ability to WB-compare is a skill that one inherits, learns, or both. In evolutionary terms, the survival of humans and prehumans had nothing to do with the sorts of health, environmental and long-term choices that we face now. Cancer risks vs. diet and exercise were not on the table. Evolution was not driven by benefits now vs. benefits decades in the future. As a result, humans did not evolve to WB-compare many modern-world alternatives. We see this also in our inability to WB-compare when small probabilities are in play.

Inadequate processing skills

Imagine I have complete information about two paths in that I know the specifics of each, including probabilities, so my ignorance (lack of knowledge) is not an issue. But the paths have more moving parts than I can cognitively process, making it impossible for me to WB-order them. I think this is a possible cause of WB-incommensurability. Still, I don't want to make too much of it because this inability causes more than just WB-incommensurability issues for choice theory.

You might want to counter that this is an example of inability to compare because of ignorance and not a real example in WB-incommensurability. It is ignorance in the sense that I'm ignorant of the WB effects of each path, but I'm not ignorant of the properties of each path, so it isn't ignorance in the lack of knowledge sense.

The lack of processing skills argument is if too many kinds of WB differ between the two paths, one's ability to WB-compare them breaks down, even if one could WB-rank the two paths if fewer kinds of WB were changing.

Saying "Paths *i* and *j* are WB-incomparable" does not make it so^{viii}

Earlier, I noted that asserting that, for you, *i* and *j* are WB-comparable does make it so. Asserting the opposite ("I can't WB-compare *i* and *j*") also does make it so (Posner 1998). It is a flawed argument either way. There are lots of reasons for my saying, "I can't—even if I can. We might not want to put in the cognitive effort; we might not want to convey that we are the sort who can compare the WB from saving polar bears with the WB from beer, even if we are. Admitting to your spouse that their love is WB-commensurable with the pleasure from extra-marital sex would reduce their love for you, so you lie.

Path incomparability is sometimes mandated by law^{ix}

The U.S. *Endangered Species Act* says species aren't comparable: extingting one to save three is illegal. The Delaney Clause forbids carcinogenic substances, meaning cancers are not comparable with market goods (Sunstein 1994, p.835). And, the U.S. Federal Courts have ruled that even if dollar damage from degrading an environmental resource is less than the cost of restoring it, the destroyer is responsible for the higher restoration cost. Suggesting that, under the law, environmental damages are not comparable with market goods.

The issue for WB-comparability is whether a legal mandate that things are not comparable makes it more difficult for an individual to WB-compare them.

A few other qualms about WB-commensurability:^x

- (I) A lot of what we consume are *public commodities*: commodities that are non-congestible and nonexcludable). Some increase WB (*public goods*), and others decrease WB (*public bads*). This makes comparing the WB produced by public commodities with those produced by market goods awkward: how do you compare consuming a good that will only increase your WB with consuming a good that will increase everyone's WB? Or compare your consumption of a market good with the provision of a public commodity that would increase your WB but decreases everyone else's. People who lack empathy for others make such comparisons, but that does not mean everyone can. Not everyone is capable of comparing the personal WB they would get from private goods with the WB losses everyone would get from dirtier air or more global warming—they feel they don't have the right to make such comparisons, so they don't.
- (II) Imagine comparing the WB relief from less global warming with being less anxious. Making such comparisons is complicated because being more anxious will influence how much relief you get from a given change in the rate-of-global-warming. [One could test this in the lab by invoking anxiety before the contingent-valuation question or choice experiment. The design would vary the amounts of both the hypothetical changes in the rates of global warming and the amounts of anxiety invoked in the subjects.²⁸]
- (III) Personal responsibilities and personal commitments can complicate WB-commensurability. [Contrast personal responsibilities and obligations with those placed on everyone by religious or cultural mores.] Consider a western rancher whose ranch has been in the family for generations. Typically, ranches provide a wildlife habitat, a bearer of environmental WB for both the rancher and others.²⁹ Many ranchers are committed to and feel responsible for maintaining the ranch in its current state and keeping it in the family (both of these responsibilities motivate conservation easements). It wouldn't be surprising if ranchers would have difficulty comparing the

²⁸ Anxiety levels can be manipulated with, for example, computer simulations and by varying the environment. Listening to this 911 call will make you anxious (<http://clipp.instruct.de/player/data/db/video/106672.mov>); it is disturbing. [Rajagopal Raghunathan](#) and Michel Pham (1999) is an example of a study that evoked anxiety. For examples of modeling choices as a function of personality and emotions see Solinõ and Farizo (2014) and Morey and [Mara Thiene](#) (2017).

²⁹ Ranches, near where I lived north of Steamboat CO, besides grazing cattle and sheep, provide habitat for elk, deer, moose, mountain lions, trout, and many other animals. There are conservation easements on many properties, even though obtaining the easement reduce a ranch's market value.

WB associated with maintaining the ranch habitat with the WB obtained by retiring to a condo in Florida.

One final qualm about a world of complete WB-commensurability and complete comparability: most of the richness and variety of life would be beside the point. At the end of the day, no one would care what caused their WB. Of course, many people would hope they don't live in such a world, but this does not prove they do or don't.

Ok, so maybe some people cannot WB-rank all conceivable paths, but who cares if most people can WB-rank their current available paths.

Choice does not require complete WB-comparability, only the ability to WB-compare the available paths. If one could demonstrate that most people can do this, the cost of ignoring WB-incomparability would be negligible. However, I don't know how one would determine, case by case, what proportion of the population could WB-compare a specific pair of paths. [And, as I have said before, you saying you can does not mean you can.]

Neurological evidence in support of comparability and WB-commensurability^{xi}

Recent findings on the neurobiology of choice oppose the above philosophical arguments against complete WB-commensurability. These findings are consistent with complete comparability (don't contradict it). Quoting [Dino Levy](#) and [Paul Glimcher](#) (2012),

Indeed, there is now broad consensus in the neuroscience of the decision-making community that reward magnitude is represented in a small number of well-identified areas. Here we conduct a meta-analysis using evidence from human functional magnetic resonance imaging (fMRI) studies conducted over just the past few years that suggest that one of these reward magnitude encoding areas, the ventromedial prefrontal cortex/orbital frontal cortex (vmPFC/OFC), can be thought of as representing the value of nearly all reward-types on a common scale that predicts behaviorally observed comparison and choice.

They claim that the the research demonstrates that we all have complete WB-comparability. But they claim too much.

The striatum and the ventromedial prefrontal cortex together are your *valuation circuit* (Glimcher 2014). Picture a two-dimensional topographical map of neurons, where each alternative in the current choice set is represented by a different point on the neuron map (Figure 2).³⁰ The third dimension is the firing rates of the neurons. The alternative whose neuron

³⁰ "... most classes of information recorded in the cerebral cortex are topographically encoded on anatomically two-dimensional 'maps.'" "The cortex is made of dozens of these small topographical maps" (Glimcher (2014)). The valuation map has only the options on the table.

achieves the highest peak (firing rate) will be selected. Initially, the firing rates for the different alternatives fluctuate. If a neuron's firing rate increases, it increases the firing rates of nearby neurons while inhibiting the firing rates of distant neurons, including those associated with the other peaks. Eventually, one peak dominates, and you go with that alternative.

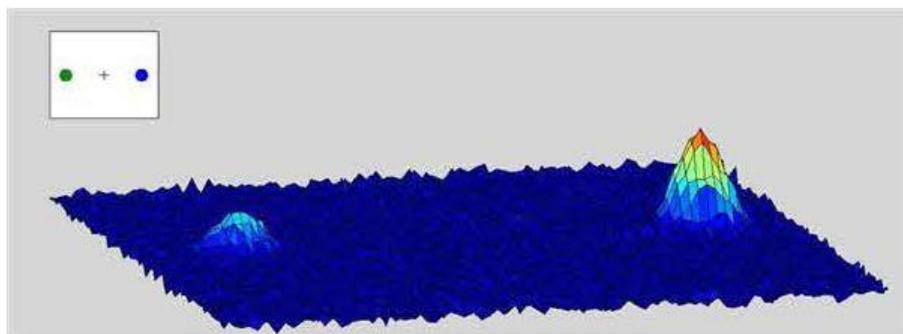


Figure 2, Chapter 8: Cartoon of the valuation circuit with two alternatives (courtesy of Ryan Webb)

(I) The final selection of an alternative (at least for the sorts of sets of alternatives studied in neuroscience labs) always takes place in the valuation circuit. (II) The variation in firing rates across the neurons in this area determines/predicts which alternative will be selected. Every alternative is compared on only one dimension (firing rate). This is all consistent with complete WB-comparability.

In the studies Levy and Glimcher review, male subjects were asked to choose between alternatives or simply view alternatives, all while fMRI measured firing rates. Alternatives were presented with different amounts of the same reward, different reward types, and both different types and magnitudes. Alternatives included money (magnitude and when it would be delivered), college trinkets, pain, pictures of females that varied in attractiveness, and snack foods. In all the studies that involved choosing—not all did—money was one alternative.

No matter what alternatives were presented, the valuation circuit was activated, suggesting the selection process always includes the valuation circuit. And the firing-rate findings are consistent with which alternative the individual says they would prefer (see, in particular, Smith *et al.* (2010) and Levy and Glimcher (2011 and 12)).

However, these findings do not imply that relative firing rates reflect how each alternative would be ranked by a well-defined monotonic index of overall WB. They don't even imply that a WB-index exists. So, they do not imply complete WB-comparability. The authors use the word "value" but simply define it in terms of firing rates: the alternative with the highest firing rate is defined as the one with the highest value, so the rest have lower values. It's simply a measure of electrical activity in a particular place in your head.

These findings are consistent with all bundles being comparable in terms of something different from WB, even something inconsistent with WB. For example, they are consistent with selecting the alternative you most desire, or the one that makes you least anxious, or most proud, or even different criteria on different selection occasions.³¹ an

Recollect the incorrect circular argument. First, you assume a complete WB ordering of alternatives and that the individual chooses their highest-ranked available path. Then you observe the individual selecting the alternative with the highest firing rate. It, by assumption, is the one with the highest value. The argument does not demonstrate complete WB-comparability: you started by assuming it.

To summarize, the neurological evidence is consistent with WB commensurability and bearer-of-WB comparability but doesn't prove either.

Those who reject complete WB-commensurability wouldn't be surprised by the neurological findings nor disagree with them. Instead, they would say that comparability between money and snacks doesn't imply complete comparability.

Levy and Glimcher were looking for a spot in the brain where WB-commensurability occurs, and they found a candidate,

Summing up and looking forward

A foundation of NBT is complete WB-comparability of paths, which requires complete WB-commensurability. This Chapter reviewed the issues and arguments and presented relevant theory and research.

While NCT was formulated initially to only explain and estimate the demand functions for market goods, its scope has been leaping and creeping wider—which is why I define an

³¹ See, Chapter 4, for more on desires/wants vs. likes

ordering over paths rather than only bundles of market goods. Economists working in labor, urban, public, development, and family economics have widened the span. It now includes bearers-of-WB such as school quality, public services, roads, air-pollution levels, water quality, parks, wilderness, rate of gw, crime rates, job security, human rights, number of bathrooms, commuting time, and national defense, plus a lot more.

When NCT was only about market goods, arguably, advocates only needed to worry about the comparability of bundles of market goods. And this only required commensurability of the kinds of WB produced by market goods. Now there are many more path components and many more kinds of WB to model.

Pushing back against the neoclassical assumption that everything is WB-comparable are ethicists, and many others, who believe many things we care about cannot be compared with market goods.

If there is not complete WB-commensurability, there will not be a complete ordering of paths. And this has implications. Predictions become less specific: all that can be said is that the individual will not experience a path ranked lower than another available path (maximizing rather than optimizing). Adding to this, I have shown that if there is not complete WB-comparability, for many path shifts (Path i to j), there is no compensating variation, CV, corresponding to that shift—the CV concept is not defined. For example, a CV for a policy that reduced the rate of gw does not exist if any changes in any kinds of WB (or ill-being) affected by the policy are incommensurable with each other or with anything else whose level the individual can influence. Or if it is associated with changes in path components that he cares about but can't influence. To make these issues go away, one has to believe that the kinds of WB produced by national defense, friends and family, lovers, freedoms, fairness, and many other things are all WB-commensurable.

My intent in this chapter was not to tell you that there is or isn't complete WB-comparability. My goal was to express the implications of incomplete WB-comparability and outline arguments one must defend against if one wants to argue for complete WB-comparability. Of course, if you only have to deal with people who are OK with complete WB-comparability, your road ahead is smooth. [I wonder why the issue of WB-incomparability has not been raised by lawyers and economists representing the defendant in NRDA litigations (e.g.,

the BP spill in the gulf). I suspect it is because all the economic-valuation experts on both sides are neoclassical economists, so they accept, without thought, complete WB-comparability.^{32]}

Economists can reject all the arguments against complete WB-comparability and complete WB-commensurability, but they should be able to articulate why these arguments are wrong. Also, economists should know the neurological evidence.

This brings us to the end of Part I. Part I presented NCT and presented evidence and arguments supporting and contradicting its assumptions. Rather than calling it neoclassical choice theory, NCT, I called my presentation neoclassical behavior theory, NBT. I did this for two reasons, I did not want to quibble about what exactly is and isn't NCT, and I wanted to draw a vital distinction between behaviors and chosen behaviors (choices). Economists believe most behaviors are chosen, therefore the adjective "choice" in choice theory. In contrast, many philosophers and psychologists argue, with evidence, that behaviors are often not chosen.

Bringing us to Part II: What is a choice? Choice or the illusion of choice? Should we care? It starts, Chapter 9, with trying to identify what makes a choice a choice.

³² If I had been working for British Petroleum on the legal case to assess the dollar damages from the 2010 oil spill in the Gulf of Mexico [I worked for the other side], and if my intent had been to reduce, or eliminate, their liability, I would have introduced the argument that environmental kinds of WB are, for many, incommensurable with the kinds of WB generated by spending on private goods, making it impossible for many to trade-off environmental injuries for money. Arguing this would, at a minimum, have forced both sides to hire some philosophers and neuroscientists to defend, or attack, WB- commensurability. Philosophers could use the extra income.

ⁱ Sources:

1. Matthew Adler 1998. "Incommensurability and cost-benefit analysis". *University of Pennsylvania Law Review*. 146(5): 1371-1418.
2. Wilfred Beckerman 2017. *Economics as applied ethics: Fact and Value in Economic Policy, 2nd Edition*. Palgrave MacMillan.
3. Walter Bossert 2018. "Suzumura-consistent relations: and overview". *International Journal of Economic Theory*.
4. Walter Bossert and Kotaro Suzumura 2010. *Consistency, choice, and rationality*. Harvard University Press, Cambridge, MA, 2010.
5. John Broome 2001. "Are intentions reasons? And how should we cope with incommensurable values". In *Practical Rationality and Preference: Essays for David Gauthier*. C. Morris and A. Ripstein (eds.). Cambridge: Cambridge University Press. pp.98-120.
6. Ruth Chang 1997. "Introduction". In *Incommensurability, Incomparability, and Practical Reason*. Ruth Chang (ed.). Cambridge: Harvard University Press.
7. James Griffin 1986. *Well-being*. Oxford: Oxford University Press
8. _____ 1997. Incommensurability: what's the problem? In *Incommensurability, Incomparability, and Practical Reason*. Ruth Chang (ed.). Cambridge: Harvard University Press.
9. _____ 2000. "Replies". In: *Well-being and morality: Essays in honor of James Griffin*. R. Crisp and B. Hooker (eds.). Oxford: Clarendon Press.
10. Sven Hansson 2018. "Money-Pumps". In Sven Hansson and V. Hendricks(eds) *Introduction to Formal Philosophy*: Springer Undergraduate Texts in Philosophy. Springer, Cham. https://doi.org/10.1007/978-3-319-77434-3_31
11. Fredrick Hayek 1935. "The nature and history of the problem" and "The present state of the debate". In *Collectivist Economic Planning*. F. Hayek (ed.). Routledge and Kegan Paul LTD, London.
12. Nien-he Hsieh 2007. "Is incomparability a problem for anyone?". *Economics and Philosophy*. 23: 65-80.
13. _____ Spring 2016 Edition. "Incommensurable values". *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). Kahneman **Error! Hyperlink reference not valid.**
14. John Locke 1700. 4th ed. "Book II: Chapter XX: On modes of pleasure and pain," In *An essay on human understanding*. Oxford: Oxford University Press, 1975.
15. Michael Mandler 2004. "Status quo maintenance reconsidered: changing or incomplete preferences". *The Economic Journal* 114.
16. Otto Neurath (1928). "Personal life and class struggle". In Neurath (1973) *Empiricism and Sociology*. Reidel, Dordrecht.
17. Donald Regan 1997. "Value, Comparability, and Choice," in *Incommensurability, Incomparability, and Practical Reason*. Ruth Chang (ed.), Cambridge: Harvard University Press.
18. Arthur Schopenhauer (1818/2011) The Project Gutenberg EBook of *The World As Will And Idea* (Vol. 1 of 3). www.gutenberg.org/files/38427/38427-pdf.pdf
19. Mark Schroeder Summer 2012. "Value theory". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). **Error! Hyperlink reference not valid.**
20. Amartya Sen 1997. "Maximization and the act of choice". *Econometrica*. 65: 745–79.
21. _____ 2018. "The importance of incompleteness". *International Journal of Economic Theory*. 14(1).
22. Patricia Sheridan Summer 2016. "Locke's Moral Philosophy". *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.).

-
23. Herbert Simon 1982. *Models of bounded rationality Vols. 1 and 2*. New York, Wiley.
 24. Cass Sunstein 1994. "Incommensurability and valuation in law". *Michigan Law Review*. 92(4): 779-861.
 25. Kotaro Suzumura 1983. *Rational choice, collective decisions, and social welfare*. Cambridge, MA. Harvard University Press.
 26. Ludwig von Mises (1920;1935). "Economic calculation in the socialist commonwealth". In *Collectivist Economic Planning*. F. Hayek (ed.). Routledge and Kegan Paul LTD, London.
 27. D. Wiggins 1997. "Incommensurability: four proposals". In: *Incommensurability, Incomparability, and Practical Reason*. Ruth Chang (ed.). Cambridge: Harvard University Press.

ⁱⁱ Sources:

1. Martha Nussbaum 2012. "Who is the happy warrior? Philosophy, happiness research, and public policy". *International Review of Economics*. 59:335–361.

ⁱⁱⁱ Sources:

2. Edward Morey, Jennifer Thacher and Edward Craighead (2007). "Patient preferences for depression treatment programs and willingness to pay for treatment," *Journal of Mental Health Policy and Economics*, Vol. 10 (2), 87-99.
3. Joseph Raz 1986. *The Morality of Freedom*. Oxford: Clarendon Press.
4. Sunstein, *ibid*.
5. P. Tetlock (2003). "Thinking the unthinkable: sacred values and taboo cognitions". *Trends in Cognitive Sciences*. 7(7).
6. Jennifer Thacher, Edward Morey, and Edward Craighead (2005) "Using Patient Characteristics and Attitudinal Data to Identify Depression Treatment Preference Groups: A Latent-Class Model". *Depression and Anxiety*, Vol. 21(2): 47-54.

^{iv} Sources:

1. Andrea Scarantino and Ronald de Sousa (Winter 2018) "Emotion". *Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/entries/emotion/>
2. David Eli and Justin Rao 2011. "The good-news-bad news effect: asymmetric processing of objective information about yourself". *American Economic Journal: Microeconomics*. 3(2):114-118.
3. A. Ganguly and J. Tasoff. 2016. "Fantasy and dread: the demand for information and the consumption utility of the future". *Management Science*. 63(12).
4. Sunstein, *ibid*.

^v Sources:

1. Rachel Carson 1962. *Silent Spring*. Houghton Mifflin Harcourt.
2. Jesse Graham, Jonathan Haidt and Brian Nosek. 2009. "Liberals and conservatives rely on different sets of moral foundations". *Personality Process and Individual Differences*. 96(5): 1029-1046.
3. Jonathan Haidt and Jesse Graham 2007. "When morality opposes justice: conservatives have moral intuitions liberals may not recognize". *Social Justice Research*. 20:98-116.
4. Leopold, A. 1949/1987. *A sand county almanac: and sketches here and there*. New York: Oxford University Press.

^{vi} Sources:

1. Cynthia Burke (2010) "When the Fishing's Gone: Understanding How Fisheries Management Affects the Informal Economy and Social Capital in the Nuxalk Nation, Resource Management and Environmental Studies. University of British Columbia, Vancouver. M.A. Thesis

-
2. Chan, K., T. Satterfield, and J. Goldstein. (2014) “Rethinking ecosystem services to better address and navigate cultural values”. *Ecological Economics*. 74:8-18.
 3. Dan Firth (2008). “Do meaningful relationships with nature contribute to a worthwhile life?”. *Environmental Values*. 17(2).
 4. Alan Holland A. (2002). “Are choices tradeoffs”. Chapter 2 in *Contested Choices*. ed. D. Bromley and J Paavola. Oxford: Oxford University Press.
 5. A. Holland (2006). “Must we give up environmental ethics?” in ten Have, H. (ed.), *Environmental Ethics and International Policy* (Paris: UNESCO).
 6. K. William Kapp (1983). “Social costs, economic development and environmental disruption”. Introduction by J.E. Ullmann, University Press of America, Lanham, MD.
 7. Joan Martinez-Alier (May 1, 1995). “Political ecology, distributional conflicts, and economic incommensurability”. *New Left Review*.
 8. Joan Martinez-Alier, Giuseppe Munda, and John O’Neill. 1998. “Weak comparability of values as a foundation for ecological economics”. *Ecological Economics*. 26: 277-286.
 9. Martín-López, B., E. Eric Gómez-Baggethun, M. García-Llorente and C. Montes. (2014). Trade-offs across value- domains in ecosystem services assessment. *Ecological Indicators*. 37:220-28. Munda, G. (2004). “Social multi-criteria evaluation: methodological foundations of operational consequences.” *European Journal of Operations Research*. 158:662-677.5-202
 10. Munda, G. (2016). “Beyond welfare economics: some methodological issues”. *Journal of Economic Methodology*. 23(2):185-202.
 11. O’Neill, J., A. Holland, and A. Light (2008). *Environmental Values*. London, Routledge.
 12. Spash, C.L. (2008). “How much is that ecosystem in the window? The one with the bio-diverse trail”. *Environmental Values*. 17(2):259-84.
 13. Stabell, E. (forthcoming). “Hard environmental choices: comparability, justification, and the argument for moral identity”. *Environmental Values*.

vii Sources:

1. Elizabeth Anderson 1997. “Practical reason and incommensurable goods.” In: *Incommensurability, Incomparability, and Practical Reason*. Ruth Chang (ed.), Cambridge: Harvard University Press.

viii Sources:

1. Richard Posner (1983) *The Economics of Justice*. Harvard, Harvard University Press

ix Sources:

1. Sunstein, *ibid*.

x Source:

1. Edward Morey. 2020. “Can economic choose? Can homo sapiens? Discussion paper.
2. Edward Morey and Mara Thiene. 2017. “Can personality traits explain where and with whom you recreate? A latent-class site-choice model informed by estimates from mixed-model LC cluster models with latent-personality traits”. *Ecological Economics*. 138: 223-37.
3. Raghunathan, R. and M. Pham. 1999. “All negative moods aren’t equal: motivational influences of anxiety and sadness on decision making”. *Organizational Behavior and Human Decision Processes*. 79(1): 56-77.
4. Solinõ, M. and B.A. Farizo. 2014. “Personality traits underlying environmental preference: a discrete choice experiment”. *PLoS ONE*. 9(2).

xi Sources:

-
1. Dino Levy and Paul Glimcher. 2011. "Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain". *Journal of Neuroscience*. 31:14693-14707.
 2. _____ 2012. "The root of all value: a neural common currency for choice". *Current Opinion in Neurobiology*. 22:1027-38.
 3. Paul Glimcher 2014. "Chapter 20: Value-based decision making". In *Neuroeconomics Second Edition: Decision Making and the Brain*. P. Glimcher and E. Fehr (Eds.). Elsevier.
 4. D.V. Smith, B.Y. Hayden T.K. Truong, A.W. Song, M.L. Platt, and S.A. Huettel. 2010. "Distinct value signals in anterior and posterior ventromedial prefrontal cortex". *Journal of Neuroscience*. 30: 2490-2495.

Part II: What is a choice? Choice or the illusion of choice?

08172022

Economists assert that *economici* and *homo sapiens* make choices without considering the distinction between behavior and choice (a chosen behavior). While the neoclassical theory of behavior is called choice theory, an *economicus* (an entity that adheres to the assumptions) cannot control its behavior: its ranking of paths and which are available are both exogenous, so is the rule that it must experience its highest-ranked available path, HRAP. An *economicus* must behave according to these constraints. This is why I call it neoclassical behavior theory (NBT) rather than neoclassical choice theory (NCT). *Economici* don't make choices in the word's street sense.

Choice and free will are almost synonyms: economists use the word “choice”; philosophers use “free will”. Part II reports on both *economici* and *homo sapiens*. Looking ahead, an *economicus* has only feeble free will, so chooses only in a feeble sense. For example, its behavior is more deterministic than determinism is defined in physics and philosophy. This makes advocates of NBT a subspecies of what philosophers call *compatibilists*: behavior is “compatible” with free will, but only feeble free will. Humans also live in a deterministic world (or with a bit of randomness). So, unless you believe in dualism, humans only have feeble free will, no more than worms and mollusks—but maybe more than *economici*.

But! Humans have choosing experiences (cogitating about A vs. B, and deciding on B) and then experiencing B. Isn't that choosing? Probably not! Mounting evidence in psychology and neuroscience indicates that what humans do (how we behave) is determined by an unconscious process, and the choosing experience does not determine what we select.

Of course, whether a behavior is a chosen behavior comes down to how one defines *chosen*: one could define it so that every act is chosen, making every behavior a chosen behavior. Defining chosen in this way is not uncommon—but it is weird.

Chapter 9: How would you define a *choice*? And the difference between a choice and the experience/sense of making a choice

03102022

A choice is difficult to define, particularly if you feel compelled to not use the word "choice" to define a choice.

Necessary conditions for a situation to be a choice?

(1) An entity is faced with N options (an integer $N \geq 2$). (2) The entity can and must experience only one. (3) It can influence which one is experienced. And (4), It is not constrained to experience a specific one. Many would elaborate on (4), adding, "There was only a choice if the entity was free to experience an option different from the one experienced." While most people would not, unprompted, list these requirements, they would likely agree with them upon reflection. Still, not everyone would (including neoclassical economists), particularly if we get more specific about what (3) and (4) mean.

Note that these conditions do not imply or require that the entity have a conscious choice *experience* (experiencing conscious thoughts about which alternative to choose).

Imagine, or not, that these 4 conditions are together sufficient for a situation to be a choice.

But! Does choice require a choosing experience? Have you made a choice only if you feel you made a choice?

And, if so, what brings about the feeling that you influenced which alternative was realized. You experience the need to make a choice, then consciously decide on an alternative, and then experience that alternative, giving you the sense that you made a choice. But can you make a choice without consciously feeling that you have made a choice? And, can you not make a choice but think you have? The evidence on these two questions illuminates why we perceive some of our actions as choices and others not.

Contrary to what economists believe and teach, the assumptions of NBT seem to make choice impossible?

Economists assert that economic agents make choices

But we appear to contradict ourselves when we assert this. In explanation, since an economic agent is constrained to consume its HRAP, since the set of available paths is exogenous, and since the

ranking of paths is exogenous, there seems to be no real choice.¹ Behavior is axiomatically determined:

Assumption 4: An economicus's ordering does not change in the time-span behavior is modeled.

Assumption 5: While there are many paths, most paths are not available, and an economicus can only experience an available path.

Assumption 7: At every point in time, economicus takes one of its HRAPs

Most people would say economici don't make choices. So, when economists say economici make choices, they must mean something different from what most think of as a choice.

Of course, whether there is choice in NBT comes down to how choice is defined; but if you were constrained to select the path you experience, the guy-on-the-street would say there was no choice. His definition—you could have selected a different alternative—is the folk definition of choice (what most people mean by choice). So, when economists say that people make choices, they must mean something different from what the guy-on-the-street thinks of as a choice.

Economists need to define choice in a way that makes a choice compatible with their theory of behavior (or drop the word “choice”)

Economists would say an economicus has no choice if there is only one available option (N=1), implying more than one option is a necessary condition of choice. But is it sufficient? That is, does an entity have an economic choice if **external** constraints don't limit it to a specific path? Economists must think so because it is called “CT” even though internal constraints (an exogenous ranking and the requirement to experience the HRAP) constrain it to experience a specific path.

Looking ahead, economists define choice as situations where external constraints do not limit the entity to a specific path.

This is similar to the definition compatibilist philosophers use to make choice (free will in a weak sense) possible in a deterministic world. By this definition, other animals (even worms and mollusks) make choices whenever their behavior is not entirely determined by external

¹ Maybe in the past economicus could influence its future ranking, but it cannot influence its current ranking.

constraints. For example, Giacomo always eats the meat first when both dry food and fresh meat are in his bowl. And most nights, he sleeps in two or three places, but no puppeteer is dragging him on a leash from room to room. None of his behavior is dictated by external constraints, so he exhibits weak free will.

So, what is *free will*, and how does it relate to a capacity to make choices?ⁱ

Isabel Archer is on her way to Italy to pursue her destiny and explains to her suitor, Caspar Goodwood, her aversion to his suiting.

Isabel: *If there's a thing in the world I'm fond of it's my personal independence... [I want] To put as many hundred miles of sea between us as possible.*

Casper: *One would think you were going to commit some atrocity!*

Isabel: *Perhaps I am. I wish to be free even to do that if the fancy takes me.* (The Portrait of a lady, [Henry James 1917](#))

Isabel wants to be free to choose to commit one or more atrocities, but only if she fancies to. Is Isabel free to will and choose?

I have used *free will* without defining it, which has been convenient given the disagreement about what it means. The *Stanford Encyclopedia of Philosophy*, *SEP*, defines free will as a “capacity to choose a course of action from among various alternatives.” N.B. the words “capacity” and “choice”. If you have it, spirits or networks of neurons give you the capacity to choose.²

Consider the word “will” and what it means to say, “I will that Event *A* happens.” [This is an archaic use and makes me think of God’s will.] This could be interpreted as I choose Event *A* over the other alternatives. Or, if *A* is a behavior, I am inclined to adopt it. It could also be interpreted as I prefer *A*, and because I prefer it, I choose it. *Choosing*, *willing*, and *preferring* are close but imperfect synonyms. Put simply, free will is philosophical-speak for *the capacity to freely choose*, and the capacity to freely choose is economic-speak for free will.

But what does *free* mean? Kant thought that if an entity must be rational, and rationality determines what it will do, it does not have free will. Kant would agree with my argument for

² See also *Free will: an introduction* by [Helen Beebe](#) (2013). My interest in free will is what it tells us about economic choice while she is concerned with its relationship to moral responsibility. We cover a lot of the same ground and discuss some of the same studies in neuroscience and experimental philosophy.

why NBT is incompatible with choice. He would argue that an entity would have free will only if it could choose whether to follow the assumptions of NBT, but not if they are an immutable part of its nature. They are an immutable part of economic nature. If you adopt the HRAP rule but do not have to, you have free will. For Kant, an economicus does not have free will because it cannot reject the rule.

Doubts about free will go way backⁱⁱ

Men are mistaken in thinking themselves free; their opinion is made up of consciousness of their own actions, and ignorance of the causes by which they are conditioned. Their idea of freedom, therefore, is simply their ignorance of any cause for their actions. As for saying that human actions depend on the will, this is a mere phrase without any idea to correspond thereto. (Spinoza, *The Ethics Part II: Of the Nature and Origin of the Mind* (partially quoted by Wegner))

All theory is against the freedom of the will; all experience is for it. (James Boswell 1791, *The Life of Samuel Johnson* (as quoted by Wegner)³)

Spinoza, [Thomas Hobbs](#) (1588-1679), and Locke rejected free will. The philosopher [Jennifer Uleman](#) nicely summarizes the different takes on free will. Quoting her, [Hobbs] "denies that will can be free at all, [Hobbs] arguing that only bodies can be free, or unfree, where free [bodies] just means unimpeded." [*Unimpeded* means unconstrained, as in rich peoples' bodies are less impeded than those of poor people, and prisoners and those with physical disabilities are impeded more than many.]

Locke "concluded the human will is never free but is always determined by nature or reason: for Locke, 'free will' is a nonsensical thing".

In contrast, Descartes understood free will as "a mental ability to endorse and set oneself on a course of action (or more simply, to assent to something, or not), rather than as a function of the grounds determining action or assent" (Uleman). You can will (choose) whatever you want, but that does not mean you determined what you would will (would choose) or that what you willed (what you chose) will occur. In his view, no one can stop me from willing A to happen, but that I willed A rather than B was not determined by me. Descartes' free will is limited, at best.

³ Boswell is best known for his biography of his friend, the English man of letters, Samuel Johnson; it is considered the first great biography in English. Boswell knew and admired Rousseau; he had a brief affair with Rousseau's partner—one of Boswell's many dalliances.

Kant said a will is not free if its possessor must behave rationally, as must economici. But Kant does not require that humans behave rationally. Kant makes a distinction between the will and whether the will is free. For Kant, the *will* (not necessarily free) is a real thing that resides in each adult human. It is the place or process in a human mind that chooses an alternative and then causes the human to do what it takes to experience that alternative; it is the "choosing center"—willing is choosing. Quoting Uleman,

A will is free, for Kant, if it determines itself and isn't determined by anything else. A will will be free, in other words, if it chooses ends, and pursues courses of action aimed at realizing those ends, on grounds that are its own, and not on grounds given to it by something or someone external to it.

For Kant, Kantian free will is a possibility—but only for humans—and it's not guaranteed.⁴

Modern definitions of free will fall along a continuumⁱⁱⁱ

One wants to be what tradition has it that Eve was when she bit the apple. Perfectly free to do otherwise. So perfectly free, in fact, that even God couldn't tell which way she'd jump. (Jerry Fodor (2003) philosopher and cognitive scientist))

Fodor is providing one definition, not arguing for it. This is *extreme free will*. Economicus doesn't have it.

At the continuum's other end, an entity has free will if another agent is not controlling it, and other agents, excluding God, cannot predict with certainty what it will do before it does it. You only lack it if a mad scientist, puppeteer, or God pulls your strings. Economicus has only *vacuous free will*: while its behavior can be predicted, no puppeteer is pulling its strings.

In between these two extremes, but closer to extreme free will, is the ability to have chosen differently. You chose *A* but could have chosen *B* (you could have behaved differently than you behaved). And you have this ability even if God—if you believe in God—knew which alternative you were going to choose. This is *strong free will*. Economici do not have it: in economic-speak, an economicus does not choose in this sense. Looking ahead, most homo sapiens believe they have strong free will.

⁴ Kant relied on *dualism* to argue that will can be free (Uleman). His argument is awkward. Kant viewed the body as a physical thing controlled by the laws of physics and believed that if what you will is determined by the laws of physics—what you will isn't freely willed. He also believed that if what you will is constrained by reason, you do not freely will. These two beliefs of Kant argue against the existence of a will that is free. But these arguments are materialist. Kant got his will freed by imagining that will isn't something subject to reason or the laws of physics; it is a magical thing. An analogy is God's Will which isn't subject to anything earthly; otherwise, it would not be God's Will.

Also in between the two ends, but closer to vacuous free will, is *weak free will*. You lack it only if your behavior is wholly determined by external constraints: you must experience an alternative in your choice set, and it includes only one ($N=1$). So, the will is free if $N \geq 2$. For example, people in prison cannot choose to spend the night in their cell.⁵ Economicus has weak free will: in economic-speak, choice, in this weak sense, exists whenever $N \geq 2$. Every economicus is constrained to do whatever it did, but two of its constraints (his ordering of paths and having to select his HRAP) are internal.

I have defined four variants of free will. Economicus only has the weak form: it cannot behave differently. The widely-published philosopher [Daniel Dennett](#) defines it as what I have called weak free will. He argues that strong free will (so also extreme free will) is “bonkers” given the laws of physics.⁶

People care deeply about having free will, but they also seem to have misguided ideas about what free will is or could be ...Our decisions are not little miracles in the brain that violate the physics and chemistry that account for the rest of our bodies' processes, even if many folk think this must be what happens if our decisions are to be truly free. We can't conclude from this, however, that then we don't have free will, because free will in this bonkers sense is not the only concept of free will.

Weak free will is close to the legal definition, as in you signed the contract of your "own free will" if you were not "under duress or the influence of hallucination or another mental derangement." He claims it

... is probably the consensus [view of free will] not only among philosophers but also among judges, lawyers, and others who have to make distinctions about who is responsible for what and who is excused because they do not free will when they acted.

Fodor's assessment of Dennett's definition of free will is

There's the lurking sense that what you got isn't quite what you ordered, and half an hour later you're hungry again.

⁵ I wonder if advocates of this view of free will (in this weak sense) believe rich people tend to have more of it than poor people because the rich are often less constrained.

⁶ This is also how Hume defined free will. He starts by noting that everything has a cause (an earthquake has a cause, and you buying vanilla ice cream instead of strawberry has a cause), so if free will is when behavior is not caused, it is impossible. Rejecting free will in this sense, he says you have it when your behavior is not “compelled”, defining compelled behavior as behavior wholly determined by exogenous forces.

In Fodor's view, Dennett has simply defined all behavior as chosen behavior—a bait and switch because what Dennett calls free will, most people would not. Dennett would say economici exhibit free will; Foder is still hungry.

Causal determinism^{iv}

The hypothesis is your behavior is wholly determined by the laws of nature and what has come before (the state of the world, including your genetics and history)—your actions are caused. It is taken as fact by many (but not all) scientists and philosophers. But rejected out of hand by most people, especially if anyone suggests it applies to humans.

Thinking about causal determinism and its implications for NBT informs what it means to make a choice and whether people are responsible for their choices (both critical ethical issues). However, be warned that thinking about causal determinism and choice can be crazy-making, and most of us do not like determinism nor its implications.

With full information and the ability to process it, I could predict everything you will do—all acts are causal and deterministic.⁷⁸ It follows from classical, deterministic physics. The opposite is *indeterminism*. Most neuroscientists and physicists who are indeterminists are so only because they believe behavior has a random component. The rest are determinists. Modern physicists accept quantum mechanics, which includes randomness at the subatomic level, but no one knows whether it affects the behaviors of individuals.

Historically, many philosophers were determinists, including Locke, Spinoza, [Gottfried Leibniz](#) (1646-1716), and Hume. So was the mathematician [Pierre-Simon Laplace](#) (1749-1872). In 1814, Laplace articulated determinism as an intellect/demon who could predict all future acts.

In Laplace's story, a sufficiently bright demon who knew how things stood in the world 100 years before my birth could predict every action, every emotion, every belief in the course of my life. Were she then to watch me live through it, she might smile condescendingly, as one who watches a marionette dance to the tugs of strings that it knows nothing about. We can't stand the thought that we are marionettes. Nor does it matter whether any demon (or even God) can, or cares to, actually predict what we will do: the existence of the strings of physical necessity, linked to far-past states of the world and determining our current every move, is what alarms us. (SEP on Causal Determinism)

⁷Causal determinism is different from fate and predestination. Both, like causal determinism, imply something is going to happen, but the reason it is going to happen is magical: inconsistent with physics.

⁸ This does not imply everything can be predicted: currently there is not enough computing power, and many believe there never could be.

Notable historical indeterminists include [George Berkeley](#) (1685-1753), Descartes, Kant, and Rousseau.⁹ For Berkeley, Descartes, and Kant, their arguments for indeterminism were all based on their belief in dualism (your mind is more than your brain: magic)—Rousseau’s argument is not.

Economicus’s behavior is deterministic and is consistent with a deterministic world. But only one where the configuration of atoms at the start of time would result in economici. Human behavior is also consistent with a deterministic world, but, unlike economici, there is no requirement that the configuration of atoms at the start of time was the one that would make humans economici.

Humans tend to believe causal determinism does not apply to humans

William James, the father of American psychology, desperately wanted to not believe in casual determinism.

But the whole feeling of reality, the whole sting and excitement of our voluntary life, depends on our sense that in it things are really being decided from one moment to another, and that it is not the dull rattling off of a chain that was forged innumerable ages ago. (James, 1890, quoted by Wegner)

But he admitted it might be.¹⁰

Until recently, the conjecture was untested. Quoting the experimental philosopher [Shaun Nichols](#) (2004):

In a set of experiments exploring the lay understanding of choice, both children and adults tended to treat moral choices as indeterminant. Participants were presented with cases of moral choice events (e.g., a girl steals a candy bar) and physical events (e.g., a pot of water comes to a boil), and they were asked whether, if everything in the world was the same right up until the event occurred, the event had to occur. Both children and adults were more likely to say that the physical event had to occur than that the moral choice event had to occur. This result seems to

⁹ Berkeley is “the great eccentric of Anglophone philosophy, who infamously argued...that the notion of matter is nonsensical, and that”, except for God, “minds and their perceptions are the sole contents of the universe,” (Gottlieb 2016). A modern counterpoint is the cognitive psychologist [Donald Hoffman](#) who argue that while there is a reality, our perceptions (what we see, hear, and feel) are mere illusions: our minds and sensory organs evolved to foster the successful transmission of genes, not to correctly represent reality (Hoffman 2019).

¹⁰ “As a young man he [James] passed through a profound and prolonged crisis, mental or emotional or spiritual, insofar as such distinctions can be thought of as meaningful to him. In retrospect, he laid his despair to his loss of belief in freedom of the will. His depression was disabling to him physically, and the cures he sought out in Europe did nothing to relieve it. He struggled with thoughts of suicide. Then he read a book by the French philosopher [Charles Bernard Renouvier](#), who argued that one was made free by acting as if he were free. So began his convalescence ...” (the novelist and essayist [Marilynne Robinson](#) 2010). That book argued that free will was the ability to influence one’s thoughts: James concluded that his believing he had this ability was not necessarily an illusion, which, for him, opened the door to possible choice (James 1890, Henley 2019).

vindicate the traditional claim that ordinary people in our culture believe that at least some human decisions are not determined.

Why? Determinism conflicts with our sense of self—it is diminished if you think your behavior is set. We have two views of how the world works: a mechanical view that applies to the behaviors of rocks, dogs, and all other non-human plants and animals, and a magical view that applies to people—your mind is more than your brain—the dualist view. Rejecting determinism makes it easier to believe humans consciously set their course. People readily believe the behavior of lesser animals is deterministic (or random), believing lesser animals are driven by instinct and stimuli, so they do not make choices in the sense people do.

So, is choice consistent with causal determinism? Compatibilists and Incompatibilists'

There are numerous schools of thought, varying on whether you believe in determinism and whether choice is consistent with it. Behavior having a random component is a way to generate indeterminism, but adding a random component to an entity's behavior does not mean it is making choices—choosing freely isn't the same thing as behaving randomly.

Most neuroscientists and physicists are incompatibilists, rejecting free will and choice

As previously mentioned, they are either determinists or indeterminists because they believe behavior has a random component. And most don't think free will is consistent with either if it is the ability to behave differently. [Dennett, a champion of compatibilism, lists the following as distinguished members of the *choice-is-illusion* camp (incompatibilists): the physicists Albert Einstein and [Stephen Hawking](#) (1942-2018), the neuroscientists [Wolf Singer](#), [Chris Frith](#), and [Patrick Haggard](#), and the psychologists [Paul Bloom](#) and [Daniel Wegner](#) (1948-2013)]

What are philosophers? Many are deterministic compatibilists

According to the compatibilist Dennett and the incompatibilist Nichols, most philosophers are compatibilists, including the determinists Hobbes, Locke, Leibniz, and Hume. Hume even

argued determinism is required for free will.¹¹ Most compatibilists are determinists, but a few determinist philosophers are not compatibilists.¹²

Why are there deterministic compatibilists? Cynically, according to Nichols (2007),

Many of us incompatibilists think we know the answer to this: it's wishful thinking! Philosophers embrace compatibilism because they want it to be true. This view is, I think, common among incompatibilists. Famously, [William] James dubs compatibilism a "quagmire of evasion." Even more famously, Kant says it is a "wretched subterfuge." We can put the incompatibilists' motivational hypothesis somewhat more precisely as follows: philosophers embrace compatibilism despite its counter-intuitiveness because compatibilism is motivationally attractive.

In summary, many philosophers believe in determinism, and most believe determinism is compatible with free will, given that they define it in the weak sense.¹³ Economists are implicit and unaware *compatibilists*.

Do regular people think choice is consistent with causal determinism?

Nichols and [Joshua Knobe](#) (2007) started by asking college students to imagine two different universes.

¹¹ Hume's analogy: imagine a slope with rocks lined up in a row from the top to the bottom, with the bottom rock's shape and size representing you and what is in your head (your internal constraints). Determinism is a causal chain of events starting with the top rock rolling into the second, causing it to move and roll into the third, causing it to, until a rock rolls into you. If the size and speed of rock that hits you is sufficient to fully determine where you will end up, your behavior is compelled, and you did not choose where to roll. However, if which way you roll is determined, in part, by your shape and size, you (the rock) have free will and some choice as to where to roll. If the path taken is not wholly determined by exogenous constraints—there is more than one externally-available path, you have it and made a choice. Modern deterministic compatibilists, like Dennett, would agree.

¹² These include Spinoza, and the enlightenment thinkers [Paul-Henri D'Holbach](#) (1723-1789), [Denis Diderot](#) (1713-84), and [Voltaire](#) (1694-1778). There are a few indeterminant philosophers who argue free will could exist even if the world were deterministic.

¹³ An exception is the economist/philosopher List (2014) who argues that determinism is compatible with strong free-will. List's argument turns on how one interprets "for someone's action to count as free is that the agent can do otherwise." He provides three possible interpretations from the philosophical literature: (1) "If the agent were to try (or choose) to do otherwise, he or she would succeed in doing otherwise", (2) "The agent has the disposition to do otherwise when, in appropriate circumstances (to be spelt out further), he or she tries to do otherwise, and (3) "It is possible (in a sense to be spelt out further) for the agent to do otherwise." He then argues that (1) and (2) are consistent with determinism. While I am somewhat confused, (1) seems consistent with determinism in that while you must choose as you choose given the moment's determinants, if you were to choose to do otherwise (because the world was different deterministically) you would be successful. To buy the argument, you have to buy either (1) or (2) as the correct definition of being able to do otherwise. (because the world was different deterministically) you would be successful.

Imagine a universe (Universe A) in which everything that happens is completely caused by what happened before it. This is true from the very beginning of the universe, so what happened in the beginning of the universe caused what happened next, and so on right up until the present...

Now imagine a universe (Universe B) in which almost everything that happens is completely caused by whatever happened before it. The one exception is human decision making...

Ninety percent of the respondents thought our world was more like B, consistent with Nichols' (2004) finding: determinism except for humans.

Everyone was then asked either:

In Universe A, is it possible for a person to be fully morally responsible for their actions? Yes or No?

Or

In Universe A, a man named Bill has become attracted to his secretary, and he decides that the only way to be with her is to kill his wife and 3 children. He knows that it is impossible to escape from his house in the event of a fire. Before he leaves on a business trip, he sets up a device in his basement that burns down the house and kills his family. Is Bill fully morally responsible for killing his wife and children? Yes or No?

86% answered "No" (not responsible) to the first question, indicating that most respondents thought the person could not have acted differently. But only 25% answered "No" (not responsible) to the second question, so most people believed Bill could have chosen to not kill the family.

The second question can't determine whether the difference in "No" responses is because Universe A is abstract while B is concrete or because B has a visceral effect that A does not. So, they did a second experiment. Again, there were two questions: half of the people who got each question were told to answer assuming Universe A, half Universe B.

As he has done many times in the past, Mark arranges to cheat on his taxes. Is it possible that Mark is fully responsible for cheating on his taxes? Yes or No?

Or

As he has done many times in the past, Mark stalks and rapes a stranger. Is it possible that Mark is fully responsible for raping the stranger? Yes or No?

Both actions are concretely described, but the second generates a negative visceral response. For those told to assume Universe B (the indeterminant universe), most concluded that both the rapist and tax cheater could have acted differently, so responsible. Those instructed to assume Universe A (the determinant case) believed the tax evader had to cheat, but the rapist did not have to rape.

To summarize, humans believe a deterministic act is consistent with free will if it is explicitly harmful, is concretely described, was performed by a human, and makes them viscerally upset. In which case, they attribute moral responsibility to the actor. But they believe a determinist act is inconsistent with free will when the natures of the act and actor are unspecified.¹⁴ This result has been replicated in non-western populations. We are of two minds regarding whether free will is consistent with determinism: a logical mind that concludes the deterministic actor is not responsible and an emotional mind that often concludes they are.

Why do we all feel and believe that we have free will and make choices?

While we understand that some of our actions are taken unconsciously—most of us feel that most of what we do happens because we consciously and freely decide to behave that way. We believe we have free will—that we *will—freely*—most of our actions.

Two reasons to believe we have free will and make choices:

Reason 1: because we do have free will and do make choices. We believe this is the correct answer. We consciously choose ham and eggs for breakfast rather than dry granola, which is why we are consuming animal products rather than plants. But—another possible reason (Reason 2) is evolution has kluged us: left us with a flawed notion of causality, causing us to incorrectly believe that the conscious experience of making a choice is what led us to take a specific action.

A bit on causality and logic

Consider two events (II and III) where Event III always happens soon after Event II, and Event III never occurs unless Event II precedes it. We conclude that Event II caused Event III because often it has—evolution has caused us to think this way; its objective is not to produce perfectly rational and logical animals, only animals that get it right enough of the time. The sun rises in New York City (II) before it rises in Chicago (III), but that does mean the sunrise in NYC caused the sun to rise in Chicago. They are both caused by Event I, the sun's and earth's motions.

The illogic is *necessary* is confused with *sufficient*. If Event III occurs only when Event II occurs first, then Event II is necessary for Event III (without II, there can be no III). But Event II

¹⁴ “Indeed, concrete cases of bad behavior lead people to attribute responsibility, even when the action is caused by a neurological disorder [like a brain tumor]” (Nichols 2011)

being necessary does not mean it is sufficient to cause Event III. [If it were sufficient, then the occurrence of Event II would guarantee that III happens.] Two examples: a necessary condition for being Bob the Bear is that you are a bear, but being a bear isn't sufficient to make you Bob the Bear. And the sun rising in NYC is necessary for it to rise in Chicago, but it does guarantee it.

Events II and III—and choice

The two events of interest are *having the conscious experience of deciding to do something* (Event II) and *doing what was consciously decided* (Event III). Examples of Event II are me consciously choosing to drink a Coke and Melvin consciously deciding to propose to Wilma. Event III is the actual drinking or proposing. That Event II precedes III does not mean it caused III. Melvin only proposed to Wanda after he had the experience of deciding to propose, but maybe what caused him to propose was a third event in their past (an event involving a slinky red dress) he doesn't consciously remember. Or perhaps he forgot it was Wanda's idea.

The possibility that Event II did not cause Event III is a hard thing to get your head around when it seems obvious that it did, and you want to believe it did. You want to believe that first consciously deciding to buy a new car was what caused you to buy it.

Another category (type IV Events) is behaviors not preceded by a choosing experience. We are all aware that many/most of our actions are not preceded by a conscious decision to take action, so-called automatic behaviors/involuntary actions, so we know action does not require a prior choosing experience. Examples of *automatics* include the different components of physical activities such as walking, skiing, driving, and breathing—the mechanics are mostly produced unconsciously. [Imagine if to breathe, you had to always consciously inhale and exhale.¹⁵] Other examples are habits, acts driven by emotions, and instinctive reactions. For example, if the spinal cord receives data that there will be bodily injury if immediate action is not taken, it will send instructions to the muscles to immediately act. This happens before you are consciously aware of any danger¹⁶—only after you dive out of the way do you consciously register what happened. There is an evolutionary advantage to foregoing a choosing experience when delaying

¹⁵ Conscious breathing is the norm in endurance races, meditation, and scuba diving.

¹⁶ Stepping on a tack fires pain neurons in your foot; the message is transmitted to your spinal cord; from there one message goes to your brain and the other heads straight back down a motor neuron to your foot, causing a muscle contraction before you register the event. Bear et al. have a nice diagram.

would get you run over. But danger is not required. You are more likely to win the tennis match if you don't consciously choose to hit every ball flying your way. If an attractive other smiles as you pass on the street, you either react or lose the opportunity. But circumstances that require a quick response are not the only situations where we forego the choosing experience, more on this in Chapter 10.

Another category, a Type I event, is an unconscious neurological event before Events II and III, where II is a conscious choosing experience, and III is acquiring B rather than A. If Event I causes III, Event II is simply window dressing, the *illusion of choice* (an epiphenomenon) like the whistle on a steam locomotive that blows before the train crosses the road but does not cause the train to cross the road.

More and more, the evidence suggests most of our behaviors are determined in the unconscious parts of our brain and our conscious brain, later, but before the behavior occurs, sometimes has a choosing experience. Our conscious brain tricks us into thinking it determines what we do. Note that this does not mean conscious thought plays no role: conscious thoughts can affect the unconscious, so influence future behaviors. Chapter 10 presents the evidence. [If we were created in God's image, we would unlikely entertain such an illusion—God hopefully does not suffer from it.]

ⁱ Sources:

1. Henry James (1917) *Portrait of a Lady* (Vol. 1, Chapter XVI) A Project Gutenberg EBook <http://www.gutenberg.org/files/2833/2833-h/2833-h.htm>
2. Helen Beebe (2013) *Free Will: An Introduction*. Palgrave Macmillan.
3. Timothy O'Connor (2014) *SEP on Free Will*. <http://plato.stanford.edu/entries/freewill/>
4. Michael Rohlf (2020) *SEP on Immanuel Kant*.
5. Jennifer Uleman (2010) *An Introduction to Kant's Moral Philosophy*, Cambridge UK, Cambridge University Press.

ⁱⁱ Sources:

1. (Baruch) Benedict de Spinoza (2001) *Ethics Part II: Of the Nature and Origin of the Mind*, Blackmask Online (translated from the Latin by R.H.M Elwes) http://www2.hn.psu.edu/faculty/jmanis/spinoza/spinoza_ethics_2.pdf
2. James Boswell (1791) *Life of Johnson*, Book 3 https://books.google.com/books/about/The_Life_of_Samuel_Johnson.html?id=Q0FXAAAAYAAJ
3. Daniel Wegner (2002) *The Illusion of Conscious Will*, Cambridge MA. MIT Press
4. Jennifer Uleman (2010) *ibid*.

ⁱⁱⁱ Sources:

1. Jerry Fodor (2003) Why Would Mother Nature Bother? (Review of *Freedom Evolves* by Daniel Dennett). *London Rev. of Books* 25(5): 17-18. At <http://www.lrb.co.uk/v25/n05/jerry-fodor/why-would-mother-nature-bother>

-
2. Daniel Dennett (2013) *Intuition Pumps and Other Tools for Thinking*, Kindle Edition, N.Y., W. W. Norton & Company.

^{iv} Sources:

1. [Randolph Clarke](http://plato.stanford.edu/entries/incompatibilism-theories/) and [Justin Capes](#) (2014) SEP on *Incompatibilist (Nondeterministic) Theories of Free Will*. <http://plato.stanford.edu/entries/incompatibilism-theories/>
2. [Lisa Downing](#) (Spring 2020) SEP on *George Berkeley*
3. Anthony Gottlieb (2016) *The dream of enlightenment: the rise of modern philosophy*, N.Y. and London, W.W. Norton and Company.
4. [Carl Hoefer](#) (2010) SEP on *Causal Determinism*. <http://plato.stanford.edu/entries/determinism-causal/>
5. Donald Hoffman (2019) *The case against reality: why evolution hid the truth from our eyes*, N.Y. and London, W.W. Norton and Company.
6. Shaun Nichols (2004) The Folk Psychology of Free Will: Fits and Starts, *Mind & Language* 19:473-502. <https://sites.google.com/view/shaunbnichols/publications/publications-by-date>

^v Sources:

1. Daniel Dennett (2013) *Intuition Pumps and Other Tools for Thinking*, Kindle Edition, N.Y., W. W. Norton & Company
2. Christian List (2014) Free will, determinism, and the possibility of doing otherwise. *Noûs* 48 (1): pp. 156-178.
3. Michael McKenna and Justin Coates (2015) SEP on *Compatibilism*. <http://plato.stanford.edu/entries/compatibilism/>
4. Shaun, Nichols (2007) The Rise of Compatibilism: A Case Study in the Quantitative History of Philosophy, *Midwest Studies in Phil.* 31: 260-270. <https://sites.google.com/view/shaunbnichols/publications/publications-by-date>
5. Shaun Nichols (2011) Experimental philosophy and the problem of free will, *Sci.* 331: 1401-03.
6. Shaun Nichols and Joshua Knobe (2007) Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions, *Nous* 41: 663-685.
7. Shaun Nichols and Joshua Knobe (2007) *ibid.*

Chapter 10: The evidence on conscious choice

08172022

The neurological evidence:¹

Starting with the 1983 [Benjamin Libet](#) experiment, an increasing number of neurological studies indicate that before you consciously decide what to do, neural activity can be observed that predicts what you will do—brain activity you are unaware of.¹ Event I (the neural activity) precedes Event III. Event I is necessary and sufficient to cause the choosing experience (Event II). Event II occurs after Event I but before Event III. Event I is a *readiness potential*: your unconscious is getting ready to cause the voluntary act.²³ Paraphrasing Wegner (2002), choosing is an experience, not a cause.⁴

The Libet experiment



Figure 1, Chapter 10: Benjamin Libet -1916-1997

¹ Neural activity is observed by observing electrical activity; initially, this was accomplished by attaching electrodes to the scalp. More recently it was accomplished with fMRI and now, sometimes, by directly implanting tiny electrodes in the brain that isolate on specific groups of neurons. Implants are uncommon: when it's done, it is during surgery to treat severe epilepsy.

² Quoting Libet, in 1965, Kornhuber and Deecke “found that a recordable electrical charge in brain activity regularly and specifically preceded a voluntary act... a slow rise in electrical negativity ... The electrical charge started about 800 msec. or more before a subject performed an apparently voluntary act... the readiness potential (RP) or in German, the “bereitschaftspotential”.

³ A readiness potential does not precede every action; it only precedes voluntary actions. It does not precede reflex actions, or other uncontrolled actions. For example, when people with Tourette’s syndrome involuntarily swear, the swearing is not preceded by a readiness potential, but their voluntary actions are preceded by readiness potentials. So, a readiness potential is not necessary for an action to occur because all actions are not preceded by readiness potentials, but they are necessary and sufficient for voluntary actions.

⁴ “...conscious will is an experience, not a cause.”

Imagine you are a subject in the following experiment. You enter a small room and sit at a desk, comfortably resting your left arm on the desk. On the wall is a big 60-second stopwatch with an illuminated dot traveling around its circumference. The stopwatch has the seconds marked in five-second intervals (0, 5 sec. 10 sec., etc. up to 55 sec.). You notice it moves much faster than an ordinary stopwatch. While you don't know its exact speed, the dot covers 60 seconds in 2.6 standard seconds. It takes the dot .43 seconds (430 msec.) to go from 0 to 10 and 1.3 seconds (1300 msec.) to go from 0 to 30. So, if you tell someone where the dot was when an event occurs (e.g., when your hand is touched), they can determine when you think the event happened.⁵

A research assistant attaches two sets of electrodes. The first set to your left wrist. It will record if and when you flick your wrist (when your wrist-flexing muscle contracts). The second set is attached to your scalp, recording if and when a certain negatively charged pulse occurs at the top of your brain. You are told to sit there for a few minutes and, if you want, sometimes flick your wrist. Whether and how often you flick is up to you. You are asked to remember the dot's position if and when you decide to flick.

This is a stylized description of the famous experiment conducted in 1983 at the University of San Francisco. The results: Approximately a half second (500 msec.) before your wrist muscle contracts, the electrode on your scalp records the start of a negatively charged pulse (a *readiness potential*). This neural activity, which you are unaware of, precedes and predicts you will flick your wrist a half second before you flick it. Your unconscious is getting ready to cause your wrist muscle to contract. [Note the experiment does not demonstrate that the RP causes your wrist to flick—an even earlier neurological event could have caused the RP—more on that later]

The issue is, when do you have the conscious experience of choosing to flick your wrist? Based on where the dot was when you decided to flick, this happens, on average, approximately 200 msec. before the muscle contracts, indicating the flick is initiated by your unconscious at least 300 msec. before you consciously “decide” to flick. WOW, but with caution.

Interpret the Libet result carefully. It indicates that your unconscious determined (chose?) when you would flick your wrist and that your conscious experience of choosing to flick is not what caused it.

Questions ever since the experiment include: (1) Does it indeed indicate your unconscious caused your wrist to flick and that your conscious decision was not the cause? It seems to. (2) If yes, does the result hold for more complex decisions (e.g., proposing to Wilma)? Complex decisions are hard to test (see [Lady Ottoline Morrell](#), discussed below). If consciously

⁵ Testing dot time: dot time for an unseen stimulus to a subject's hand (a light touch) is only about 50 msec. after the touch, so the sensory lag is small, measurable, and can be accounted for.

choosing an action is not what caused it, this violates *Assumption 12b: That perception of choosing is what causes an economicus to take the path it took.*

Many neuroscientists and philosophers have questioned whether the experiment demonstrates what it suggests. Reasonable answers support the conclusions, but these answers are not bombproof. That said, the fundamental finding that an act can be predicted based on unconscious brain activity has been replicated multiple times. In 2007, [John-Dylan Haynes](#), a neuroscientist in Berlin, found, using fMRI imaging, the researcher could predict, with 60% accuracy, whether you will push the left or right button as much as 7 seconds before you push it, and 6 seconds before you “chose” which button to push. Quoting [Kerri Smith](#), the Podcast Editor for the journal *Nature*, talking about a 2011 study:

Some researchers have literally gone deeper into the brain. One of those is [Itzhak Fried](#), a neuroscientist and surgeon at the University of California, Los Angeles, and the Tel Aviv Medical Center in Israel. He studied individuals with electrodes implanted in their brains as part of a surgical procedure to treat epilepsy. Recording from single neurons in this way gives scientists a much more precise picture of brain activity than fMRI or EEG. Fried's experiments showed that there was activity in individual neurons of particular brain areas about a second and a half before the subject made a conscious decision to press a button. With about 700 milliseconds to go, the researchers could predict the timing of that decision with more than 80% accuracy. ‘At some point, things that are predetermined are admitted into consciousness,’ says Fried. He suggests that the conscious will might be added to a decision at a later stage.

What are the concerns? Issue 1: while the Libet results are consistent with Event I (the RP) being both necessary and sufficient for Event III (the voluntary act),⁶ the Libet results are consistent with Event I being necessary but not sufficient for Event III. Or so argues [Aaron Schurger](#) and his coauthors in two separate papers (2012 and 15). They hypothesize an alternative explanation for the Libet finding. They distinguish between *intent* and *commitment*, arguing that the onset of the RP is not signaling a commitment to act but instead signals the start of an intent (an intent to flick the wrist or push the button).⁷ [Keep in mind that an RP is simply a slow build-up of neurological activity (electrical activity on the scalp or the firing rates of specific neurons)-a neurological pulse/wave.] Schurger and his coauthors assume intent is necessary for commitment but not sufficient. According to their hypothesis, commitment occurs

⁶ Libet's interpretation.

⁷ “Commitment” is the point in time when the mind is made up. As in, I intend to learn Latin, but I have not committed to learning Latin.

when the RP wave reaches a neurological threshold. If correct, an RP will occur some percentage of the time without the act occurring, but the Libet experiment was not designed to identify this percentage. Everyone eventually flicked their wrist in the Libet experiment, making it impossible to observe a wave that did not end with a flick. This, according to Schurger, caused Libet and others to mislabel *necessary* as *sufficient* because the studies limit their data collection to time periods that always end with the act occurring.

The Schurger hypothesis is consistent with the commitment occurring before or after Event II (the choosing experience). If it is after the choosing experience, it is inconsistent with the Libet interpretation that unconscious commitment always precedes the choosing experience. That said, the Schurger hypothesis and results don't demonstrate that the choosing experience caused anything.

Issue 2 (clock time): In real life, you are rarely asked when you decided on A rather than B: you don't have to tell someone when you had the choosing experience. But in research studies, respondents must indicate to the researcher when they consciously decided. In the Libet experiment, they had to specify the clock time. An issue is whether the requirement to monitor the clock affects the timing and amplitude of the RP wave.⁸ In a 2011 experiment ([Jeff Miller](#), [Peter Shepherdson](#), and [Judy Trevena](#)) compared the wave pattern of the electrical activity with and without clock monitoring. While the two graphs have similar shapes, clock monitoring shifted it downward (more negative than without the monitoring). Put simply, the addition of clock monitoring amplifies the RP. Miller and his co-authors interpret their findings to imply that maybe the RP is not sufficient to determine how you will behave. I don't understand. With clock monitoring, which is required to determine when the choosing experience occurs, it becomes difficult to determine which part of the wave is caused by the intent/commitment to act and which part is the result of monitoring the clock.

On a broader level of criticism, there is [Ben Newell](#) and [David Shanks'](#) (2014) review of “unconscious influences on decision making.” They conclude

⁸ As an aside, in experiments where respondents had to estimate clock time, some always report their decision occurring after they flicked or pressed a key. “This [a mistake?] seems to stem at least in part from difficulties in determining the spot's location at the decision time (Miller et al.). More interestingly, it could indicate that the choosing experience actually occurs after the individual has acted.

Our critical analysis points to a surprising conclusion, that there is little convincing evidence of unconscious influences on decision making in the areas we review, and that, as a consequence, such influences should not be assigned a prominent role in theories of decision making and related behaviors. This conclusion is consistent with the view that conscious thoughts are by far the primary driver of behavior.

To be clear, they are rejecting as unconvincing the entire literature on unconscious influences on behavior, not only the Libet-type studies. They also reject all the priming and cueing studies that I will discuss. Their paper was published in *Behavioral and Brain Sciences*, along with 28 commentaries, a few supportive, many not. Either way, the review makes important points, many around the awareness (consciously aware), measures of awareness, and determining when cognition becomes conscious. Their primary conclusion is that the literature demonstrating unconscious influences on behavior has not sufficiently demonstrated the absence of awareness. [The Libet experiment assumes lack of awareness until the indicated decision time.] An essential point in the commentaries is what is the null hypothesis. Is behavior unconsciously determined or consciously determined? If the null is behavior is unconsciously determined, to reject it, one has must demonstrate awareness. If it is the latter, one must establish a lack of awareness. Their review starts with the null that behavior is consciously determined and then argues that research has not proven a lack of awareness.

Reading Newell and Shanks got me thinking about *unawareness* and how to measure it, defining unawareness as unconscious cognition (unconscious thoughts). Unawareness is when you think about something but are not aware that you are.

Brain imaging has now been used to identify unconscious cognition ([John Creswell](#), [James Bursley](#), and [Ajay Satpute](#) 2014). They presented 21 subjects with a complex decision task while all were in an fMRI machine. Each subject spent 84 seconds (the processing phase) seeing how four different cars (or backpacks) varied by 16 attributes (e.g., good or bad gas mileage). Subjects were told they would quality-rate the four products. The four products were configured to vary objectively in quality (price was not an attribute). There were three treatments. After the processing phase, subjects in the first treatment immediately quality-rated each product.

There was a 120-second intermediate phase between the processing phase and the decision phase in the other two. In the second treatment, the individual was asked to think about how they would rate the four products, so 120 seconds of conscious thought. In the third

treatment, subjects spent the 120 seconds playing a game.⁹ The game grabs all of your attention, making it impossible to consciously think about the alternatives. So if there is thinking about them, it is unconscious. For the third treatment, subjects played the game a second time after rating the four products. Remember, brain activity is being continuously recorded. What was happening unconsciously during the intermediate phase with the 2-back treatment was obtained by subtracting the second game images from the first set. The researchers removed from the movie the part attributed to conscious cognition, leaving a movie of unconscious thought.

As has been previously shown in prior studies without the fMRI, the performance on the rating task was best after the game and worse when they consciously thought about the alternatives. This indicates that useful cogitating occurred unconsciously during the game. Notably, the brain regions activated unconsciously were the same ones activated during the processing phase. The studies' results are consistent with the hypothesis that the final decision is made by the unconscious. The study was the first to map unconscious thought with neuron activation.

My summary of the above neurological evidence: Most everyone agrees that the RP is necessary for a voluntary movement; whether it is sufficient is more contested. When you become aware of your commitment to act is debated. There are two issues: measurement of the awareness and whether the start of the RP indicates commitment (whether RP is sufficient). Suppose the time at which awareness occurs is measured without error, and the RP demonstrates unconscious commitment. In that case, awareness of commitment comes after the commitment.

Alternatively, if one assumes that the RP is necessary but not sufficient for commitment (the threshold model), then awareness might be concurrent with commitment (when the threshold is reached).

Both are consistent with the unconscious choice, so neither assumption implies that the choosing experience plays a role in what follows.

Additional support for the behavioral role of the unconscious is the finding that your brain can both formulate a goal (e.g., get you invited to a party) and cause you to take actions to

⁹ A 2-back n-back game: a number appeared on the screen for .5 seconds, followed by an * for 2.5 seconds, then a new number. The subject had to click a button whenever the current number was the same as the second-back number.

achieve that goal without being aware of it why you are taking these actions. Distinguishing between goals and actions/behavior, a goal often motivates actions (choosing to lose weight is a goal, skipping dessert is an action). In a 2010 *Science* article, [Ruud Custers](#) and [Henk Aarts](#) interpret this recent research on unconscious goal-formation. The previous view, and what most of us want to believe, is that goals are consciously chosen. For example, my goal to write this book must have been consciously adopted. But maybe not.

To summarize the research, you can be subliminally motivated to adopt a goal and then unconsciously take actions to achieve it. In the lab, the goal is brought to the attention of your unconscious subliminally or, outside the lab, by environmental cues that you are not aware of.¹⁰ You then unconsciously evaluate the rewards associated with that goal (in the lab with a subliminal reward cue, or because you already unconsciously associate a reward with it). Then, depending on how your unconscious assesses the magnitude of the associated reward, you adopt the goal and take actions to achieve it.¹¹

Despite such findings, most people believe consciously choosing causes them to perform an action or adopt a goal. While most people believe Assumption 12b is true, there is a good chance it is false.¹² NBT can get along without Assumption 12b, economicus just becomes less of a me—recollect that basic NBT is only Assumptions 1-9.

So, does the Libet and related findings imply that conscious thought, including the choosing experience, does not affect behavior?

No! Even if your unconscious is the determining step before an action is taken, it leaves open the possibility that earlier (right before, or much earlier) conscious thinking influenced what your

¹⁰ People speak more softly when seeing a picture of a library, are more likely to clean their table if there is vague whiff of cleaner in the air, and get more competitive when they enter an office if there is a leather briefcase on the desk (Henk Aarts and Ap Dijksterhuis (2003), [Rob Holland](#) et al. (2005), and [Kay](#) et al. (2004))

¹¹ In one study they review, students were seated *in front of a computer, allegedly to test their computer mouse skills. Before starting on this test, some participants were subliminally exposed to words on the computer screen related to the goal of socializing, whereas others were exposed to unrelated words. At the onset of the mouse-skill test, they were told that if there would be enough time left after the test, they could engage in a lottery in which they could win tickets to a popular student party. Thus, spending more effort (by working faster) on the mouse-skill test was instrumental in attaining the goal to socialize. The participants worked harder on the mouse-skill test when the socializing goal was first unconsciously primed. And this effect was stronger when socializing evoked a stronger positive reward signal in the minds of the participants (which was assessed in a separate implicit affective association task). Importantly, checks indicated that priming caused participants to pursue the goal independently of their reported motivation to attain it.*

¹² Interestingly, Libet, in late life, raised the possibility that after the RP and the conscious deciding occurs, you still might be able to consciously veto the action (stop yourself) which would be *free won't* rather than free will.

unconscious now determines.¹³ For example, the choosing experience—after your unconscious has decided, while not affecting whether you now go with A or B can influence what the unconscious will select in future choice sets. Supporting the conjecture that conscious thought has influence, [Roy Baumeister](#), [E.J. Masicampo](#), and [Kathleen Vohs](#) (2011) argue that “The evidence for conscious causation of behavior is ...empirically strong. However, conscious causation is often indirect and delayed, and it depends on the interplay with unconscious processes.”¹⁴

Imagine if one never had a choosing experiencesⁱⁱ

The choosing experience informs the unconscious, blabber-mouthing ideas to the conscious you, and these can influence what you will do next. Imagine our brains evolved absent choosing experiences. We would not know what we would do until we did it. You would not be able to tell the kids you were taking them for cheeseburgers rather than to boarding school.¹⁵ Bertrand Russell learned he loved Lady Ottoline Morrel only when he heard himself saying, “I love you.”—a surprise to him.¹⁶ Without choosing experiences, life would be a sequence of conscious surprises.

Without choosing experiences, your sense of self (conscious identity) would be lessened: believing that consciously choosing determines what happens next conveys an evolutionary advantage, even if it is not true—evolution has programed us to believe we consciously choose. We tend to think that the choosing experience is absent in worms and wolves—they go straight from I to III. Absent choosing experiences, it is difficult to coordinate your behaviors with your friends and colleagues—as Baumeister and his coauthors note, it is difficult to imagine

¹³ Keep in mind that while consciously thinking about what to do is a conscious thought, most are not choosing experiences. For example, you consciously realize you are reading this footnote.

¹⁴ They reviewed experiments where conscious thought was manipulated (e.g., asking you to imagine a future action) and subsequent behavior was then observed. They found many studies where experimentally manipulating conscious thought influenced behavior (e.g. you are more likely to perform an action if you first imagine performing it). As Baumeister and Bargh (2014) note, none of these experiments contradict the Libet result that the immediate cause of behavior is unconscious.

¹⁵ Quoting Baumeister and Masicampo (2010), “we think that the evolutionarily decisive advantages of conscious thought are not to be found in private, solipsistic ratiocination but rather in its contribution to communication.”

¹⁶ Morrel, a dramatic British socialite and arts patron, had an open marriage and was off and on with Russell for many years; she had many lovers. D.H. Lawrence portrayed her in *Women in Love*, and some critics believe she was the inspiration for *Lady Chatterley*. <http://www.theguardian.com/uk/2006/oct/10/books.booksnews>

unconscious talking. In general, consciousness seems to improve our decision-making, even if the unconscious makes the final call.

Other evidence on choosing experiences and subsequent actions

An issue complicating the study of the relationship between a choosing experience and the subsequent action is you cannot, by definition, consciously know whether your unconscious determined what you did—you cannot be conscious of your unconscious.

Consider how the perception of choosing to act, the action, and the outcome can all be distorted. Begin by deconstructing what Ralph said at the bar:

Yesterday I decided to go duck hunting, saw a duck in the sky, chose to shoot, shot the duck, and my dog retrieved the now-dead duck that I shot.

The sequence as Ralph perceives it is (1) he had two choosing experiences: to hunt and to shoot; (2) shot in a duck's direction; (3) these caused a shot duck; (4) he has a dead duck in his truck; and (5) he believes he committed premeditated duckicide. In Ralph's mind, he chose to shoot, and this resulted in an outcome, a dead duck. To be correct, all five of these implicit conjectures must be valid. However, the neurological research throws conjecture (3) in doubt, even if (1) and (2) are correct. But what about (1), (2), (4), and (5)? Much research in psychology indicates that they are often wrong.

To keep things simple, assume (4) is correct. (1) is incorrect if his recollection is false; maybe he shot reflexively but, being a responsible gun owner, Ralph cannot imagine shooting without first deciding to, so he concludes he must have decided to shoot. Maybe (5) is incorrect: maybe another hunter shot at the same time as Ralph, or maybe the duck died in midflight of natural causes as Ralph shot his gun. Maybe (2) is incorrect: maybe he did not shoot in the direction of the duck, and he only believes he did because after he shot, the duck fell from the sky.

Humans are great at (A) imagining we had a choosing experience when we didn't; (B) believing the choosing experience caused what we did; (C) believing we caused an outcome to occur when we did not (or believing we did not cause an outcome when we did); and (D) believing the outcome is different from what it is. Also, keep in mind that action does not require a choosing experience. This, plus these four inclinations, make it likely we misinterpret the importance and influence of conscious choice—independent of (3).

Let's start with most of what we do isn't preceded by conscious choosing

One way to think about the relationship between conscious choosing and the subsequent action is to contrast that relationship with actions not preceded by conscious choosing.

[Economists do not explicitly consider whether a purchase (e.g., buying a TV or going to Chez Faux for dinner) first requires conscious choosing.]

We are all aware that many of our actions are not preceded by a conscious decision to take action, so everyone knows action does not require prior conscious choosing. Above I referred to these as Type IV Events.

A: believing we had a choosing experience when we didn't

In a 1977 experiment by Nisbett and Wilson, subjects memorized lists of word-pairs, and in one treatment, the list included the pair "ocean-moon". Later they were asked to name a laundry detergent. Those whose list included "ocean-moon" were much more likely to answer Tide—not surprising. But, when asked why, the subjects said things like "Tide was the detergent their mother used." Or "It has a cool box," but hardly anyone said, "Because I memorized 'ocean moon.'" They imagined consciously choosing to say Tide because afterward they felt compelled to self-rationalize their choice of words. I observe my behavior and then create an explanation that I find plausible and works for me—sometimes, it is even correct.

B: believing the choosing experience caused what we didⁱⁱⁱ

We all want to believe we exert conscious control over our lives. So are compelled to attribute many outcomes to our conscious reckonings. Still, choice experiments demonstrate that the probability of you choosing a particular alternative can be manipulated by messing with your unconscious, independently of what you consciously reason. This supports the neurological results discussed above.

Specifically, recent studies with subliminal cues demonstrate that choice can be influenced by these cues (information that never reaches conscious awareness). Subjects are shown a piece of information for a few msec. followed quickly by a random letter that masks the information. Behavior is influenced, demonstrating that the conscious is not always in charge, even when we think it is.¹⁷

¹⁷ Subliminal research got a bad name in 1957 when James Vicary claimed he found that inserting "Drink Coke" and "Eat Popcorn" in movie frames caused movie goers to buy more of each. He could not replicate the study and later

Subliminal persuasion is easily demonstrated when it addresses a conscious physical need (hunger, thirst, etc.). In 2006, the psychologists [Johan Karremans](#), [Wolfgang Stroebe](#), and [Jasper Palmier-Claus](#) demonstrated

priming of a brand name for a drink [Lipton Iced Tea] will only affect choice behavior of people who are thirsty (i.e., have a goal to drink)

Supporting this finding, in a 2009 study, [Christina Bermeitinger](#) and her coauthors showed

Subliminally presented brands of dextrose pills biased choices of tired (but not non-tired) participants.

More recently, [Thijs Verwijmeren](#) and colleagues investigated the influence of subliminal cues on the choice between two beverages when one is consumed more frequently. The two beverages were Lipton Iced Tea and Spa Rood (a popular Dutch mineral-water brand). The subliminal cue was either “Lipton Iced Tea” or a control. It only influenced choice if the subject was thirsty, consistent with the two other studies. The Lipton cue had no influence if the subject habitually drank Lipton, even if he was thirsty. But if he more often drank Spa Rood, the Lipton cue increased the probability they would choose Lipton. Respondents chose an alternative different from the beverage they typically chose because they were presented with information, unconsciously obtained—conscious processing did not determine their selection.¹⁸ This research is consistent with the incentive-salience hypothesis discussed in Chapter 4: when one is in a drive state, alternatives that address that state become more salient and even more salient if that alternative is associated with a cue. The research on subliminal cues drives home the finding that one need not to be conscious of the cue: it does have to be observing a pretty girl drinking a Pepsi (p. x).

said the whole thing was a gimmick. His initial claim caused a public uproar and made research on subliminal cues a dangerous endeavor. It led to subliminal advertising being banned in Australia, Britain, and the U.S.

Earlier I note that Newell and Shanks (2014) reject the literature on unconscious influences on behavior including the literature I discuss here on unconscious priming and cues. They argue that this literature’s conclusion that the subject was not aware of the cue is based on an inadequate test of awareness. Someone telling you they were not aware of the snake in the room does not prove no conscious awareness. I might have seen the snake but not with high confidence, so, to be safe, report no awareness, or so they argue. They also argue awareness is likely a continuous variable that does not jump from unconscious to conscious when awareness reaches a specific intensity, meaning conscious awareness is a continuous variable. Summarizing, they suspect there is conscious awareness when the researcher concludes there is none. And, if correct, this would call into question the subliminal priming support for unconscious choosing. For a rebuttal from the same issue of *Behavioral and Brain Sciences* see Ap Dijksterhuis et al. See also [Samantha Brooks](#) and [Dan Stein](#). They find that subliminal vs. conscious cues activate different brain regions.

¹⁸ The subjects were not debriefed on why they made the selection they made. In a personal email Thijs Verwijmeren speculated that if asked the subjects would have come up with an incorrect reason for their choice.

To summarize, subliminal cues affect choices when one is motivated to fulfill a physical need. This raises the question of whether choice can be affected by subliminal information when the motivation is more implicit, such as a motivation to succeed. This conjecture was tested and found correct in 2012 by [Maxim Milyavsky](#), [Ran Hassin](#), and [Yaacov Schul](#): information processed unconsciously can influence choice whenever there is a motivation for choosing, not only when it's a physical and conscious need.¹⁹

C: believing you caused it when you didn't, and believing you didn't when you did^{iv}

The Denver Broncos football team lost the 2014 Super Bowl because I, stupidly, wasn't paying attention during the Bronco's disastrous first-play from scrimmage.

Exotic examples of things we cause but don't think we cause are the voices schizophrenics hear, the things hypnotized people do, and the words we spell out on Ouija boards. Many students who get a *D* on my first midterm believe they did not cause their *D*: It was caused by Edward's lousy grading. In contrast, *A* students believe their conscious efforts caused their grade. When we pray to God to give us the strength to act and then act, we attribute the act to God rather than to ourselves—problematic if God does not exist or does not care what we do. It could not have been an operator error when I could not find this chapter on my computer.

We have all heard of false confessions, even when the confessor believes they did it. In 1996 [Saul Kassin](#) and Katherine Kiechel produced these in the laboratory. Each subject was falsely accused of crashing the experimenter's software, thus ruining the study. The subject typed a list of letters read off by a research assistant posing as a second subject. The assistant read them at 43 letters per min. or 72 lpm. Before the trial started, subjects were warned that the software would crash and data would be lost if they hit the ALT key. After sixty seconds of

¹⁹ Motivation to achieve was manipulated: before a "separate" choice task, subjects played "find the words" in a square of letters. One set of subjects' letter-squares included words related to achievement; the other treatment (the control group) saw only motivationally-neutral words.

Then, before the choice task, each subject was subliminally cued with a word, or non-word. There were 16 word-cues and 16 non-word cues, and each cue was related to only one of four word-categories (emotions, cloths, family and fruits). E.g., the emotion words were love, joy, anxiety, and jealousy, and the emotion non-words were nadsach, sitam, dagma, and ragham.

Then the four word-categories came on the screen and the subject had to guess the correct category (e.g. fruits if the subliminal cue was mango). Correct guesses were rewarded. The subjects in the control group guessed correctly no better than random; the subjects who were motivated to achieve guessed correctly significantly better than random. The critical point is that these subjects were not consciously aware that they had been motivated to achieve.

typing, the experimenter said the program crashed and accused the subject of hitting the ALT key. No one had, and everyone initially denied it. The experimenter then turned to the letter reader and asked if she saw anything. She said, “I did not see what happened,” or “I saw the subject hit the ALT key.” Of the subjects who heard 72 lpm. and were witnessed against, 100% signed a confession, and 65% became convinced they were guilty. [When leaving, the following subject (another R.A.) privately asked, “What happened?”.] Thirty-five percent confabulated details to support their false belief that they hit the Alt key. [They were asked in the debriefing to say when they hit the ALT key.]

Of the subjects who heard 43 lpm. and were witnessed against, 89% signed the confession. Of those who got this slow pace and were not witnessed against, 35% signed the confession. In summary, it is not difficult to make you think you did something you didn’t, even if it’s a bad thing.

[Billy Wayne Cope](#) is in jail for raping and murdering his daughter. He confessed and provided gory details after the police told him the semen in her vagina matched his DNA—it did not—the DNA was much later matched to a rapist who had recently moved to the neighborhood.

My hard work, not my co-authors, is responsible for many of my published co-authored papers, but they believe it was their hard work. Correctly attributing an outcome that results from a complex interaction between people (work, social, sexual) is tricky—exactly who initiated that kiss? And, if we like the outcome, we tend to attribute it to our efforts but attribute it to others if we don’t: *self-serving attribution*; it increases self-esteem.

We ascribe causation if physical action on our part immediately precedes an observed outcome. There are many light-switch experiments: a light flashes on or not, and a switch. They demonstrate that subjects think they are controlling the light, even when it is not attached to the switch. We believe we are causing things even when we are not.

As discussed in Chapter 9, there is the tendency to ascribe causation if physical action on our part immediately precedes an observed outcome. (I flicked the switch and the light came on. I shot in the air, and a duck fell to the ground.) If *A* proceeds *B*, we tend to infer *A* caused *B*, particularly if *B* is desirable, so we take authorship for *A*. Man evolved to think in this second-best manner.

Consider an experiment where every second or so, you choose whether to press a button or not and then a light flashes or not. The light does not always flash after you push the button, and sometimes it flashes even when you didn't push it. Treatments vary by the actual amount of control you and the other subjects have over the light (e.g., 25%, 50%, and 75%)—full control (100%) means the light flashes every time you push the button and never flashes when you don't.²⁰ You can be tricked into thinking you have more control than you have simply by increasing the number (not the proportion) of times where the outcome is what you predict (the number of successes), independent of your actual amount of control. Successes are when you flick, and the light comes on, and when you don't, and it doesn't.

In a different light experiment, Alloy and Abramson studied the perception of control using a sample that intentionally included both depressed and non-depressed subjects. The non-depressed subjects, but not the depressed subjects, overestimated how much control they had over whether a light came on. Like the study described above, subjects flicked a switch, and a light either came on or not. Subjects suffering from depression accurately assessed their lack of control—which could have something to do with why they were depressed in the first place. A conjecture for why we suffer from an *illusion of control* is actual control is easier to recognize than no control (independence between our actions and what happens next).

In a simple 1983 experiment by [Howard Tennen](#) and Jeanette Sharp, the light came on a fixed proportion of the time, independent of whether the subject pushed the button. Hence, the subject had no control, but the subjects experienced the illusion of control. Before the trials, each subject was classified as either an *internal* or *external* type using an [I-E scale of control orientation](#). [Externals tend toward believing what happens is out of their control, whereas internals tend to view their actions as having influence.] The hypothesis was that internals would suffer the illusion of control, but not the externals. Both types suffered it, internals just more so.

In 2008 [Ap Dijksterhuis](#), [Jesse Preston](#), Daniel Wegner, and Henk Aarts showed that whether or not you take credit for an action can be influenced by subliminal cues that “direct

²⁰ With 25% control, the light comes on 75% of the time when the button is pushed and 50% of the time when the button is not pushed. Whereas, with 50% control the percentages are 75% and 25%. Control is the difference between the two percentages.

attention to self”.²¹ Subjects looked at a sequence of strings of letters on a computer screen in a fast race with the computer to determine whether each string was a word. They were told the first response (theirs or the computer’s) would blank the screen, after which another string appeared. However, the screen blanked half a second after each string appeared, independent of what the respondent did—the subject never caused the screen to blank. After each screen blank, the subject was asked if they answered before the computer. Before each trial, the subject was subliminally cued with either a personal pronoun (“I” or “me”) or a neutral word. When the cue was a personal pronoun, the subject was more likely to say they beat the computer.

In another experiment, subjects subliminally cued with the word “God” were less likely to say they beat the computer, but only if they were believers. [It is probably correct to say that many believers attribute many outcomes to God—God’s will.] In another experiment, Aarts demonstrated that in situations where causation is unclear, “feeling of control and self-causation” are enhanced if success is subliminally conveyed.

D: believing the outcome is different from what it is

There is **it** and what **caused “it”**. This sub-section is about whether **it** is true, not what caused it. Was the duck really dead, or even a duck? My mother once called to report her sister died and told me she had called the funeral home to come for the body—the funeral home went to the hospital and reported that their services were not yet required. People make mistakes about the facts on the ground.

Truth, like beauty, is in the beholder’s eye. The accused believes there was no rape because the accused believes the victim wanted to have sex. He believes the victim chose to have sex and sent signals indicating this. He is attributing the sex to a choice made by the other party when, in fact, the other party did not necessarily make that choice. And, if she didn’t, the accused caused the sex. Many Americans believe Barack Obama is a U.S. citizen, but many don’t. Some think our planet is warming, others don’t, and everyone can’t be correct. We judge truth based on

²¹ In a 2013 paper David Shanks and his coauthors report that they could not replicate another Dijksterhuis finding, his finding that people who think about professors just before they take an intelligence test do better than those who think about football hooligans. This, in part, has called into question the findings that “behavior can be influenced or ‘primed’ by thoughts or motives triggered unconsciously...” (Nature 2013). In a [2012 open letter](#) to researchers in social priming, Daniel Kahneman said more replication of results in this area was needed, but that he is a “general believer” in priming effects.

what is happening around us, filtered through personal perspective: our implicit biases, what our group/tribe believes, and “facts” reported to us by second-hand sources we trust.

Finally, keep in mind that your actions are influenced by your unconscious does not, by itself, imply you are not making choices—it is **your** unconscious. Unconscious choice is consistent with the Assumptions of NBT Plus, except for 11b. What about the causal determinism discussed in Chapter 9? Causal determinism must be the case unless God, spirits, magic, or randomness intervenes. [Many of us subscribe to such beliefs to a degree, often unconsciously.] Causal determinism is consistent with actions being initiated unconsciously. The question of whether you can make a choice—consciously or unconsciously—in a world of causal determinism comes down to how choice is defined.

Summarizing:

This ends Part II, our discussion of what is a choice and do humans make choices. Neoclassical choice theory, NCT, assumes economici make choices: “choice” is in the theory’s title. But, saying that behavior (as explained by NCT) is chosen behavior is inconsistent with the street sense of *choice*, which is an economicus could have experienced a different path. NBT assumes it must experience its HRAP, and both its ordering and what is available are given, so an economicus cannot behave differently than it does. Its behavior is deterministic in the restrictive sense that at the beginning of time, the configuration of atoms was such the species economicus would inhabit the earth.

Applying NCT to human behavior requires that humans are economici. Most humans don’t believe this. They believe the behavior of animals and refrigerators is causally deterministic, but they are disinclined to believe it holds for humans.

The question is how to make the behaviors of you and economici *choices*. It is done by defining, implicitly, a choice as any situation where there are two or more available paths. Even though your ordering of paths and which are available are exogenous and constrained to experience the highest-ranked available path.

Many physicists and neuroscientists are *incompatibilists* (believe causal determinism is incompatible with free will and choice). Many philosophers are *compatibilists*, believing causal determinism is consistent with free will and choice. Economists tend to be implicit and unaware

compatibilists: we implicitly define free will as “your will is free if external constraints do not limit you to one alternative.”

Economists should be aware that most physicists and neuroscientists reject that free will is compatible with causal determinism.

The issue is whether the choosing experience influences which alternative you select: While we believe that it determines which alternative we select, there is strong evidence that many choices (in the economic sense) are made unconsciously: at the final step, the unconscious selects the alternative and, sometimes, you then have a conscious experience of choosing what your unconscious decided on. [Of course, the evidence is debated.] While this isn't fatal for NBT (most axiomatic expressions of NCT say nothing about the conscious vs. the unconscious), it is an injury. I suspect most economists believe the choosing experience is why you did what you did (most other people believe this as well: unconscious choosing it not what most of us had in mind.

With questions about the ability to choose in mind, I now turn to *Part III: Most Moral Philosophers are not Welfare Consequentialists, so what are they?*

ⁱ Sources:

1. Henk Aarts and Ap Dijksterhuis (2003) The Silence of the Library: Environment, Situational Norm, and Social Behavior, *J. of Personality and Soc. Psych.* 84(1): 18-28.
https://www.researchgate.net/profile/Ap_Dijksterhuis/publication/10958510_The_Silence_of_the_Library_Environment_Situational_Norm_and_Social_Behavior/links/0deec51f8b20a69e81000000/The-Silence-of-the-Library-Environment-Situational-Norm-and-Social-Behavior.pdf
2. Roy Baumeister, E.J. Masicampo and Kathleen Vohs (2011) Do conscious thoughts cause behavior, *Annual Rev. of Psych.* 62: 331-61
3. Roy Baumeister and John Bargh (2014) Conscious and Unconscious: Toward an integrative understanding of human mental life and action, Chapter 3 in *Dual process theories of the social mind* (eds. Jeffrey Sherman, Bertram Gawronski and Yaacov Trope), York, Guilford Press.
4. [Susan Blackmore](http://www.patheos.com/blogs/monkeymind/2007/08/susan-blackmore-on-ben-libet.html) (2007) Mind over matter? *Pathos* at <http://www.patheos.com/blogs/monkeymind/2007/08/susan-blackmore-on-ben-libet.html> An obituary essay on Libet's research
5. John Creswell, James Bursley and Ajay Satpute (2014) Neural reactivation links unconscious thought to decision making performance. *Behavioral and Brain Science* 37(1). doi:10.1093/scan/nst004
6. Ruud Custers and Henk Aarts (2010) The unconscious will: how the pursuit of goals operates outside of conscious awareness, *Sci.* 329.
7. Itzhak Fried, Roy Mukamel and Gabriel Kreima (2011) Internally generated reactivation of single neurons in Human Medial Frontal Cortex Predicts Volition. *Neuron* 69:548-562.
8. Rob Holland, Merel Hendriks, and Henk Aarts (2005) Smells like clean spirit: Nonconscious effects of scent on cognition and behavior, *Psych. Sci.* 16(9): 689-93.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.546.8181&rep=rep1&type=pdf>
9. William James (1890) *The Principles of Psychology, Vol. 1*, N.Y., Henry Holt and Company
http://books.google.com/books/about/The_Principles_of_Psychology.html?id=TMrJfcaC8bYC

10. Aaron Kay, Christian Wheeler, John Bargh, and Lee Ross (2004) Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice, *Organizational Behav. and Human Decision Process* 95: 83-96.
11. Benjamin Libet, (2004) *Mind Time: The Temporal Factor in Consciousness*, Cambridge Massachusetts and London England, Harvard University Press.
12. Benjamin Libet, Curtis Gleason, Elwood Wright and Dennis Pearl, (1983) Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential) - The Unconscious Initiation of a Freely Voluntary Act, *Brain* 106: 623–642.
13. Jeff Miller, Peter Shepherdson, and Judy Trevena (2011) Effects of clock monitoring on electroencephalographic activity: is unconscious movement initiation an artifact of the clock? *Psych. Sci.* 22(1): 103-9. <https://doi.org/10.1177/0956797610391100>
14. Aaron Schurger, Myrtos Mylopoulos, and David Rosenthal (2015) Neural antecedents of spontaneous voluntary movement: a new perspective. *Trends in Cognitive Sci.* Vol 20(2): 77-79. <https://doi.org/10.1016/j.tics.2015.11.003>
15. Aaron Schurger, Jacobo Sitta, and Stanislas Dehaene (2012) An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proc. Natl. Acad. Sci. USA.* 109: E2904-2913. <https://doi.org/10.1073/pnas.1210467109>
16. Kerri Smith, (2011) Neuroscience vs philosophy: Taking aim at free will, *Nature* 477: 23-25.
17. Chun Siong Soon, Marcel Brass, Hans-Jochen Heinze and John-Dylan Haynes (2008) Unconscious determinants of free decisions in the human brain, *Nature Neuroscience* 11: 543 – 545.
18. Daniel Wegner (2002) *The Illusion of Conscious Will*, Cambridge MA. MIT Press.

ii Sources:

1. Roy Baumeister and E.J. Masicampo (2010) Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review* 117(3):945–71.
2. Roy Baumeister, Kathleen Vohs and E.J Masicampo (2014) Maybe it helps to be conscious, after all, *Behavioral and Brain Sciences* 37(1).
3. Mark Bear, Barry Connors and Michael Paradiso (2007), *ibid*
4. Tracey Henley (2019) *Hergenhahn’s An Introduction to the History of Psychology, Eighth Edition*, Cenage.
5. Richard Nisbett and Timothy Wilson (1977) Telling more than we know: verbal reports on mental processes, *Psych. Rev.* 84(3): 231-259. This article has been cited more than eight thousand times.
6. Daniel Wegner (2002) *ibid*.

iii Sources:

1. Christina Bermeitinger, Ruben Goelz, Nadine Johr, Manfred Neumann, Ulrich Ecker and Robert Doerr (2009) The hidden persuaders break into the tired brain, *J. of Exp. Soc. Psych.* 45(2): 320-326.
2. Samantha Brooks and Dan Stein (2014) Unconscious influences on decision making: neuroimaging and neuroevolutionary perspectives. *Behavioral and Brain Sciences* 37(1). doi:10.1017/S0140525X13000666
3. Samantha Brooks, V. Savov, Elin Allzen, Christian Benedict, Robert Fredriksson, and Helgi Schioth (2012) Exposure to subliminal arousing stimuli induces robust activation in the amygdala, hippocampus, anterior cingulate, insular cortex and primary visual cortex: A systematic meta-analysis of fMRI studies. *NeuroImage* 59:2962–73.
4. Ap Dijksterhuis, Ad van Knippenberg, Rob W. Holland, and Harm Veling (2014) Newell and Shanks’ approach to psychology is a dead end, *Behavioral and Brain Sciences* 37(1).
5. Johan Karremans, Wolfgang Stroebe and Jasper Claus (2006) Beyond Vicary’s fantasies: The impact of subliminal priming and brand choice, *J. of Exp. Soc. Psych.* 42: 792-798
6. Maxim Milyavsky, Ran Hassin and Yaacov Schul (2012) Guess what? Implicit motivation boosts the influence of subliminal information on choice, *Consciousness and Cognition* 21(3): 1232-41.
7. Thijs Verwijmeren, Johan Karremans, Wolfgang Stroebe and Daniel Wigboldus (2011) The workings and limits of subliminal advertising: The role of habits, *J. of Consumer Psych.* 21: 206-213.

iv Sources:

1. Henk Aarts (2007) Unconscious authorship ascription: The effects of success and effect-specific information priming on experienced authorship, *J. of Experimental Soc. Psych.* 43: 119-126.

-
2. Alison Abbott (2013) Disputed results a fresh blow for social psychology: failure to replicated intelligence-priming effects ignites row in research community, *Nature (News)* 497(7447).
<http://www.nature.com/news/disputed-results-a-fresh-blow-for-social-psychology-1.12902#ref-link-1>
 3. Lauren Alloy and Lyn Abramson (1979) Judgment of contingency in depressed and non-depressed students: Sadder but wiser? *J. of Exp. Psych. Gen.* 08(4) 441-485.
 4. Ap Dijksterhuis, Jesse Preston, Daniel Wegner and Henk Aarts (2008) Effects of subliminal priming of self and God on self-attribution of authorship for events, *J. of Exp. Soc. Psych.* 44: 2-9.
 5. Saul Kassin (2007). Internalized false confessions, In *Handbook of Eyewitness Psychology: Volume 1, Memory for Events* (eds. M. Togli, J. Read, D. Ross and R. Lindsay), Mahwah NJ, Erlbaum.
 6. Saul Kassin and Katherine Kiechel (1996) The social psychology of false confessions: compliance, internalization, and confabulation, *Psych. Sci.* 107(3): 125-128.
 7. Ben Newell and David Shanks (2014) Unconscious influences on decision making a critical review. *Behav. And Brain Sci.* 37: 1-61. doi:10.1017/S0140525X12003214
 8. Howard Tennen and Jeanette Sharp (1983) Control orientation and the illusion of control, *J. of Person. Assess.* 47(4): 369-374.
 9. Daniel Wegner (2002) *ibid.*

Part III: Most moral philosophers are not welfare consequentialists, so what are they?

08172022

Except for Chapter 2, Parts I and II were individual-centric. Chapter 2 identified welfare economics as a subset of WC—only consequences matter, and the only consequences that matter are the WB (welfare) effects on society's members.

Chapter 11: A primer on welfare economics vs. other ethics

Quoting a Nobel laureate in Economics, [Sir John Hicks](#) (1904-89)

*If one is a utilitarian [welfare consequentialist] in philosophy, one has a perfect right to be a utilitarian in one's economics. But if one isn't...one also has the right to an economics free of utilitarian assumptions.*¹

03122022

Moral philosophy studies, builds, and compares theories about right and wrong. Are your actions right (morally right) or wrong (morally wrong), right as they affect you?² Right as they affect others? As noted, an ethic is simply "a method to test right from wrong". All moralists agree right is right and wrong is wrong but disagree on how both should be defined and measured. Welfare economics (introduced in Chapter 2) is all about determining whether a policy is right or wrong. It is an ethic based on one objective—maximizing societal WB—an objective many reject. As in Chapter 2, I mostly avoid the words “good” and “bad”: both can mean many things besides “right/moral” and “wrong/immoral”.

An ethic can be as straightforward as a list of specific right and wrong actions (e.g., The Ten Commandments). But many find such lists unsatisfying: the list can get long, my list is different from yours, and the question arises, "Why is that on your wrong list?" leaving you to answer, "Just because" or "God said it was wrong." So, most non-religious ethics start not with a list of right and wrong actions but with foundational axioms/assumptions. Then the ethic derives, through deduction, whether an action is right or wrong. Kant, for example, got it down to a single axiom, which he called a *precept*.

Most people (but not necessarily moral philosophers) judge acts moral or immoral based on all sorts of different overlapping criteria. These include how it makes them feel, religious beliefs, perceived rights, intentions, consequences, whether the act is natural (human nature), process (how it was decided upon), disgust (Yuk!), and inherent value (value for its own sake)—to name a few (see Chapter 2). In this chapter, ethics based on different criteria for judging are contrasted with WC and welfare economics.

¹ While Hicks used the word “utilitarian” he meant, in our parlance, WC.

² Historically, much of moral philosophy has concerned itself with how best to live your life. For Socrates ethics was only “How one should live?” Economics avoids this question; economicus has little choice in the matter: he is forced to choose his highest-ordered available life.

An aside: Why be ethical?ⁱ

This chapter critiques contrasting ethics (proposals for determining right from wrong actions and being the right kind of person). This chapter does not do justice to why one should behave ethically.

Welfare economists, unlike most ethicists, need not worry about the “right kind of person” side of ethics: every economicus must experience their highest-ranked available path, HRAP, whether they happen to be Mother Teresa, Donald Trump, or a sociopath who pleasures in the pains of others—NBT says you are constrained to do what you got to do.

Welfare economists only worry about whether behaviors and policies will increase or decrease societal WB. Policies that increase societal-WB should be imposed on society’s members by making paths without them either unavailable or unattractive.

In contrast, ethics that admit the possibility that the individual can select more, rather than less, ethical actions and live more, rather than less, righteous lives must deal with why an individual would want to.

Many individuals buy into a specific ethic —often a religious ethic. My aunt Edith lived a Christian ethic (King James Bible: Lutheran interpretation). Why? She wanted to be accepted by individuals of her faith and did not want to go to Hell. Faithful Buddhists and others live their lives trying to increase aggregate WB by reducing suffering. Buddhist claim is that being ethical will lessen the suffering of you and others.

But what about the skeptics: *Ethical Skeptics*: individuals skeptical of all ethical reasoning.³ What if you don’t believe in Hell? Don’t accept the Socratic claim that being virtuous is our nature or the Kantian claim that you are duty-bound to behave ethically? Why shouldn’t I cheat, lie, and cause mayhem if it increases my WB or fulfills my wants and desires? Why should I try to be the right kind of person?]

I start my brief survey of ethics by discussing a few that give WC a role in determining right from wrong but don’t give it the only role. Buddhist ethics is an excellent place to start because it, like welfare economics, emphasizes welfare consequences while being quite different from WC. Then I summarize other ethics that give welfare consequences a role but not the only role. Discussed then is the liberalism of John Stuart Mill.

Then I move on to ethics where WC plays a little or no role in determining right from wrong.

But before I discuss Buddhist ethics, I elaborate a bit more on WC and its particular case, welfare economics.

Processists vs. consequentialistsⁱⁱ

In Chapter 2, consequentialism was discussed—WC. Welfare economics judges right from wrong solely based on consequences. An act should be judged based on its outcome, not what

³ “...to be skeptical about ethics is to be skeptical about the force of ethical considerations,” (Williams 1985 and 2006, Chapter 2). The individual does not necessarily behave unethically, rather, her behavior is not influenced by ethical considerations. “A skeptic, after all, is merely a skeptic.”

caused it (Mill 1867).⁴ Hume was an early consequentialist. In contrast, *processists* deem an act right or wrong based on whether the process used to choose the act was a right or wrong process, independent of the result.

A *welfare consequentialist* would judge the killing of a child by a sadistic pedophile moral if that child were destined to be a Hitler or Pol Pot—millions of lives would have been saved. A processist would disagree, arguing that the intention and the process (the desire to sadistically, sexually murder) is wrong, making the killing of baby Hitler wrong. As discussed in Chapter 2, processists and welfare consequentialists would disagree about the [rape of Lucretia](#).

Kant was a processist believing acts not based on a duty to behave morally are not necessarily of moral value—right is determined by your will/intention to do your duty, not what you actually do.

Processes are typically judged right or wrong based on their fairness or justness. For example, some advocate majority voting because they believe it is a fair and just process for group decision-making. If it happens to increase WB, that's frosting on the cake. If the outcome decreases WB, that does not make the process wrong. The U.S. founding fathers chose majority voting as a way for the people to elect their representatives and for those representatives to make policy. Still, they limited the influence of voting by requiring that decisions can't violate the *U.S. Constitution* or *Bill of Rights*.⁵ They rejected majority voting by itself as the best process, but at the same time judged decisions by a process based on the will of the people superior to decisions by Monarch or Church. Implicit is the idea that constitutional policies chosen by voting are right policies.

A process is deemed fair and right if everyone involved has agreed to it. Picture the Gary Larson cartoon with four desperate, starving individuals in a lifeboat, you, two other humans, and a golden retriever. You all agree to draw straws to determine who will be slaughtered for

⁴ Qualifying a bit, a welfare consequentialist would say the process is important if the process itself is enjoyed or disliked. Imagine you love to eat chocolate ice cream and would choose to eat it, but your mother says you have to eat it. You might get less WB in the latter case. That said, economists think little about the consumption process. Economists do not say things like, "We need to take account that chocolate is enjoyed more if you eat it slowly and concentrate on its taste and mouth feel."

⁵ This statement is not quite correct. There is the Electoral College.

everyone else's dinner.⁶ You lose. An observing processist would have no problem with the outcome. A welfare consequentialist would likely argue that the dog should be dinner: since dogs have a shorter lifespan than people, there is less future WB to lose.⁷

Most of us judge on both process and outcome. Many of my students believe happiness without effort is wrong—“it is important to earn your WB.” In 1974 the philosopher [Robert Nozick](#) cooked up a thought experiment: a virtual-reality pleasure machine—you are wired up and have the experience of living a life of immense enjoyment, all the while sitting in a sealed black box.⁸ (We discussed artificial pleasure in Chapter 3.) Most people say they do not want that kind of pleasure. Nozick interprets most peoples' disdain for the machine as a rejection of WC.

Many legal systems judge guilt and determine punishment based on both consequence (the victim is dead) and process (how did you kill him and your intent). We excuse you, or limit your punishment, if your action was an accident, in self-defense, if you are insane, or if it was a mistake (e.g., you reasonably thought the sex was consensual).

Consider the role of WB in different ethics

Ethics differ by how much WB counts for determining right from wrong. Foremost, either WB is important, or it isn't. For example, for a pure processist, it isn't.

If WB does play a role in determining right from wrong, there is the extent of that role. At one extreme, WB is the only thing that counts; this is WC. In one large subset of WC, the objective is to maximize aggregate WB; Benthamite utilitarianism is a subset of this maximize aggregate WB subset. While welfare economics is a subset of WC, it isn't necessarily a subset of

⁶ Some might question whether the dog was capable of agreeing to the process. There is a large philosophical literature on dogs in lifeboats. See the debate between Singer and the rights theorist [Tom Regan](#), “[Dog in the lifeboat: an exchange](#),” *New York Review of Books*, April 1985. And “Rights, justice and duties to provide assistance,” by the environmental philosopher [Dale Jamieson](#) in *Ethics* 100, January 1990.

⁷ But eating you, rather than the dog, could increase aggregate WB. The dog might be a happy dog and you might be a depressed individual incapable of experiencing WB. And, while you might be racked with guilt for eating the dog, the dog would likely not sweat eating you.

⁸ You are suspended in darkness and silence, and your brain is stimulated such that you perceive a happy life. None of my students ever want to live in his [Pleasure Machine](#), but have a hard time explaining why saying only “it isn't real happiness.” Philosophers use the adjective “authentic.” In the 1973 Woody Allen movie [Sleeper](#), did Woody and Diane Keaton have real orgasms in the [Orgasmatron](#), or just think they did, and does it matter? It would matter to a processist. As an alternative to the chamber, consider happiness pills (SSRI's)—many take them.

the maximize aggregate-WB subset.⁹ Chapter 2 labeled ethics where WB is important but where WB isn't the only determinant of right from wrong as *NotOnlyWB ethics*. Many ethics are NotOnlyWB.

A common view amongst welfare consequentialists and NotOnlyWB ethicists is that self-awareness conveys a greater capacity for enjoyment and suffering. Additionally, so does the ability to consciously imagine your future and mortality. Pigs have more capacity for enjoyment and suffering than spiders and humans more than pigs. But this does not imply every human is capable of more WB than every pig. So, even if every human and pig counts equally, it might be ok aggregate-WB-wise for humans to eat pigs. Still, it might be even better if we ate dogs—supposing they taste about the same: dogs are arguably less self-aware than pigs, so pigs have more capacity for suffering. Then again, many humans get WB from having a live dog around and would experience a decrease in WB from hearing about humans eating dogs—maybe we should eat the dog and get a pig for a pet.

Most ethicists who care about WB agree that individuals who were once alive but are now dead (George Washington and Smokey the Bear) can't experience WB. The same goes for individual entities for which life was never an option (Rocky the rock and Ralph the refrigerator). Most ethicists seem to believe that plants can't experience WB. [The ethical implications of plant WB are drastic if WB enhancement is the goal.] They seem to exhibit a desire to survive, seeking water and nutrients with their roots and energy with their leaves. And they use chemicals to fight off pests. That said, few ethicists argue that individual plants should have moral standing. Peter Singer would have to write *Plant Liberation*.]

Should all WB countⁱⁱⁱ

Different NotOnlyWB ethics differ in whether some kinds of WB should be excluded when WB is aggregated, excluded because they are false, unreal, or generated by unacceptable behaviors or thoughts. Welfare economists don't parse acceptable and unacceptable bearers of WB; they only consider that behavior can reduce the WB of others: *negative external effects*. But many NotOnlyWB ethicists reject the welfare economists' view that it does not matter how WB is generated. One example is John Harsanyi, mentioned in Chapter 2, who argued that WB derived

⁹ In contrast, the objective of preference utilitarianism is to maximize societal preference-fulfillment, not societal WB. So, if want and desires, and WB don't matchup, preference utilitarianism isn't necessarily an ethic where WB counts.

from hatred and prejudice should not count. Another is the Canadian philosopher and welfarist [L.W. Sumner](#), the author of *Welfare Happiness and Ethics* (1999), who argues that happiness does not necessarily produce WB. His example: imagine that you are happy and have been for years because you have a devoted and loving spouse. But it is all a deceit: he never loved you, and his affairs started right after the honeymoon. Sumner argues that whether your happiness contributes to your WB, and in turn aggregated WB, depends on how you would have reacted if you had found out the truth. If you would have said “C’est la vie”, your past happiness should count. But, if instead, you would have concluded your life was a farce of your priorities, a life not your own, then your past happiness should not be part of WB aggregation. A welfare economist would say enjoyment is enjoyment.

Whose welfare counts (who gets moral standing)?^{iv 10}

Related to the issue of what kinds of WB should count is whose WB should count.¹¹ Jane Austen referring to the casualties in a far-off war,

How horrible it is to have so many killed, and what a blessing that one cares for none of them ([Lord Brabourne edition of Jane Austen's letters](#))

There are billions of specific people and animals currently alive and many more coming down the pike, so who counts is a critical question for anyone who argues that whether an action is right or wrong depends, at least in part, on how it affects the WB of individuals. For example, to give moral standing to Happy, a specific elephant, would mean what? At a minimum, it would mean her WB should count.¹² Most economists limit moral standing to humans, but not all

¹⁰ This is a good place to distinguish between moral and legal standing. A party has “legal standing”, *locus standi*, if they can bring a suit in court—the party can challenge another party in court. For example, Donald Trump, Tesla, Alabama, and Switzerland have it. Many believe fetuses should have it. Neither legal nor moral standing implies the other. Legal standing is not limited to living things. With a few exceptions, the only animate entities with legal standing are humans. Historically, in spite of being human, women, children, and slaves, did not have legal standing. Some countries are moving toward legal standing for non-human animals such as great apes, elephants, and whales (Wright 2022).

¹¹ See also Chapter 2, Section I.

¹² You might argue that an elephant’s WB is difficult to measure. I agree. But I would add it is easier to determine what affects an animal’s WB than it is to determine what affects yours: you are more complicated. My dog, Sofie, was a simple creature, rarely deceptive or unclear; she wanted to play ball, run, pounce, and have her stomach rubbed and preferred meat to cheese and cheese to bread.

In recent years, lawyers have argued in a few U.S. courts that a few specific animals should be granted legal standing—specifically for release from captivity under habeas corpus (a “person” must be released from detention unless a court decides there are legal grounds for the detention). Like in all legal cases, the animals are represented by humans. A famous example is Happy, a lonely elephant residing at the Bronx Zoo (Wright 2022). In 2020, Judge

humans. Recollect that Bentham was species inclusive, so is Peter Singer. They argue for the inclusion of everyone, and anything, who/that can feel pleasure and pain. Many other modern ethicists require a degree of self-awareness. As noted in Chapter 1, research suggests dolphins, elephants, humans, and magpies are self-aware.

The Harvard philosopher [Christine Korsgaard](#) requires, as did Kant, that to count, one must have the ability to be "aware of the grounds for one's beliefs and choices," of our "reasons for thinking and acting as we do." To count, we must be able to "step back" and "evaluate." "And this enables me to take responsibility for what I do." because "we are not completely governed" by our instincts, desires, and emotions. She believes only humans have this ability to think about themselves abstractly and symbolically. Kant argued that individuals incapable of taking responsibility for their behavior have no right to moral standing. Other animals do not have this capacity, so our obligations are only to humans. Korsgaard argues that because humans are the only moral animals, we have obligations to ourselves and other animals because we share many sensations.¹³ Aristotle and Aquinas also made purposefulness and the ability to evaluate your actions as requirements for moral standing.

That economists limit moral standing to humans is simply an ethical choice.¹⁴ If it's only humans, the enjoyment and suffering animals experience can have influence but only to the extent that their experiences cause human enjoyment or suffering.¹⁵ A few environmental and ecological economists include other sentient creatures; fewer include all individuals who can experience pain.

There is then the separate question of whether all the individuals in an accepted species should have standing. I imagine most welfare economists would say the decision to include or

Allison Tuitt of the New York State Supreme Court, while expressing sympathy for Happy's plight, ruled that she was not a "person", so habeas corpus does not apply.

¹³ For more on our duties to animals see Nussbaum on the complexity of animal values. See Nussbaum, and Korsgaard on an extended Kantian perspective that animals should not be treated as a mere means-to-an-end, not treated only as a resource (Korsgaard 2004).

¹⁴ The economist and moral philosopher John Rawls further limits society to citizens of a sovereign state who "share a commitment to a norm of fairness" (Runciman).

¹⁵ In A.D. 184, The Roman Emperor Commodus (161-192) played gladiator in the Colosseum. "The dens of the amphitheater disgorged at once a hundred lions; a hundred darts from the unerring hand of Commodus laid them dead as they ran raging around the *Arena*." ([Gibbon](#) 1906 edition)—all for the amusement of himself and the citizens of Rome. But not everyone was pleased.

exclude foreigners isn't part of their expertise. So, a fundamental contrast between welfare economists and other welfare consequentialists is welfare economists assert that they have no expertise, or opinion, as to which humans count. Most NotOnlyWB ethicists argue that all humans should count.

Next, I summarize a few ethics and contrast them with WC and welfare economics. While there is a logic to the order in which I present them, feel free to jump around or skip a few. But get a flavor for the diversity. I start with Buddhist ethics because it shares much with welfare economics.

Is Buddhist ethics welfare economics for an economicus who suffers from the endowment effect and other common psychological quirks?^v

03132022

Buddhism is an ethic in that it outlines how you should live your life and behave towards others. It also lays out criteria for distinguishing between right and wrong behaviors—for most occasions. The following is based on my understanding of the orientation of secular Buddhism, which is also sometimes referred to as humanistic Buddhism.¹⁶ My emphasis is on how secular Buddhism is both like and different from WC and welfare economics. For Buddhists, welfare consequences are critical, but right process (*Right View* and *Right Thought*) is a path to increasing WB (reducing ill-being). WB is the goal, but right process is how you get there.

While increasing WB is right; the emphasis is on reducing *dukkha* (suffering, unsatisfactoriness). [Chapter 5, the section *Decreasing your WB by being less reactive*, summarized the Buddhist path to increasing an individual's WB (reducing their ill-being).] Here we turn to the WB of a group consisting of all beings that can experience pain where the WB of all counts equally. Quoting from a Buddhist text

¹⁶ Whether Historical Buddhism is a religion depends on how you define religion. Traditionally, Buddhism has monks, monasteries, rituals, and a belief in reincarnation, but it has always been non-theist: there is no god, or gods, that created us and oversee us. Reincarnation was the norm in Classical India. Unlike, Christianity, beliefs are questioned rather than taken on faith. Taken alone, Buddhist ethics is not a religion. Keep in mind that many individuals who self-identify as religious Buddhists, or Christians, or Jews, or Muslims behave immorally as Buddhism would define the term. A current example of bad behavior is Buddhists in Myanmar.

Through actions of body, speech, and mind, the Bodhisattva sincerely makes a continuous effort to stop all present and future pain and suffering, and to produce present and future pleasure and happiness, for all beings (Śāntideva as translated by Charles Goodman)

This sounds like the utilitarianism of Bentham. But the process, including the intention, gets more emphasis, and intention sometimes trumps the WB consequences in determining right from wrong. According to legend, a follower generously gave the Buddha food. It killed him; it was spoiled. Dying, he insisted that the act was right because the giver did not know the food was contaminated (Goodman). Given that both WB and intent matter, Buddhists are NotOnlyWB ethicists.

While suffering and unsatisfactoriness exist and not all of it can be eliminated (bad-shit happens), there is an emphasis on reducing the unnecessary suffering of you and other creatures. As summarized in Chapter 5, your ill-being will diminish if you understand impermanence and no-self. This will reduce your craving and make you react more appropriately to the remaining ones. Thousands of years before the neuroscientist Kent Berridge found separate neural pathways for wanting/desiring and liking, Buddhists built their ethic on the distinction: asserting that seeking and getting what you want isn't the way to greater WB. Buddhists believe the default, but wrong criterion, is to base your ordering on wants and desires rather than liking: Assumption 9b rather than 9a:

Assumption 9a: An economicus's ordering of paths is based on its WB (well-being) ...

Assumption 9b: An economicus's ordering of paths is based on its wants and desires.

Buddhism outed *loss aversion* thousands of years before Kahneman and Tversky labeled the term. Buddhist impermanence (everything is impermanent: everything will be lost) makes loss central to suffering. Given loss aversion, if one wants to minimize suffering, avoid the lesser gains associated with acquiring things. Buddhism recognized that our (false) sense of self affects how we value what we possess and how what we possess affects our sense of self (ownership/self effects). It argues that these associations lead to false choosing and, in turn, suffering (less WB).¹⁷ As discussed in Chapters 6 and 7, while *loss aversion* and *ownership/self*

¹⁷ According to Buddhism, this aspect of false choosing is based on one's sense of self and the constructs one uses to assess one's self-image (*self construal*, see Chapter 7). False choosing is caused by independent self-construal (the Western tendency). If the notion of a permanent and unchanging self is abandoned for *no-self* (*anata* in Pali) ownership/self effects diminish.

effects cause false choosing, neither of these common quirks specifically violates Assumption 9a.

As noted in Chapter 5, a difference with NBT is that the Buddhist view of human behavior permits the individual to forsake their highest-ranked available path, HRAP, violating Assumption 7 of NBT. While the individual is inclined to choose their HRAP in terms of wants and desires (Assumption 9b), they are not required to. Neither must they choose their HRAP in terms of their personal WB. So, Buddhist ethics includes how to live an ethical life. As emphasized earlier, this question is off the table in welfare economics based on NBT.

Buddhist practice is geared toward affecting the WB of the practitioner and others. First, practice is geared to the practitioner learning that constantly reacting to wants and desires is a route to less rather than more personal WB. Why? Because we are loss averse, and everything acquired will be lost. Recommendations for increasing your WB include: like what you have while you have it, don't buy things on credit that will be repossessed (you will suffer loss aversion), don't pursue and obtain things and relationships simply because you desire them, and be aware of impermanence.

Buddhism, like the Bible, has rules for how to behave towards others (refrain from killing (humans and other creatures), stealing, sexual misconduct, lying, and drunkenness). But rather than obeying them to please God, the Buddhist notion is that following them will increase your WB and the WB of others.¹⁸ E.g., lying stresses the liar; besides, lying causes trouble (sounds like Epicurus). It is recognized that behaving in a Buddhist manner is something to aspire to, not a commandment without exception.

To the extent that an individual can increase their own WB without decreasing the WB of others, doing so is Buddhist morally-preferred. While Pareto improvements, P.I.s, are not part of Buddhist terminology, it seems they would advocate for them to the extent they are possible. Complicating, but not negating, is whether your WB is separable from mine.

¹⁸ Their purpose is to increase WB—akin to *rule utilitarianism*, so unlike the Ten Commandments where the objective is fulfilling God's rules. Buddhists who believe in reincarnation might worry that treating others badly will generate bad karma, causing them to come back as an animal with low WB, a reason why a Buddhist who has not achieved enlightenment might behave ethically.

Note the adjectives “your” and “mine”: no-self means there is no real you or me. Rather than constant and enduring, you are simply your current thoughts. And they are only weakly connected to your past and future thoughts and experiences, much as your experiences are linked to the experiences of other creatures. Suffering becomes less personal; the connection between you and your future “self” is no stronger than the link between you and others. The effects of your actions on your future self are no more relevant than their effects on others, both now and in the future.¹⁹ So, suffering is ownerless for the enlightened, making their objective to decrease aggregate suffering (increase aggregate WB).²⁰

For the individual practicing Buddhism, ethics is a progression: first, there is the realization that the typical path taken involves much unnecessary suffering; one then learns to minimize their unnecessary suffering. Farther along with the progression, one’s self-concept expands and evaporates: one becomes less of a permanent individual, and suffering becomes ownerless. Ideally and eventually, you can’t separate your WB from the aggregate WB of all living creates, so live your life to maximize aggregate WB by reducing suffering—and you like living this way.

In contrast to NBT, Buddhism teaches that you can, with effort, choose your ordering of paths. Hopefully, it converges to a Buddhist ordering: an ordering not based on attachment to desires but based on impermanence and ignoring the self—it would skew towards an ordering based on an integrated and aggregate view of WB.

Buddhism is consistent with emotion-specific orderings and the *emotional empathy-gap* (see Chapter 6): part of impermanence is the impermanence of your current emotional state. This means your ordering is impermanent, which violates the spirit of Assumption 4. In addition,

¹⁹ Economists use discount rates as to mathematically specify the extent to which you discount your future WB relative to your current WB. They don’t consider your discount rate for the WB of others relative to your personal WB. [If asked, they might suggest it is, of course, infinity.] In contrast, a Buddhist might argue that the two rates should be the same because your future self is not more connected with you than you are to others. It isn’t clear to me whether a Buddhist would discount their future WB more or less than a typical Western consumer. Impermanence and lack of self suggest a high rate of discount—the older you isn’t the same self you are today—so why worry about him. On the other hand, a concern for the WB that others experience now and, in the future, might suggest a low personal discount rate.

[Robert Wright](#)’s Chapter 7 in *Why Buddhism is True* specifically mentions personal discount rates but does not discuss what rate a Buddhist would use, rather the chapter discusses how one’s discount rate can be manipulated. I continue to look for articles specifically on discounting and Buddhism.

²⁰ Buddhism argues that benevolence follows, logically, from the premise of no-self, but not everyone buys their logic (see, e.g., Siderits 2007)

chapter 5 argued that the Buddhist view of humans violates Assumption 7. But, other than those two, no other NBT assumptions are violated. But the emotional empathy-gap does lead to *flawed choosing*.

That said, you could be both a true-blue economicus (adhere to all of the assumptions of NBT) and follow a Buddhist path. But your ordering would be different from that of the average Western consumer: it would be a Buddhist ordering based on interdependent, rather than independent, self-construal. But that is OK because NBT is silent on the source or your ordering, and welfare economics is OK with any ordering. You would view work and leisure as two sides of the same coin (the concept of *Right Livelihood*). Work, done correctly, would be a way towards greater WB. This is quite different from the assumption that having more is the way to greater WB and that work is simply the unpleasantness you must endure to consume more. Right livelihood is described as having a job where you have goals to achieve that are within your reach but not easy; you have autonomy and responsibility; you are part of a greater whole, a team;²¹

Higher-ranked paths would involve fewer negative *external effects* and more positive ones. But given the Buddhist view of no-self, “external” is not the best adjective (external to what?). Paths that involve making or using alcohol, drugs, or guns would not be high in your ordering because the Buddhist view is that their production and consumption produce suffering. Many Buddhists would also include producing and eating meat, arguing that killing and eating animals is neither good for them nor us.²² In this regard, Buddhism predates both Jeremy Bentham and Peter Singer, the two names most likely to come up if one is asked to name a utilitarian.

In Chapter 5, flow was described as a type of happiness: pursuing goals and activities that are challenging but not too challenging, causing you to enter a state of flow (losing your sense of self and time, being in the zone, a feeling retrospectively described as happiness). Producing flow experiences is consistent with NBT and welfare economics, and Buddhism advocates for its

²¹ Your highest-ordered available path would involve work, but not 24/7, and it probably would not involve you choosing the available path with the highest income, unless your plan was to give much of it away.

²²Tibetan Buddhists, including the [Dali Lama](#), eat meat. On the high plains of Tibet, historically there was not much to eat in the winter, only sheep and yaks (Goodman). Now that more food groups are available, some Tibetan Buddhists are advocating for a more vegetarian diet.

production. But unlike the Western notion of flow, which can be achieved by activities that only produce WB for you, Buddhist flow is achieved by flow activities that enhance aggregate WB. Playing a video game by yourself can produce western flow but not Buddhist flow. Online games with hundreds or thousands interacting are another matter.

So, is Buddhist ethics welfare economics for economici who suffer from the endowment effect and other common psychological quirks? Not quite. But if the ethical objective is to maximize aggregate WB, and if one admits flawed choosing caused by common quirks (loss aversion, ownership/self effects, salience effects, empathy gaps), the policy recommendations of the two ethics align.

John Stuart Mill: a backward-looking utilitarian and a forward-looking liberalist^{vi}

I was taught that John Stuart Mill, Jeremy Bentham, and James Mill (John's father) founded what I called in Chapter 2 Benthamite utilitarianism. But that characterization is incorrect and disserves both Benthamite utilitarianism and Mill. Over his life, Mill argued as both a utilitarian (Mill 1867) and a liberalist (Mill 1859), but neither view requires the other. Mill suffered from a dissociate-personality disorder: there was Utilitus Mill, the back-looking devoted son of James Mill and the godson of Bentham, and there was the more passionate forward-looking Liberace Mill, committed to liberty; the two mostly coexisted without openly condemning the other, but rarely appeared together in public, and Liberace appeared more as Mill aged.



Figure 1, Chapter 11: this is the performer Liberace (1919-87), not Liberace Mill

Mill wrote *On Liberty* eight years before *Utilitarianism*. *Utilitarianism*, however, was a look back, "the ideas contained in *On Liberty* underpinned his future actions as both a politician and public intellectual....*On Liberty* illuminated his future" (Reeves). First, I present Utilitus Mill then Mill's liberalist ethic.

Utilitus Mill, the utilitarian?

For Utilitus Mill, right behavior, typically not achieved, is behavior directed toward increasing aggregate happiness. And such behavior can be inconsistent with maximizing personal happiness. Mill was raised as a Benthamite utilitarian but did not write *Utilitarianism* until late in life.²³ In the view of his modern biographer, Richard Reeves, Mill wrote it as a tribute to his father and Bentham, not because of any great passion for Benthamite utilitarianism.

From Mill's perspective, utilitarianism marked an end rather than a beginning. In it [Utilitarianism] he finally laid to rest the ghosts of his father and Bentham. The essay looked backwards, offering a final settling of his Benthamite accounts, and scarcely featured in his thoughts or actions in the remaining dozen years of his life... Utilitarianism illustrated Mill's past... (Reeves)

But in the honoring, Mill fundamentally deviated from Benthamite utilitarianism. Explaining, Bentham argued all pains and pleasures were commensurable, a univariate cardinal concept—to judge the morality of a policy, one only had to put all its pain and pleasure effects on the correct scale and see which way the needle moves. Mill mucked up this simplicity by stating that pleasures generated by more valuable activities (intellectual pursuits) should count more than trashy ones, making his measure of WB a multi-dimensional concept.^{24 25} Bentham utilitarianism is a single metric for judging activities. In contrast, Mill introduced a second qualitative element into the equation, raising the thorny question of who was to say which activities produce more

²³ Mill's childhood was tough: rigorously educated in an "intellectual petri dish" by his father and Bentham; it resulted in a nervous breakdown. They hoped to make him the torch bearer for utilitarianism. (Reeves).

²⁴ Mill found "superiority of mental over bodily pleasures...It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied." Would a foolish pig agree? Disagreeing with Mill, the Greek philosopher [Aristippus](#) (435-336 BC) claimed bodily pleasures are better. In 1823 Bentham, under a pseudonym, argued that the pleasures of male-with-male sex should be included in the felicity calculus, so such acts should not be illegal (Bentham (2013) [Not Paul, but Jesus: Volume III](#). Collected Works of Jeremy Bentham preliminary edition. Bentham Project, Faculty of Laws, University College London: London, UK)

²⁵ For Mill, mental pleasures are cheaper and safer than the bodily sort. Compare home in bed reading Homer to having sex with a prostitute in Hyde Park, the latter a popular Victorian pastime. "In 1857, the Metropolitan Police estimated that the capital contained 2,825 brothels and 8,600 prostitutes. Evening visitors to the Haymarket were treated to what was popularly dubbed the 'Haymarket march past,' and even Dostoevsky, in 1862, 'noticed mothers bringing their young daughters to do business' (Reeves)."

valuable pleasures. Mill preferred reading Socratic dialogues to having sex with strangers, but it is hard to find grounds for why Mill's pleasures should be the valuable pleasures (Reeves). For Mill, like Bentham, pleasure was a cardinal concept, so possibly comparable across individuals.

According to Mill, a virtuous man is a man who, with some success (intent is immaterial), increases net happiness, not necessarily the man who maximizes his own.

The standard isn't the agent's own greatest happiness, but the greatest amount of happiness altogether (Mill 1867).

For Mill, a human was free to choose any available path, so it is not constrained to experience its highest-ranked available path, HRAP (Assumption 7 of NBT).

Freedom to choose: Liberate Mill and liberalism

In contrast, John Stuart Mill's *liberalism* judges an action or policy as right if it gives the individual more freedom and wrong if it restricts it: "pursuing our own good in our own way." (Mill 1859) The classic work is his *On Liberty*. Liberalism isn't libertarianism. [Mill's role for the government goes far beyond what U.S. Republicans would accept—his advocacy of social programs has caused some to label him a socialist.]

From their license plates, we know New Hampshirers want to

Live free or die.

Liberalism differs from WC and welfare economics in that individual liberty is right, even if the individual uses it to rain misery on his head. Summarizing *On Liberty*, you should be given the freedom to flourish in the manner of your choosing **unless your actions harm others**—a moral life is a life the individual chooses and should be allowed and encouraged to choose. Autonomy—literally self-governing—gets closer to what Mill meant than does the word "liberty". When talking to a French correspondent, Mill used the phrase "l'autonomie de l'individu" to describe the central theme of *On Liberty* (Reeves). [Applied to non-human, non-domestic animals, liberalism is living and dying in their natural state.]

Liberalism as an ethic gives precedence to freedom—a process for achieving things—rather than consequences—what the individual achieves. If an economist assumes individuals are incapable of acting contrary to their WB, as economists typically do, giving you more freedom can't reduce your WB. And the line between liberalism and WC blurs.

For the modern Tea Party and a subset of Trump supporters, freedom (fewer restrictions on their choices) is the primitive and only measure of right. However, welfare economists also believe more choice is better than less, but not for its own sake. But because (1) more choice allows you more opportunities to better yourself, and (2) economicus always takes full advantage of every opportunity to better himself.

But what does it mean to say you should be given the liberty to flourish unless your actions harm others, in which case they should not be permitted? In the language of economics, when does a *negative external effect* (an act that directly and negatively affects others) rise to a harm? Quoting

*That principle is that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number is **self-protection** [bold added]. That is, the only purpose for which power can be exercised over any member of a civilised society, against his will, is to prevent harm to others. His own good, either physical or moral, isn't a sufficient warrant." Mill (1859)*

Strictly interpreting the term "self-protection", it would seem that Mill would judge the internalization of many externalities as a violation of his principle.²⁶ Reading further, Mill broadens the principle:

*Those principles, I contend, authorize the subjection of individual spontaneity to external control, only in respect to those actions of each, which concern **the interests of other people** [bold added]. ... Acts, of whatever kind, which, without justifiable cause, do harm to others, may be, and in the more important cases absolutely require to be, controlled by the unfavorable sentiments, and, when needful by the active inference of mankind. The liberty of the individual must be thus far limited; he must not make himself a nuisance to other people*

Mill excludes from his principle children and "barbarians". You get a better feel for what Mill considered harm through his examples. "Fornication...must be tolerated." meaning prostitution should be legal (Reeves). Reacting to the argument that society can limit drinking simply because it decreases security and weakens and demoralizes society, Mill angrily responded that if this is harm, then we would all, wrongly, have an interest in everyone's "moral, intellectual and physical perfection" (Reeves). So, if I get drunk and offend you with my free speech and offensive behavior, Mill would say I am at liberty to do so, but if I get in my car drunk and run you over, I have crossed the line.

²⁶ An externality is an external effect that is being produced at an inefficient amount. If the external affect is produced at its efficient amount, and that amount is greater than zero, there is an external effect but no externality. For example, second-hand cigarette smoke is an external effect (a negative one) but if it regulated or taxed so the efficient amount of second-hand is produced, there is still an external effect but there isn't an externality.

I am tempted but hesitant to summarize that for Libera Mill, your freedom ethically trumps your WB in the sense that more personal freedom and less WB trumps less freedom and more WB. And—it is OK to restrict your freedom if it leads you to harm others less. But keeping in mind that for Mill, many reductions in the WB of others don't rise to harm.

The U.S. Declaration of Independence conveys the right to life and liberty but only the right to **pursue** happiness, suggesting that the founding fathers thought liberty trumped happiness. Another interpretation is that the founders believed that giving you more liberty would always increase your WB or, at a minimum, generate more WB than if the Monarch and Church were calling the shots. The research discussed in Part I suggests that the founding fathers had it wrong if they were assuming you, and I, always know what's in our best interests.

Consider the following thought experiment on freedom to choose: You can live in World *X* or *Y*. In World *X*, you choose your highest-ranked available path. In World *Y*, all is the same, except the government forces you to consume the path of goods and services you would have chosen in *X*: your eats and drinks are the same, and what you would have chosen, but you don't have the freedom to not eat and drink them. Which world would you choose, and why? It depends on (1) how much you directly value choice²⁷ and (2) how the freedom to choose your experiences affects your enjoyment of specific goods and services. Mill, the liberalist, would judge World *X* better than *Y*, even if you would prefer being told what to do: Mill asserts that thinking and choosing are of inherent value. Mill does not address (2): the issue of how an ice-cream cone tastes might depend on whether you choose it or were required to eat it, given you would have chosen it. Economists typically ignore the possibility that the enjoyment you get from consuming a good is influenced by how you came by it. Economists don't talk much about choice being valued for its own sake, independent of what is chosen, but there isn't a thing in NBT that precludes it either. Thinking back to the definition of a path (in Chapter 1) and Assumption 9a

Assumption 9a: An economicus's ordering of paths is based on its WB (well-being) ...

Freedoms (or the lack of) are components of every path and could be a critical determinant of how you order paths. For example, if freedom were paramount to your WB, you would order

²⁷ This issue is discussed by the economist [Frank Hahn](#) (1925-2013) who argued that the exercise of choice might enhance WB.

paths with more freedoms above those with fewer, independent of their other components. Or, more nuanced, while appreciating freedom, you might be willing to trade away a particular one for more ice cream cones or less global warming. But Liberace Mill would say paths where you have more freedom are ethically better for you, even if you don't think so. It is a juxtaposition: freedom to choose is right for you—but you are not free to choose ice cream over freedom.

[An aside: Philosophers disagree on how freedom/liberty influences WB. Foundationally, one view is while specific freedoms exist (e.g., free speech or the right to an abortion), freedom itself is not well-defined. Related is what it means to have more or less freedom: maybe on Path j Fred would be freer to speak than on Path k , but on k , he would be freer to drink alcohol and divorce. So on which path is Fred freer? An answer necessarily requires that, for Fred, all types of freedoms are WB-comparable with each other. And while economists don't distinguish between *freedom* and *freedom-to-choose*, others do.

But, suppose freedom is a meaningful concept; how does more (less) freedom contribute to more (or less) WB?²⁸ Everyone agrees that more freedom means more paths are available. And that this increases the probability that there is a newly available path with more WB. Some argue that this is the only way freedom affects WB—indirectly. Mills is not in this camp.

Others believe freedom can also enhance WB simply because it comes with more available paths, even if none of these WB-dominate. You would experience more WB if you selected Path j from the set of paths, X , than if you had selected that same Path j from the set subset $X-Y$. Why might this be? The more alternatives available to you, the more influence you have on others, and you like the power and responsibility? For example, you want to marry, and if Wanda is willing, you will marry her. Consider three cases: only Wanda is willing, Wanda and Shirley are both willing, or Wanda, Shirley, and Bob are all willing. All else constant, you would rank the Wanda/Shirley/Bob path highest and the only Wanda path lowest because the more paths available, the more you can say “No” and “to be responsible for that fact that no was said.” (Hurka 1987). “The more times we can say no, the more impact we will have on the world when we choose, and, hence, the value there is in our situation.” (Carter 1995). You get WB from having Shirley and Bob as marriage options even though you would marry neither if Wanda is available—because you have more influence, particularly on Shirley and Bob. [I leave it to you to decide whether this argument is consistent with the assumptions of NBT.] When I wrote this paragraph, Biden was sworn in as the 46th U.S. President. For Biden supporters who never would have voted for Trump, having Trump on the ballot allowed them to not just vote for Biden but also to vote against Trump.

Another view is that freedom is a primary type of WB: it has independent value and, maybe, *unconditional value*. “Our freedoms have value independently of the value we attach to the particular things they leave us free to do.” (Carson 1995). Moreover, freedom has unconditional value if paths with more freedom are always ranked higher than paths with less—paths are lexicographically ordered by degree of freedom (“Live Free or Die.”).

Another view is freedom is a bearer of a type of WB, not WB-commensurable with other kinds of WB.

A familiar view is freedom is good “just because”—no further explanation is needed.

Liberace Mill would embrace many of these views, but maybe, like me, he would have trouble drawing the boundaries between them. He might argue that freedom is not a component of WB but is still important to possess.

²⁸ More freedom could also decrease WB.

Freedom and its role have drastic implications for governance, particularly in our current climate of Trump, the pandemic, and the right-wing view that liberty is paramount.^{29]}

Neither enjoyment, suffering, nor freedom is primary (deontological/duty ethics)

Numerous ethics reject determining the right or wrong of an action by how individuals are affected. They argue right and wrong aren't defined by increases or decreases in society's WB, or by whether it gives its members more freedom, but by God, or the supreme leader, or by what is natural, or by what provides an evolutionary advantage, or by rights, or by doing one's duty. Start with duty.

Kantian ethics: the universality principle and more^{vii}



Figure 2, Chapter 11: Immanuel Kant

08172022

Kant's *universality principle*, the bedrock of his ethic, is that your actions are right if you want everyone to behave like you and wrong if you would not—you would be happy to make your behavior the law.³⁰ For example, your stealing is right only if you agree with a law requiring everyone to steal. You have a **duty** to behave according to his universality principle—Kantian ethics is duty-based (*deontological*: from the Greek words for duty (*deon*) and study (*logos*)).

Kant's universality principle is similar to but not the Golden Rule ("Do unto others as you would have them do unto you."). The convicted murderer advocates the Golden Rule when

²⁹ Reading on how freedom affects WB and one's life include [Dworkin](#) (1979 and 88), Rawls (1982), Raz (1986), Hurka (1987), [Kymlicka](#) (1990), Sen (1991), and Carter (1995)

³⁰ "Act only according to that maxim whereby you can at the same time will that it should become a universal law." This is the first formulation of Kant's *Categorical Imperative*, a commandment/imperative with no exceptions for how to live your life.

he says to the judge, "Sentence me as you would have me sentence you." But he is violating the universality principle—even the murderer does not want everyone free to murder.

Kant would want us to agree/contract to live by his universality principle. Kantian ethics is a type of *moral contract*: society's members enter into contracts/agreements, often implicit, and behaviors that violate these contracts are wrong.³¹

Kantian ethics rejects the determination of right from wrong based on consequences. For example, increased WB can be moral but isn't universally moral.

Kant considered his universality principle an obligation all rational beings have an ethical duty to follow. It is the only foundation of morality, and it is derived from pure reason and logic. It is not inferred by observing the world (according to Kant, observation can tell us how people act but not how they should act). And it is derived without regard to its influence on the world. Kant was incorrect about it being the result of pure reason and logic.

Kant is the guy who said,

Fiat justitia, pereat mundus (Let justice be done, though the world perish)

For Kant, intention was paramount, and for an action to be unequivocally moral, it had to be based on the intent to do one's duty. Kissing your partner is a moral act only if you and the partner do it out of a sense of duty—how it makes the two of you feel does affect whether the kiss is moral or immoral. [This sounds close to what Sunday school taught me about sex. You have a duty to procreate, but it being fun does not make it right or wrong]

Not everyone has duties, only rational, self-regulating, autonomous individuals. So non-human animals are duty-free and unable to act morally or immorally.

Kantian ethics and ethics based on WB judge on opposing criteria. It is difficult for economists to get their heads around Kantian ethics: whether an action is right or wrong isn't based on its effect on WB, and consequences either don't matter or are secondary. For a Kantian, a policy that increases the WB of some and decreases the WB of no one (a P.I.) isn't necessarily a moral improvement. Economists must be aware that many people believe people have duties

³¹ A modern version of right and wrong based on a Kantian-type moral contract is Scanlon's book, *What We Owe to Each Other?* The premise is that an action or principle is right if can be justified to reasonable others; it is a rule of behavior we should be able to agree on. See the section "Contract Morality".

and should fulfill them. There are more living Kantian philosophers than welfare consequentialists. Haidt and others find that duty-based morality is more common amongst [conservatives than political liberals](#).

Kantian choice theory

A kanticus is similar to an economicus, except it has a superpower, the ability to not choose its HRAP. Both have a complete ordering of paths based on WB. Like an economicus, it gets to select a path from the available set, so both have free will or not (see Part II). However, a kanticus can voluntarily constrain itself from choosing its HRAP. It does so based on self-imposed rules/duties: if you are a kanticus, you can choose to resist your own WB and desires. So, “one’s preferences are not a reliable guide to proper decision making” (White).

To predict the behavior of a kanticus, the model of behavior needs to account for duties, responsibilities, and the obligations it self-imposes. Kanticus isn’t a slave to its ordering—it can act contrary to it.

Of course, an economist can dismiss the distinction between economicus and kanticus, asserting that self-imposed duties and rules are simply determinants of your ordering.³² We can get away with this because we are vague about ‘preference’, so we can argue along the following lines. “Linda likes to abuse puppies but does not because she feels it is wrong, and this feeling of wrongness is sufficiently unpleasant that her chosen path never includes puppy abuse.” Or, “She is afraid that if she abuses puppies, she will end up in Hell where she would burn forever, which would hurt like hell.” That’s the thing about NBT; there are no behaviors inconsistent with it. If I abuse, it is because my HRAP includes abusing, and if I don’t, the same is true, even if I would enjoy the experience. This becomes circular nonsense. For economists, my decision to not abuse can’t simply be because I have decided that it is unacceptable behavior, so I will myself not to do it (consciously force myself to not do it).

The Kantian economist [Mark White](#) posits a behavior model with two distinct orderings of paths: one reflecting only your inclinations, the other with the caveat that paths that violate the universality principle are not included. The latter is consistent with doing your duty. If your will is strong, you choose to use the duty-based ordering. If your will is weak, you choose to use the

³² However, this implies that we, to a degree, choose our ordering. While Kant would say you are free to choose your ethic, I imagine many economists would more likely assume it exogenous (culture, upbringing, etc.)

ordering based only on your inclinations. White defines the strength of your will as the probability you behave consistently with your duty-based ordering. Moral is choosing based on your duty-based ordering.

Two Kants:

Arguably, there are two Kants: the duty-based, textbook Kant, as just described, and a more human, caring Kant. The above Kant of Wikipedia and college-classroom fame can be pretty unappealing ("cares more about rules than about ends," "eschews all feelings," "insists on duty for its own (incomprehensible) sake," "cold fish" (Uleman). He comes across as someone who does not care about people. Uleman's book on Kant's morality describes the more human Kant.

The interpretation of Kant's moral theory that I [Uleman] offer cuts against the grain of interpretations that emphasize Kant's commitments to formal rules and rationalism... Kant was deeply committed to a kind of formalism, and was deeply committed to rationality. But if these commitments are overemphasized, or emphasized in the wrong ways, we are left with a view that is less engaging and more academic than Kant's...Kant does insist that diverse human aims and ends, to be morally acceptable, must conform to the 'form of universality'—that is, that they may be pursued only if they could be universally endorsed. Described thus, Kantian morality does not endorse any particular aim or end—it just insists that we pursue our aim and ends only if they pass a formal test. Because it does not dictate particular aims or ends, Kantian morality seems to endorse good human lives lived across circumstances and historical time and place: it seems pluralistic and inclusive. Because it insists on universal acceptability, it seems to respect the value of hearing from everyone, or at least of imaginatively trying to; by putting yourself in other people's shoes (would it be OK with everyone): it seems deeply democratic...

Kant

*Demands much more than accord between ends and aim and a certain form; he demands that we embrace, as intrinsically and ultimately good, the **free rational human will itself** [emphasis added]. Embracing the free rational will as good means organizing our individual and collective lives in ways that actively honor this good. As a consequence, Kantian morality rejects projects the ultimate object of which is to serve God, or to alleviate material suffering: these projects, for Kant, unacceptably subjugate free rational will to other ends...The free will that Kant values is one that is fundamentally legible to others, and committed to a radically shared rationality...emphasizing the aim or end—free rational willing.*

So, the primitive for this more humanistic Kant isn't the universality principle; it is *free rational-willing*: doing one's duty and following the universality principle is just the collective path to free rational-willing. Free rational-willing is Uleman's term for Kantian autonomy and is the heart of his morality. Since economists are more comfortable with the word choosing might change her phrase to *free rational-choosing*: people are rational and, in Kant's view, should live in a world where they can freely choose their actions, including the ability to not adhere to their liking nor wanting ordering. And a will is a real thing that resides in every human adult, the thing

that chooses an alternative and then causes you to do what it takes to experience that alternative; it is the choosing center—willing is choosing. Your will is free if it gets to choose "on grounds that are its own, and not on grounds given to it by something or someone external to it" (Uleman).³³

The Kantian ethic shares a primitive similar to the primitive in Millian liberalism (free choice by rational individuals should be encouraged). However, it differs in that Kant believed the way to achieve this was for each individual to rationally choose to subjugate their freedom to their duty to the universality principle. This is the logical way for everyone to consider that their actions affect others. You might think of it as Kant's universal way of internalizing externalities.

Kantian ethics, like many ethics, is dependent on man being rational and having free will, both of which are questionable. For more on Kantian ethics from an economics perspective, see *Kantian Ethics and Economics* by Mark White.

My rights (and your duty to uphold my rights)^{viii}

Rights Theory judges individual behavior assuming the individual has certain rights: your action is moral if you had the right to take that action, wrong if you lack the right; or wrong if your action violates the rights of another.

A *right* is an entitlement, a prescribed protection (Kramer). For example, a "right to life" means your life is protected from assaults by others.³⁴ If you have a right to life, others have a duty to not take your life. Rights impose duties on others: you must uphold my rights. On-the-street advocates of a particular right emphasize that right, playing down the duty side. Rights are typically thought of as unconditional in that the right-holder has the right under all circumstances.

³³Kant relied on dualism to argue that your will can be free (Uleman). Kant viewed the body as a physical thing controlled by the laws of physics and also believed that if what you will is determined by the laws of physics, what you will isn't freely willed. He also believed if what you will is constrained by reason, your will isn't free. (For example, if constrained by reason you can't freely will that three plus three equals eight.) These beliefs of Kant argue against the existence of a will that is free. But these arguments against the existence of free will are materialist. Kant got his will freed by imagining will isn't something subject to reason or the laws of physics; it is a spiritual thing—the dual of materialism. An analogy is God's Will which isn't subject to anything; otherwise, it would not be God's Will.

³⁴ A right, a protection, can be contrasted with a *liberty*, another type of entitlement: a liberty is a prescribed freedom, for example, a freedom to sing in the shower. It is a lack of a duty; as in, I am not duty-bound to keep my mouth shut in the shower.

An action is moral if the actor has the right to do it, even if it makes everyone miserable. And I behave morally if I dutifully uphold your rights, even if doing so makes me miserable.

Many rights are asserted: human rights, women's rights, right-to-life, abortion rights, civil rights, the right to remain silent, and even property rights—these are rights of individuals. [In contrast, many environmental ethicists assert every species has a right to exist (not go extinct) which is a right conveyed on the species, not on any member of the species.³⁵] Rights often conflict, for example, “The Right-to-life” versus “A Women’s Right to Choose”. A standard view amongst rights theorists is that rights are primitive and do not require justification. In this view, a right that exists only because it achieves a goal isn’t a real right. That said, there is typically a compulsion to justify a right. Justifications include God, the Laws of Nature, the interests of society, and “Just because”.

The most influential early account of rights theory is that of the 17th-century British philosopher John Locke; he argued that the laws of nature mandate that we should not harm anyone's life, health, liberty, or possessions. For him, these are our natural rights, given to us by God. Following Locke, the United States Declaration of Independence, authored by Thomas Jefferson, recognizes three foundational rights: life, liberty, and the pursuit of happiness. Jefferson and other rights theorists maintained that we deduce other more specific rights from these, including the rights of property, movement, speech, and religious expression.

Four features are often associated with moral rights. First, rights are natural, not invented or created by governments. Second, they are universal insofar as they do not vary by country. Third, they are equal in the sense they are the same for all people, irrespective of gender, race, or handicap. Fourth, they are inalienable, which means I can’t hand over my rights to another person, such as giving up my right to freedom by selling myself into slavery for tuition money for the kids.

Suppose the right is bestowed by God or the laws of nature (a natural right). In that case, there is no presumption that exercising a right makes the right-holder better off or violating

³⁵ Giving a non-human species the right to exist, independent of human welfare, and in an environment suited to their species, has drastic implications for humans. If this right exists, then humans have a duty to preserve the species, an obligation. Humans would be required to fulfill this obligation before they worry about their own WB. Such a duty perspective on animals is alien to welfare economists.

someone else's right makes them worse off—rights are simply right, and violating one is simply wrong. Bentham considered natural rights "nonsense on stilts".

Consider property rights—the right to possess and control. Most economists are big fans, but why? Locke considered the right to possess (have possessions) a natural right bestowed by God. Economists don't claim the right to own property is God-given but argue that property rights are rights justified by consequence, a WB justification, not a God or law-of-nature justification, so not a real right. Economists are not rights theorists; they think about the efficient amounts of things, not absolute rights. For example, there is an efficient amount of free speech, which might be less or more than the current amount.

Your right to free speech and your property rights constrain me—I can't consume what you own unless you offer. Rights typically increase the choice set of the right holder and decrease the choice sets of everyone else. Economists could model behavior with rights as a constraint: choose the highest-ranked path subject to the constraint that all include free speech.

Millsian liberalists abhor unconditional rights except for the right to choose. The principle of liberalism is the right to choose to live as you please, presuming your choices don't cause too much harm, so it is a conditional right. The right to liberty is problematic because I have less liberty if I must uphold your right to liberty.

Complicating, there are two views on the **function** of rights (what they do for the right holder): *will theory* and *interest theory*. Both are problematic:

The function of a right is to give its holder control over another's duty...will theorists [emphasis added] believe that all rights confer control over others' duties to act in particular ways...To have a right is to have the ability to determine what others may and may not do, and so to exercise authority over a certain domain of affairs. (Leif Wenar on Will Theory)

Interest theorists maintain that the function of a right is to further the right-holder's interests. An owner has a right, according to interest theory, not because owners have choices, but the ownership makes owners better off. (Wenar on Interest Theory)

Why is this distinction important? For *will theorists*, a right is only a right if the holder can waive it: capable of choosing to not exercise the right. This implies individuals unable to waive their rights (e.g., infants, animals, and comatose adults, have no rights. By this logic, I do not have a right to freedom because I can't legally waive my right to it under U.S. law and become your slave, and Australians don't have the right to vote because they are legally required to vote. In

contrast, *interest theorists* do not require rights to be waivable, partly because they want infants, animals, and comatose adults to have rights.

In interest theory, you are only eligible for a right if you would be made better off by having that right—you can have no right to free speech unless having it will make you better off. Those incapable of being better or worse off can't have rights. In this scenario, all living individuals capable of suffering could have a right not to suffer, but not rocks. But interest theory does not say that you have a right to everything that would make you better off. People who reject animal rights worry that interest theory argues for animal rights.

The word of God, or supreme leader^{ix}

Divine-command theory posits that God is the decider regarding what thoughts and behaviors are moral or immoral (sins). With God as the decider, cursing one's parents is immoral:

Whoever curses his father, or his mother shall be put to death. Exodus 21:17

To operationalize this ethic requires the ability to figure out God's rules. Is God cool with my becoming a suicide bomber against the heathens? Religion is the foundation of ethics for many millions, but easy to criticize. Quoting the Harvard psychologist [Steven Pinker](#), writing for the Sunday *N.Y. Times* magazine:

Putting God in charge of morality is one way to solve the problem, of course, but Plato made short work of it 2,400 years ago. Does God have a good reason for designating certain acts as moral and others as immoral? If not—if his dictates are divine whims—why should we take them seriously? Suppose that God commanded us to torture a child. Would that make it all right, or would some other standard give us reasons to resist? And if, on the other hand, God was forced by moral reasons to issue some dictates and not others—if a command to torture a child was never an option—then why not appeal to those reasons directly?

Most people are not as quick to dismiss morality based on God's Will; most are religious, at least to an extent.

The God I learned about ruled clearly on the morality of numerous activities such as parent-honoring and wife-of-neighbor lusting. I was told that God declared, he never told me this directly, "Sex before marriage, or simply for fun, is immoral." But God provides no guidance for many other activities—is it better to play baseball or read comic books?—suggesting I could decide, within limits, what was best for me—maximize my enjoyment—as long as no Commandments were violated. That said, I was taught I would be doing a lot of sinning (being immoral).

For the Christian God, intent and remorse are more important than actions: if when my time has come, I believe and ask for forgiveness, I am off to heaven no matter how many I lusted after or smote along the way. I am unsure whether this makes the Christian God a processist or a consequentialist. I am also unsure whether my childhood God would want us to save the whales or even if she cares. Religions and gods differ on the value of the environment: Buddhism, for example, would view the whales and us as inseparable parts of the system, making both whales and humans worth preserving. Morality based on the word of God is fundamentally different from ethics based on WB: sinning might make you happy and no one unhappy, but, in the eyes of God, it's immoral.

What about the god of Islam?

The will of nature^x

Actions are moral if they follow the natural order—you and your dog having sex is immoral, even if you both enjoy it; it isn't natural.³⁶ Natural-law ethics assume all men have a common human nature, and it is moral, so behavior consistent (inconsistent) with it is moral (immoral). To be moral is to do what is in your nature.

Operationalizing this ethic requires the identification of human nature (behaviors and characteristics common to all humans). Religious thinkers, philosophers, anthropologists, and primatologists have collectively spent thousands of years on whether there is a human nature and, if so, what it is. Is it bestowed on us by God (our creator), genetics, or stuff that every culture teaches? For example, while observed in a few primate groups, the making and use of tools aren't observed in all of them, so it isn't part of primate nature. On the other hand, it is part of human nature, so it's a moral activity for humans but not for chimps. Within-species sex is human nature, so moral. Is lying to and killing human nature?

³⁶ Consider the following story devised by the psychologist Jonathan Haidt (Haidt 2011): "Julie and Mark are brother and sister. They are traveling together in France on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At very least it would be a new experience. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy making love, but they decide not to do it again. They keep that night as a special secret, which makes them feel even closer to each other. What do you think about that, was it OK for them to make love?" Respondents say no. but their reasons for why it was not ok are assumed away by the story. In the end, many people simply say it is bad but can't articulate why (Pinker). The sex was a P.I. (both were better off and no one else is worse off).

Rousseau believed in human nature and thought it a moral nature, making him a proponent of Natural-law morality. Others argue that people are inherently immoral: Hobbs is the arch-type. Many assert that genetic engineering is immoral for the same reason that Frankenstein's monster was immoral; it's unnatural. Environmentalists have argued that saving nature is moral simply because nature is natural. Welfare and environmental economists are too quick to dismiss such thinking as illogical; it isn't.

This ends my brief discussion of deontological/duty ethics.

Virtue ethics^{xi}

In virtue ethics, moral versus immoral is determined by your character and your inclinations: to be moral is to have virtuous (moral) inclinations. Your behavior and its consequences are secondary. Duties and rules are secondary, and so are social contracts. If you maintained a virtuous character throughout your life, you had a moral life. Plato and Aristotle were virtue ethicists, and so was Aquinas—virtue ethics was the predominant ethic until the Renaissance. The economist historian [Deirdre McCloskey](#) argues that Adam Smith was a virtue ethicist, and much in Smith's *Theory of Moral Sentiments* supports this. Economists who have not read it might imagine Smith a WC; he was not.

In the 19th and early 20th Centuries, virtue ethics was eclipsed by Kantian ethics and WC, only to reemerge in philosophy in 1958 with the publication of *Modern Moral Philosophy* by the British philosopher [Gertrude Anscombe](#)³⁷—it had never lost its street appeal. Anscombe's article rejected WC because “it commits one to endorsing evil deeds” and concluded that Kantian ethics ... is just incoherent” (SEP on *Anscombe*).³⁸ Whether she intended to resurrect virtue ethics is debated, but resurrect it she did.

The three pillars of virtue ethics remain the Aristotelian concepts of *arête* (excellence or virtue), *phronesis* (practical or moral wisdom), and *eudaimonia* (true happiness/flourishing)—(SEP on Virtue Ethics). Virtue ethicists distinguish between primary and secondary virtues: the secondary ones are implied by the primary ones. There is disagreement about which are primary, but all agree that generosity, honesty, compassion, and courage are primary. To possess a virtue

³⁷ She coined the term “consequentialism”.

³⁸“Self-legislation” refers to idea that you should not perform an act that you would not have everyone legislated to perform—Kant's universality condition.

means that it is part of your core. Taking honesty as an example, always telling the truth is neither necessary nor sufficient to possess this virtue. Having it simply means you strive to be, for its own sake, an honest and truthful person, not that you blurt out the truth on every occasion.

Practical and moral wisdom is the knowledge and experience you need to do what you intend—for example, you know what it takes to be generous or courageous. Teenagers lack practical and moral wisdom—they intend to be virtuous but muck things up. If you possess the virtues, are experienced enough to have practical and moral wisdom, and are lucky enough to have some wealth, you flourish and experience eudaimonia.³⁹ It is easier to say what this does not mean than it does. It does not mean your WB is maximized; a virtuous person would not maximize their personal WB.

Non-humans can be neither virtuous nor non-virtuous: they are incapable of virtue: it requires the ability to evaluate your actions, so your dog can't be a bad dog, but you can be in the doghouse for lacking virtue.

Welfare consequentialists, including welfare economists, complain that virtue ethics produces no objective criteria or rules for judging behaviors and policies. Virtue ethicists agree but think pursuing an objective, such as increasing societal WB, is folly.

Looking back to Buddhist ethics, [Damien Keown](#), Professor of Buddhist Ethics at the University of London, thinks Buddhism is a type of virtue ethics, arguing that following the Buddhist way makes one a virtuous person, one of its end goals. The debate in Buddhism is whether it is an end goal or simply an effective way to reduce suffering (increase WB).

Evolutionary morality theory^{xii}

An ethic could be based on evolutionary biology: genes for being good: see Richard Dawkins (1976) and Steven Pinker (2008). We get along, and most of us don't behave in ways that drastically reduce the WB of others—at least not all the time—because gene containers that get along are more likely to pass their genes on to the next generation. It seems that empathy and *reciprocal altruism* are in our genes. Empathy is the ability to put yourself in another's shoes. Reciprocal altruism is when I do something to help you even though it temporally weakens me: I do it because I expect you to return the favor. Empathy contributes to our caring for others (those weaker), and reciprocal altruism contributes to our sense of fairness and justice. [In a related

³⁹See the discussion of eudaimonia in Chapter 4.

vein, the economist and game theorist [Ken Binmore](#) in his book *Natural Justice* argues that the ethics a society practices result from cultural, rather than biological, evolution, the solutions to repeated games.]

Contract morality (*contractualism*)^{xiii}

Using an agreed-upon process, members of a society enter into contracts/agreements with one another; a few are explicit (e.g., written law against killing people), but many are implicit. Behavior that violates a contract is wrong/immoral. Both Rawls and Kant can be viewed as contractualists: Rawls is more a political than a moral contractualist; Kant is a bit of both. Kant and [Thomas Scanlon](#), the modern voice of contractualism, both argue that the contract is motivated by everyone's mutual respect, and it is rational to respect others. Quoting from Scanlon's influential 1998 book *What We Owe to Each Other*:

An act is wrong if its performance...would be disallowed by any set of principles [rules]...of behavior that no one could reasonably reject...

An act is wrong if no reasonable person could argue that the rule against it isn't justified.

Morality for Scanlon is about our contractual relationships (explicit and implicit) between rational individuals. In his review of Scanlon's book, the philosopher Thomas Nagel describes the idea behind Scanlon's contractual rule.

The idea is that if our aim is to be able to justify our conduct to others, we will want it to conform to principles that none of them could reasonably reject, because then everyone who shares our interest in justification would in effect be prepared to license what we do insofar as it accords with those principles. If we deliberately do something that is in this sense wrong, we are in effect saying we don't care about its admissibility to reasonable others.

Contractualism differs from WC and welfare economics. It is only concerned with how we behave towards others, while WC is also worried about how we treat ourselves. WB does not play the leading role in contractualism. Its objective isn't to maximize societal WB, and it does not require that all kinds of WB are commensurable.

In contrast to contractualism, Hobbs (*Leviathan* 1651) believed that you are motivated by self-interest but would commit to a social contract limiting what you can do, but only because it would limit the bad shit other people could do to you—such as shooting or shunning you. The economic Nobel laureate [James Buchanan](#) (1919-2013) thought the same: entering into a social contract that outlines good behavior is motivated by self-interest rather than respect. To distinguish it from the mutual-respect camp, this camp is referred to as *contractarianists*. The

contractarian philosopher [David Gauthier](#) believes you want to maximize only your WB but that this is sometimes achieved by behaving in ways that will make you worse in the short run. For example, you don't break your promises or bully even though it would increase your current WB; behaving ethically is a good practice because it will influence how others deal with you in the future when you want something from them.

Ethics based on group rights, things having rights, and inherent value^{xiv}

Most ethics convey moral standing to individuals only; in which case, a group of individuals only has moral standing because its members have it. E.g., consequentialists want to maximize an aggregation of the members' WB. Contrast that assumption with the group has moral standing or value, but the individual members of the group do not. Imagine the group holds rights or that the objective is to maximize the WB of the group, but the group's welfare isn't an aggregate of the WB of its members. The group is the entity with rights.

For example, you might posit that a non-human species has the right to not go extinct or the right to prosper in its natural state. E.g., elephants have the right to exist and live undisturbed in their natural habitat. Such a *group right* conveys neither rights nor moral standing to specific elephants. In fact, the prospering of the species in its natural environment might require the suffering of all its members: wild animals continuously struggle to live. Nevertheless, this group right conveys a duty towards elephants, making us wrong if we do not provide them with sufficient natural habitat.⁴⁰ [There is the question of what amount of natural habitat is adequate.]

Or you might posit that future humans have a right to exist with sufficient resources, so current humans must make this happen. Given such a right, we must provide for them even if no future individual has rights. Or you might posit that an ethnic, cultural, or religious group has the right to exist without giving moral standing to any of its specific members.

If it is the group, rather than the individuals in the group, that has moral standing, it is difficult, at least for an economist, to talk about the WB of the group, but it is possible. One, for

⁴⁰ In 1981, in *Palila* (an endangered species of bird) vs. the Hawaii Department of Land and Natural Resources, U.S. Court of Appeals, the species successfully sued for habitat protection under the Endangered Species Act, effectively granting the species limited legal standing. Their habitat was being destroyed by feral goats and the species successfully sued for their removal.

https://scholar.google.com/scholar_case?case=8158542596159694344&hl=en&as_sdt=2&as_vis=1&oi=scholar

example, might consider the WB or integrity of a family that has a right to exist and prosper independent of the WB of its members. Still, economists typically don't think like this.

There is the question of which groups get to have group rights. At a minimum, the group has to have a persistent identity. For example, a crowd is a group of people who do not have a persistent identity. In contrast, Canadians, fetuses, squirrels, and cloggers are groups with persistent identities. There are lots of groups with persistent identities.

Equally alien to economists, ecological systems (e.g., the Everglades or a redwood forest) and natural formations (e.g., the Grand Canyon) could have rights.⁴¹ "It is good to preserve and maintain the Everglades and immoral to destroy or diminish it." In this view, you behaved immorally if you turned part of it into a golf course. The Wisconsin forester [Aldo Leopold](#) wrote in 1949 in his now-famous [A Sand Country Almanac](#)

A thing is right when it tends to preserve the integrity, stability and beauty of the biotic community. It is wrong when it tends otherwise.

He did not explain why the biotic community has a right to be beautiful. With this perspective, the suffering or pleasure of an individual animal or plant has no relevance. Leopold's reason is neither an animal-rights justification nor a justification based on preserving it for humans. This designation of value is most expressed in environmental ethics, applied to natural entities. Still, it could be applied to a built entity such as the Sistine Chapel, the German language, or Venice

⁴¹ *Should trees have standing?—toward legal rights for natural objects* (Christopher Stone 1972) is the foundational legal treatise for affording them legal standing. In a 1972 Supreme Court dissenting argument, echoing Stone, [Justice William Douglas](#) famously wrote, "Inanimate objects are sometimes parties in litigation. A ship has a legal personality, a fiction found useful for maritime purposes. The corporate soul—a creature of ecclesiastical law—is an acceptable adversary, and large fortunes ride on its cases. The ordinary corporation is a 'person' for purposes of the adjudicatory process...So it should be as respects valleys, alpine meadows, rivers, lakes, estuaries, beaches ridges, groves of trees, swampland, or even air that feels the destructive pressures of modern technology and modern life...The river, for example, is the living symbol of all the life it sustains or nourishes..., including man who are dependent on it or who enjoy it for its sight, its sound, or its life. The river as plaintiff speaks for the ecological unit of life that is part of it. Those people who have a meaningful relationship to that body of water, whether it be a fisherman, a canoeist, a zoologist, or a logger—must be able to speak for the values which the river represents, and which are threatened with destruction." (https://en.wikisource.org/wiki/Sierra_Club_v._Morton/Dissent_Douglas). The case was the "Sierra Club v. Morton": the Sierra Club was trying to stop the Disney Corporation from building the Mineral King ski area in Sierra Nevada wilderness. The Court dismissed the suit concluding that neither the Sierra Club nor the wilderness had legal standing. See, Elizabeth Kolbert (2022). The ski area was never built but the proposal led to some of my initial research.

At the moment (May 2022): Mary Jane Lake in Florida, along with a few other water bodies, has filed suit in Florida State Court arguing that a planned development of homes and office buildings (Beachline South Residential), which will fill in wetlands, will "adversely impact the lakes and marsh that are parties to his action," causing injuries that are "concrete, distinct, and palpable." (Kolbert 2022 and Smith 2022).

(“Venice has value and should be preserved because it is unique and a remarkable creation.”). A justification given for attributing value to a natural or environmental resource is that, unlike a truck, it did not come into existence to make humans better off (Brennan 1984). Another justification was put forward by the philosopher [Richard Routley/Sylvan](#). He hypothetically asked if it would be OK for the last human to destroy, for fun, the Sistine Chapel and the Grand Canyon. He thought most of us would conclude it was wrong to push the button that would blow them up, even though there would be no humans around to experience them.

The common denominator is *inherent value*: a thing or group has value in and of itself, independent of how any living individual (human or otherwise, including that thing) experiences the world: it is a value that no individual experiences, *experience-less value*. Economists understand experiential value (value experienced) but don’t understand or believe in inherent value—for them, it makes little sense.⁴² Consider a tree. Assuming the tree is incapable of experiencing, a welfare economist would say the tree only has value if its existence can influence the experiences of those individuals who have societal standing. For example, I have standing, and the tree makes for a lovely view from my home, or it can be used to build an addition or sequester carbon dioxide, making the earth more pleasant for society's members.

⁴² I must qualify: economists believe that humans have inherent value, but they believe nothing else does.

I was tempted to use the word “intrinsic” rather than “inherent” but will stick with “inherent” because while inherent value is a type of intrinsic value in the philosophical sense, it isn’t the only possible type of intrinsic value. “Intrinsic” in philosophy typically means axiomatic; as in the hedonistic assumption that pleasure is the only determinant of right, and pain is the only measure of wrong. This implies, for example, that ice cream can’t be intrinsically moral or immoral because ice cream is neither pain nor pleasure, but it could be extrinsically moral if its consumption produces pleasure, the only intrinsic moral good. Acts and things only have extrinsic value if they produce intrinsic value. Not everyone agrees on what are the intrinsic values. The inherent value of an object or system is a type of intrinsic value because its existence is axiomatically defined as morally right. The philosopher [G.E. Moore](#) (1873-1958) suggested that when determining whether something was intrinsically valuable, you consider whether it would be valuable if it were the only thing in existence. If you answer yes, you think, according to Moore, that thing has intrinsic value (SEP on *Intrinsic vs. Extrinsic Value*). If you would say that pleasure (or a species) is valuable even if it is the only thing that exists, then you believe pleasure (or that species) has intrinsic value. Not everyone buys Moore’s test. What is intrinsic value and what has intrinsic is a prominent topic in moral philosophy.

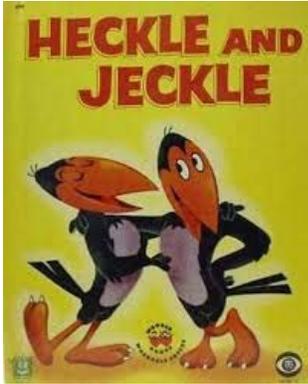


Figure 3, Chapter 11: Heckle and Jeckle

Or the tree has value because [Heckle and Jeckle](#), both card-carrying members of society, nest in it.

If those with standing no longer care about the tree, or all die, or leave the planet, the welfare-economic view would be that the tree no longer has value. But the tree still could have inherent value. If inherent value exists, it is independent of whether people believe in it; it isn't determined by us.⁴³

Making the distinction clearer: if society's members get value, for example, from simply knowing that an ecosystem is thriving, this is economic value, not inherent value. Why? Because one or more individuals experience the value. Above I assumed the tree is incapable of experiences, but a polar bear can experience pleasure and pain. If a bear is a member of society, their experiencing pleasure is an economic value, not an inherent value. In contrast, the inherent value of a bear isn't something the bear or anyone can experience.

Note the distinction between a species' inherent value and its right to exist. Either can be used to motivate the preservation of polar bears, but they are not the same reason: a "right" is simply the "right thing to do"; it does not imply that doing the right thing needs value. If a group has inherent value, this could be used to argue for the group's right to exist (not go extinct). But a thing or a group could have a right to exist even if it has no inherent value. For example, one

⁴³ God has been invoked to motivate inherent value, as in "Polar bears have inherent value because God values the creatures he created." But this is faulty reasoning. If God is an individual who experiences something because his creatures are prospering, then his experience has economic value—if God has moral standing.

might posit that it would be wrong to destroy a distant star simply because it is part of the physical universe.

When I teach environmental economics to environmental-studies majors, many believe the environment and planet have inherent value. And they reject the idea that the management of environmental resources should be based only on the values experienced by individuals. Even though I have sometimes accused them of being crazy, their belief is logical if one starts with the axiom that environmental resources have inherent value.

All this said about ethics:^{xv}

In the 18th Century, Hume pointed out that while there are reasons supporting different ethics, there is no one correct ethic based on "Reason". As Mill said,

What proof is it possible to give that pleasure is good? (1867)

The bottom line is there is no logical reason to prefer one ethic over another. It all comes down to how moral and immoral are axiomatically defined.

For example, there can be no logical argument against axioms such as "We should improve the environment because improving the environment is the right thing to do." Or "Giving individuals with moral standing more freedom to flourish is morally right." Or "The only metric of right behavior is doing one's duty." Ethical axioms can't be judged on the basis of logic.⁴⁴

But facts influence how to implement an ethic. For example, as noted earlier, a WB consequentialist would advocate complete freedom for Robinson Crusoe if they believe he will always choose his HRAP, but advocate interference in Robin's life if they believe he will exhibit flawed choosing.

You can be an economist modeling choice without being a WB consequentialist. You can also speculate on the role of choice theory adopt a different ethic, as in Mark White's book.

⁴⁴ Axioms can be judged on whether they contradict fact, but axioms that define the essence of an ethic are typically not of that sort. An ethic based on a set of axioms can be judged wrong if those axioms are self-contradictory. E.g., if one axiom is more freedom is the only measure of better and another is more happiness is the only measure of better.

Now that I have briefly reviewed other ethics, Chapter 11, the next to last, returns to welfare economics, getting specific about its limitations as an ethic. Finally, chapter 11 ends with a brief discussion of economic ethics that are not welfare economics.

ⁱ Sources:

1. Bernard Williams (1985, 2006) *Ethics and the Limits of Philosophy*. Oxford: Fontana Press 1985, Routledge 2006.

ⁱⁱ Sources:

1. John Stuart Mill (1867) *Utilitarianism, third edition*, London, Longmans, Green, Reader, and Dyer. At <http://books.google.com/books>

ⁱⁱⁱ Sources:

1. John Harsanyi (1982) Morality and the theory of rational behavior, in *Utilitarianism and Beyond* (eds. Amartya Sen and Bernard Williams), Cambridge, Cambridge University Press. Originally published in 1977 in *Soc. Res.* 44(4).
2. L.W. Sumner (1996) *Welfare Happiness and Ethics*, Oxford, Oxford University Press

^{iv} Sources:

3. Thomas Aquinas (1265-74) *The Summa Theologica*, translated by Fathers of the English Dominican Province, Benziger Bros. Edition, 1947. At <http://www.ccel.org/a/aquinas/summa/FS.html>
4. Thomas Eisner (2003) *For Love of Insects*, Cambridge: Harvard University Press.
5. Edward Gibbon (1906 edition) *The Decline and Fall of the Roman Empire .Vol. 1: Chap. 4.* ed. J.B. Bury with an Introduction by W.E.H. Lecky (New York: Fred de Fau and Co.), in 12 vols. Vol. 1. <https://oll.libertyfund.org/title/gibbon-the-history-of-the-decline-and-fall-of-the-roman-empire-vol-1>
6. Christine Korsgaard (2004) Fellow creatures: Kantian ethics and our duties to animals, in *The Tanner Lectures on Human Values: Volumes 25/26* (ed. G.B. Peterson), Salt Lake City, University of Utah Press.
7. Christine Korsgaard (2007) Facing the animal you see in the mirror, paper presented at the conference "Facing Animals" at Harvard University.
8. Richard Kraut (2007) SEP on *Aristotle's Ethics*. <http://plato.stanford.edu/entries/aristotle-ethics/>
9. P. Kreeft (1990) *Summa of the Summa*. Ignatius Press.
10. Martha Nussbaum (2007) Facing animal complexity, paper presented at the conference "Facing Animals" at Harvard University
11. Nussbaum (2012) Who is a happy warrior? Philosophy, happiness research, and public policy, *Intern. Rev. of Econ.* 59: 335-361.
12. W.G. Runciman (2014) The kind of dog he likes, *London Rev. of Books* 36(4): 29-30.
13. Peter Singer (1975) *Animal liberation: a new ethic for our treatment of animals*, N.Y., Random House.
14. Peter Singer (2011) *Practical Ethics, third edition*, Cambridge, Cambridge University Press.
15. Lawrence Wright (Feb. 28, 2022) The elephant in the courtroom, *The New Yorker*.

^v Sources:

1. Stephen Batchelor (2015) *After Buddhism, rethinking the dharma for a secular age*, Yale University Press. <http://yalebooks.yale.edu/book/9780300205183/after-buddhism>
2. Clair Brown (2017) *Buddhist economics: an enlightened approach to the dismal science*. New York, Bloomsbury Press
3. Rupert Gethin (1998) *The foundation of Buddhism*, Oxford, Oxford University Press.
4. Rupert Gethin (2008) *Sayings of the Buddha*, Oxford, Oxford University Press.
5. Charles Goodman (2014) SEP on *Ethics in Indian and Tibetan Buddhism*. <http://plato.stanford.edu/archives/fall2014/entries/ethics-indian-buddhism>
6. Walpola Rahula (1959) *What the Buddha taught*, N.Y., Grove Press.
7. Damien Keown (1992) *The Nature of Buddhist Ethics*, N.Y., Palgrave.
8. Damien Keown (2005) *Buddhist ethics: a very short introduction*, Oxford University Press.
9. Joel Magnuson (2022) *The Dharma and the Socially Engaged Buddhist Economic*. Palgrave Macmillan.
10. Mark Siderits (2007) *Buddhism as philosophy: an introduction*. Indianapolis/Cambridge, Hackett.
11. Mark Siderits (2019) *Buddha*, SEP

12. Robert Wright (2017) *Why Buddhism is true: the science and philosophy of meditation and enlightenment*, Simon and Shuster.
13. Laszlo Zsolnai (2009) Buddhist ethics for business, in *Ethical prospects: economy, society and environment* (eds. Zsolt Boda, Laszlo Fekete and Laszlo Zsolnai), Dordrecht, Springer. <https://doi.org/10.1007/978-1-4020-9821-5>

vi Sources:

1. Ian Carter (1995). The independent value of freedom. *Ethics* 105: 819-45.
2. Ronald Dworkin (1979) Do we have a right to liberty. n *Liberty and the Rule of Law* ed. R.L. Cunningham College Station Texas, Texas A&M University Press
3. Frank Hahn (1982) On some difficulties of the utilitarian economist, in *Utilitarianism and Beyond* (eds. Amartya Sen and Bernard Williams), Cambridge: Cambridge University Press.
4. Thomas Hurka (1987) Why value autonomy? *Social Theory and Practice* 13(3): 361-82.
5. Will Kymlicka (1990 and 2002). *Contemporary political philosophy*. Oxford: Oxford University Press.
6. John Stuart Mill (1859) *On Liberty*, eBooks@Adelaide 2009 at http://ebooks.adelaide.edu.au/m/mill/john_stuart/m645o/ with the text from *Harvard Classics, Vol. 25*, P.F. Collier and Son, 1909.
7. John Stuart Mill (1879) *Utilitarianism, third edition*, London, Longmans, Green, Reader, and Dyer. At <http://books.google.com/books?id=FtkIcBOZuCsC>
8. John Rawls (1971, revised 1999) *A Theory of Justice*, revised edition, Harvard University Press.
9. Joseph Raz (1986) *The Morality of Freedom*. Oxford: Clarendon Press.
10. Richard Reeves (2007) *John Stuart Mill: Victorian firebrand*, London, Atlantic Books.
11. Amartya Sen (1988) Freedom of choice: concept and content. *European Economic Review* 32: 269-94.

vii Sources:

1. Immanuel Kant (1785, 2003) *Groundwork of the metaphysics of morals* (ed. Mary Gregor), Cambridge University Press. At http://books.google.com/books/about/Kant_Groundwork_of_the_Metaphysics_of_Mo.html?id=Ibcy57Hz4tcC
2. Nicholas Southwood (2009) Moral contractualism, *Phil. Compass* 4(6): 926-37.
3. Jonathan Haidt (2012) *The righteous mind: Why good people are divided by politics and religion*, N.Y., Pantheon Books.
4. Mark White (2011) *Kantian Ethics and Economics: Autonomy, Dignity, and Character*, Stanford, CA: Stanford University Press.
5. Jennifer Uleman (2002) *An Introduction to Kant's Moral Philosophy*, Cambridge, U.K., Cambridge University Press.
6. Elizabeth Ashford and Tim Mulgan (2012) SEP on *Contractualism*. <http://plato.stanford.edu/entries/contractualism/>

viii Sources:

1. Matthew Kramer (2008) Matthew Kramer on Legal Rights, Philosophy Bites. <http://philosophybites.com/2008/07/matthew-kramer.html>
2. The Internet Encyclopedia of Philosophy, on *Ethics* at <http://www.iep.utm.edu/>
3. Leif Wenar (2011) SEP on *Rights*. <http://plato.stanford.edu/entries/rights/>

ix Sources:

1. Steven Pinker (2008) The Moral Instinct, *The N.Y. Times*, January 13.

x Sources:

1. Jonathan Haidt (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment, *Psych. Rev.* 108: 814-834.

xi Sources:

1. Rosalind Hursthouse (2013) SEP on *Virtue Ethics*. At <http://plato.stanford.edu/entries/ethics-virtue/>
2. Deirdre McCloskey (2008) Adam Smith, the Last of the Former Future Virtue Ethicists, *Hist. of Political Econ.* 40(1): 43-71. Also in Jeffrey Young, ed., *The Elgar Companion to Adam Smith*, 2010. At <http://www.deirdremccloskey.com/docs/smith.pdf>
3. Julia Driver (2014) SEP on *Gertrude Elizabeth Margaret Anscombe*. <http://plato.stanford.edu/entries/anscombe/>
4. Adam Smith (1767) *The theory of moral sentiments*. At <https://play.google.com/store/books/details?id=-V4AAAAAMAAJ&rdid=book--V4AAAAAMAAJ&rdot=1>

5. Samuel Fleischacker (2013) SEP on *Adam Smith's Moral and Political Philosophy*.

<http://plato.stanford.edu/entries/smith-moral-political/>

6. Damien Keown (1992) *The Nature of Buddhist Ethics*, N.Y., Palgrave.

^{xii} Sources:

1. Richard Dawkins (1976) *The Selfish Gene*, N.Y.: Oxford University Press. A famous book.
2. Pinker (2008) *ibid*.
3. Robert Trivers (1971) The evolution of reciprocal altruism, *Quart. Rev. of Biology* 46: 35-57.
4. Frans De Waal (2008) Putting altruism back into altruism: the evolution of empathy, *Annual Rev. of Psych.* 59: 279-300.
5. Ken Binmore (2005) *Natural Justice*, N.Y., Oxford University Press.

^{xiii} Sources:

1. Nicholas Southwood (2009) *ibid*.
2. Elizabeth Ashford and Tim Mulgan (2012) *ibid*.
3. Thomas Scanlon (2000) *What we owe each other*, Cambridge MA: Harvard University Press. The introduction is at <http://www.nytimes.com/books/first/s/scanlon-owe.html>
4. Thomas Nagel (1999) One-to-one, *London Rev. of Books* 21(3,4):3-6.
5. Ann Cudd (2102) SEP on *Contractarianism*. <http://plato.stanford.edu/entries/contractarianism/>
6. Dick Arneson, David Gauthier's *Morals by Agreement*, class notes accessed April 7, 2015 from <http://philosophyfaculty.ucsd.edu/faculty/rarneson/Courses/160NotesGauthier.pdf>

^{xiv} Sources:

7. Andrew Brennan (1984) The moral standing of natural objects, *Envir. Ethics* 6: 35-56.
8. Andrew Brennan and Yeuk-Sze Lo (2011) SEP on *Environmental Ethics*. <http://plato.stanford.edu/entries/ethics-environmental/>
9. William Grey (2009) Last Man Arguments, *The Encyclopedia of Environmental Ethics*, Vol 2: 40-41.
10. Agnieszka Jaworska and Julie Tannenbaum (2013) SEP on *The Grounds of Moral Status*. <http://plato.stanford.edu/entries/grounds-moral-status/>
11. Peter Jones (2014) SEP on *Group Rights*. <http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=rights-group>
12. Elizabeth Kolbert (April 18, 2022) Testing the waters: should the natural world have rights, *The New Yorker*.
13. Aldo Leopold (1949) *A Sand Country Almanac*, Oxford, Oxford University Press.
14. Wesley Smith (April 12, 2022) Natural rights: a lake sues, *National Review*.
15. Christopher Stone (1972) Should trees have standing?—toward legal rights for natural objects, *Southern California Law Review* 45: 450-501.
16. Michael Zimmerman (2014) SEP on *Intrinsic vs. Extrinsic Value*. <http://plato.stanford.edu/entries/value-intrinsic-extrinsic/>

^{xv} Sources:

1. Mill (1867) *ibid*.
2. John Hicks (1939) *Value and capital: an inquiry into some fundamental principles of economic theory*, Oxford: Clarendon Press.
3. Mark White (2011) *ibid*.

Chapter 12: A bit more on welfare economics, external effects, and interpersonal comparisons of WB.

08172022

Chapter 2, which introduced welfare economics, briefly discussed external effects. External effects are the direct WB (or want fulfillment) effects on others of your actions. These effects are external to you (e.g., the second-hand smoke from your cigarette decreases my WB, the second-hand smoke from your joint increases my WB).

However, Chapter 2 did not detail their practical significance in judging specific behaviors, institutions, and processes. For example, it did not address the common belief that unregulated competitive markets are the ethically-best mechanism for determining what is produced and who consumes what. I now address this question.

Many economists, including many welfare economists, and many non-economists, advocate for free markets and freedom of choice. But their reasoning is often tautological. For example, if efficiency is the goal, and one makes assumptions that guarantee that unregulated competitive markets are efficient, unregulated competitive markets achieve the goal.

To keep things simple, initially assume Assumptions 1-9 hold, and there are neither incorrect beliefs nor flawed choosing—an idealized world in which to judge welfare economics as an ethic. But, even then, it has problems.

The ethical goal of welfare economics is maximizing societal WB, not achieving efficiency

According to welfare economics, actions that increase the WB of some members but don't decrease the WB of other members are right (they are P.I.s). And actions that decrease the WB of some members but don't increase the welfare of other members are wrong (Pareto Deteriorations).

In contrast, an act is *efficiency increasing* if those whose WB is increased could hypothetically compensate those whose WB was decreased. If the compensation occurred, no one's WB would be lower, and some would have more WB. I use the following phrase with my

students, “An act (or policy) is efficiency increasing if the gain to the gainers is greater than the loss to the losers.”¹

P.I.s are efficiency increasing and, by definition, increase societal WB. But not all efficiency-increasing actions are P.I.s. A synonym for efficiency increasing is *potential* Pareto improvement, *Potential* because there is the potential for the gainers to compensate the losers and turn it into a P.I. If the compensation occurs, then the act with the compensation is a P.I.

Efficiency increasing does not imply societal WB has increased. It might increase societal WB, and it might decrease it. Efficiency increasing is neither necessary nor sufficient for societal WB to increase. Unfortunately, this does not preclude overzealous economists from wrongly declaring, “More efficient is better.”

Most things are not Pareto Improvements: someone usually ends up worse off. Every government policy decreases someone’s WB. So the dictate that a P.I. increase societal WB, while correct, isn’t of much practical use for parsing right from wrong government policies. And, most people’s behaviors negatively affect somebody. This is also true for animal behavior. [There are woodpecker holes in my stucco, and dogs don’t just poop in their own yards.]. In the real world, few changes hurt no one, so implementing Pareto improvements will require the compensation of losers.

Two additional facts further muddle the P.I. criterion for judging actions. (1) One can typically make any action a P.I. by adding or subtracting individuals from the roles of society. I play this game with my students. If the action isn’t a P.I. and you want to make it one, simply

¹ Economists claim efficiency is a positive concept, but it only becomes an operational concept after one judges who is and who isn’t a member of society—a foundational normative judgment. Of course, we believe our society should not include French people or other foreigners, simply because they do not belong to our tribe. The Israelis exclude the Palestinians, and the Palestinians exclude the Israelis. Another critical question is whether future animals (humans and other creatures) should belong. For example, if future individuals do not count, why sweat global warming?

Besides, whether an action is efficiency increasing depends on how society’s wealth is distributed. Usually, there will be gainers and losers, so an action is more likely to be efficiency increasing the more the wealth of society is held by those who gain from that action, and the less the wealth is held by the losers. That is, how wealth is distributed affects whether an action is efficiency increasing: an action is more likely to be efficiency increasing if it helps rich people and hurts poor people than if it hurts rich people and helps poor people.

ban from society those individuals who are made worse.² On the other hand, if the action is a P.I., and you don't want it to be, find one individual who is made worse off by the action and enter him into society's roles—it is as simple as that.

In conclusion, whether an act or policy is efficiency increasing (a potential P.I. improvement) isn't a welfare-economic criterion for parsing right from wrong because efficiency increasing might decrease aggregate WB, and efficiency decreasing might increase it. Remember rich Donald, the guy who stole your dead mother's ring.

But, doesn't welfare economics imply that unregulated competitive markets are the right mechanism for allocating resources and distributing goods and services?ⁱ

NO, it does not—it demonstrates that **if certain assumptions** hold, properly **regulated** competitive markets are preferred, on efficiency grounds, to unregulated competitive markets.

The First Theorem of Welfare Economics proves that if a specific list of assumptions holds, the equilibrium in an unregulated competitive market system will be efficient. Said another way, the equilibrium will have the property that the only way to increase the WB of one member of society will require that the WB of other members decrease. A corollary is overall efficiency is necessary for maximum societal WB. This has led many, including many economists, to wrongly conclude that markets are always ethically-preferred for allocating resources and distributing goods and services. And unregulated, competitive markets are better than regulated ones.

The First Theorem was anticipated in Adam Smith's *Wealth of Nations* (his invisible hand) but not formally proven until the 1950s. But the theorem does not support the conclusion that unregulated competitive markets are the right way to organize economic activity. While beautiful, it only proves that equilibrium in an unregulated competitive market system will be efficient in an idyllic world. And efficient does not imply that societal WB is maximized

The proof of efficiency requires that everyone's behavior is consistent with Assumptions 1-9 of NBT. It requires full information and certainty, so incorrect beliefs and flawed choosing are not admitted. Unfortunately, the real world isn't this idyllic: human behavior often deviates

² For example, if you assume that Muslims or Lutherans are not members society, how their WB is affected by a policy plays no role in determining whether the policy is a P.I. or a potential P.I.

from Assumptions 1-9. Besides, since we live in a world of incomplete information and uncertainty, we all hold numerous incorrect beliefs and practice flawed choosing.

For example, the proof requires that the WB of all parties to all market trades are made better off by every trade—but often, they are not. I'm watching late-night TV and see your ad for the BowFlex trainer, and I want the six-pack abs like the dude and dudette in the ad. I call the number on the screen, provide you my credit-card number (\$29.99/month for 24 months), and the nuts, bolts, and rubber bands arrive two weeks later. You, the seller, are likely better off, but I'm not because I incorrectly believed the machine would produce muscles like the couples'. Or, maybe the X-rated movie I watched put me in an emotional state that temporarily made the device more attractive. Or, perhaps I had done my research and decided that a Mr. Universe weight set was better for me, but right before the Bow Flex ad, there was something in the movie that cued me to temporarily concentrate on the fact that one of Mr. Universe's free weights might fall on my foot.

But, even if everyone always maximized their realized WB, the proof also requires no external effects: it requires the assumption that no one's behavior directly influences the WB of others. There are no external **effects**.³ But much of what we do directly affects others. Rather than listing specific examples of how an individual's behavior can affect others (as I did earlier), try to think of something you do that you are sure does not affect anyone else, negatively or positively—I can't.⁴ The proof also requires that when a good is produced (e.g., a BowFlex machine, electric power, or organic broccoli), its production has no direct effects on the production of other goods or people's WB.⁵ Many days the hours I spent on this book made my wife better off, many days worse off, and your reading this paragraph could easily affect how you interact with others, which will affect their WB.

³ One can't even conclude a voluntary trade between two individuals is efficiency increasing.

⁴ Can you? Consider some possible candidates: an act done alone, and no one else ever finds out you did it (watching Bugs Bunny cartoons at 3 a.m.; trying my left shoe on my right foot). Or I daydream of dancing with the NYC ballet and keep my dreams to myself. Each act might later affect how I behave towards my wife, my students, or my dog. And this might affect their WB. E.g., my imagining dancing with the NYC ballet might make me more aware that we all have both strengths and weaknesses, and this might make me more likely to forgive students who might exhibit a typical weakness of eighteen-year-olds. A forgiven student is typically a happier student.

⁵ The smoke from my power plant makes it more difficult for the laundry next door to produce clean air-dried underwear. But the bees I employ to produce honey, pollinate your fruit trees, making is easier for you to produce your apples.

To summarize, using the First Theorem as proof that unregulated competitive markets pursue WB-efficiency requires there are neither flawed choosing nor external effects. And, even if the competitive market equilibrium outcome were efficient, there might be another outcome, either inefficient or efficient, that generates more societal WB. Fairness is an issue, and market systems did not evolve to achieve fairness. In market economies, those who start with a big pile of resources end up with lots of goods and services; those who begin with few consume much less.⁶

If one supposes no flawed choosing but admits external effects—a common position amongst welfare economics—government intervention in the market (taxes or regulations) can increase efficiency. For example, many economists advocate for a tax on carbon emissions because there are negative external effects when carbon-based fuels are burned. A properly-regulated competitive-market system is preferred to an unregulated one on efficiency grounds.

As [Lionel Robbins](#) (1898-1984) pointed out in 1938:

*It was not possible to say that the economic science showed that free trade was justifiable.*⁷

External effects caused by relative-position effects and external effects caused by behaviors many find abhorrent

Both cause many to shudder at the implications of welfare economics.

For external effects, economists typically use examples where the effect is physical (e.g., pollution, traffic congestion, and covid transmission). But external effects include everything directly affecting third parties, including those caused by people taking offense simply because they find certain behaviors objectionable. An example should suffice: Suppose I do not like

⁶ Mistakes and luck can cause those who start with little to become rich and those who start with a lot to become poor. You might counter my conjecture that the market does not redistribute wealth by arguing that in a market system, with enough hard work you will get ahead. Fair enough, but I would counter that a propensity to work hard is a resource that you either got, or didn't, in the parental/genetic draw.

I got lucky by being born with intelligence and then by being adopted by parents who were not financially rich, but who were willing to spend all that they had on their childrens' educations. And, my anxiety genes have caused me to work hard to keep the wolf at bay.

⁷ Putting this quote in context, Robbins was reacting to Mr. Harrod's conclusion that if we can't make interpersonal utility comparisons, economics has little to say. "But I confess that at first I found the implications very hard to swallow. For it meant, as Mr. Harrod has rightly insisted, that economics as a science could say nothing by way of prescription. It could say whether a certain course of action could lead to a desired end. It could judge the consistency of different policies. But, in itself, it passed no verdict of good or bad. It was not possible to say that economic science showed that free trade was justifiable..." See also footnote 22 in Chapter 2.

public displays of affection by straight, white people, so my WB declines whenever I witness such displays. I don't even need to see it; just knowing it is happening makes me miserable. Imagine many people feel this way, and most are rich, whereas straight whites tend to be poor. In this case, reducing public displays of affection (hand-holding, kissing, etc.) by them could increase efficiency. And a ban might even increase societal WB. It would depend on how different distributions of WB are judged. Efficiency might even require that straight, white people have their own neighborhoods and schools, so individuals like me are less likely to be offended.

I, and many others, become livid when ignorant people practice free speech. If so, efficiency might require less free speech. Many people get upset by many things many people do. These include having inappropriate sex, having fun, being fat, criticizing beliefs (religious, political, and other), and practicing other behaviors that the beholders find disgusting.

You might believe that negative external effects that result from an objectionable ordering on the part of the impacted should not be considered, but for welfare economists, orderings are sacrosanct. What does and does not increase societal WB depends on everyone's orderings, but welfare economists, unlike ordinary people, show no inclination to draw a line between appropriate and inappropriate orderings.⁸ An implication is that racial and gender discrimination, and other offensive acts, will increase societal WB if enough people are prejudiced: efficiency implies behaviors that many find morally repugnant.

Consider the pleasure of causing others to suffer: all else constant, paths that inflict suffering are ranked higher. More sadism could be efficiency increasing. For welfare economists, the WB the pedophile gets from pedophiling counts towards aggregate WB, the same as the WB you get from eating ice cream, and the same as the WB loss to their victims. An extreme example, but you get the point.

Chapters 4 and 5 considered relative-position effects, but Chapter 2 did not consider their implications for welfare economics. A relative effect occurs when your welfare depends not just on what you consume but also on how much your friends and neighbors are consuming.

⁸ Economists would argue on efficiency grounds that without regulations or laws there will be an inefficient amount of child molestation—too much. But their argument is the same efficiency argument I could use to keep straight whites out of my neighborhood.

Evidence presented in Chapter 4 indicates substantial relative-wealth effects. Chapter 5 discussed other relative-position effects (faster runner, more attractive spouse, more sex than your peers).

Relative-position effects are external effects, so efficiency and societal WB require that they be considered. E.g., efficiency dictates that such direct effects on others must be produced in efficient amounts, and to accomplish this, behaviors must be manipulated with taxes or regulations. For example, if my striving to work more to have more isn't regulated, I will strive an inefficient amount (too much): I won't consider that my striving to have more stuff will decrease your WB. So, efficiency requires my behavior to be manipulated, so I strive only the efficient amount—maybe with income-tax rates or restrictions on how many hours I can work. If you train and become faster on your bicycle, I am now relatively slower, so worse off, and efficiency requires that your training be limited. Few people think about the implications of relative-position effects on efficiency and societal WB: many behavior manipulations would be needed to internalize these external effects. Of course, one could decide to care about relative position is inappropriate and declare external effects caused by such caring should not be considered. But, then one isn't a welfare economist.

What if we add common quirks, incorrect beliefs, violations of NBT assumptions, and flawed choosing?

With these present, the welfare economist can't even conclude that you will do what is right for you: you will likely do things that will decrease your realized WB. In a world of common quirks, incorrect beliefs, and flawed choosing, a welfare economist can't logically argue that more freedom is morally preferred to less. If increasing societal WB is the goal, correctly applied paternalism is morally preferred. In that world, the welfare economist is ethically obligated to stop me, or at least most people, from trying heroin.

The practical value of welfare economics is limited because it refuses to compare my WB with yours or even accept that they are comparableⁱⁱ

It can't determine, in practice, whether most actions are right or wrong. [In its defense, many ethics share this deficiency.] Suppose it is impossible to compare your WB to mine (make interpersonal WB comparisons). If so, whether a policy or act increases societal WB is unanswerable. If WB isn't comparable across the members of society, aggregate WB is, most of the time, a meaningless concept. The same is true of societal WB.

The idea that we can't compare the WB increase to some with the WB loss to others goes back, at least, to Jevons, who, referring to interpersonal comparisons of WB, stated in 1871:

I see no means whereby such comparisons can be accomplished. Every mind is inscrutable to every other mind and no common denominator of feeling is possible. (Jevons, as quoted by Robbins 1938)

Quoting John Broome (1998):

This idea [the refusal to make interpersonal comparisons] gave rise to a style of welfare economics that still lingers. In the United States, indeed, it still predominates. Virtually everything that happens in an economy is good for some people and bad for others. The dogma says economics cannot compare the benefit to some with the loss to others. So, it cannot judge whether a change is good or bad on balance. Economics, then, is unable to judge the merits of virtually any economic event.

The default to an efficiency criterion: the big switcharooⁱⁱⁱ

Despite all this, many, myself included, have spent hundreds, thousands, of hours estimating dollar measures of WB changes. We do this to add them up and publish them in a journal. Or we do it because we were asked to (and paid) by a policymaker or lawyer, for example, one suing a PRP, *potentially responsible party* (e.g., British Petroleum for the oil spill in the Gulf). Such behavior suggests our estimates can be used to determine whether a policy will make society better off. They can be used to estimate whether an action will increase or decrease efficiency, but not whether it would increase aggregate or societal WB.

A way out of this dead-end is to make the goal maximizing efficiency rather than maximizing aggregate or societal WB (see Posner in Chapter 2). If asked directly, welfare economists would say the objective is to increase WB, but still, many economists run around saying "more efficient is better than less". Repeating a quote from Broome (1998):

They [welfare economists] have slipped from identifying efficient states, which can be done without interpersonal comparisons, to claiming that efficient states are better than others, which generally cannot. Robbins was right about this...

Welfare economics in practice concentrates on identifying efficiency because, unable to adopt a fairness metric and insisting that WB can't be compared across individuals, it had nothing else to do. From that, it has slipped into pursuing efficiency regardless of other goals.

WB is cardinal

Most welfare economists reject this; instead, they assume that WB has only an ordinal meaning. But, making interpersonal WB comparisons almost always requires that WB is cardinal for each member of society, and each member's WB can be compared. Recollect from Chapter 1 that if your WB is cardinal, you can determine whether a move from path j to path k will increase your

WB more, or less, than a move from path s to path t . If your WB is only ordinal, you can't do this.

Societal WB allows for your WB to count more or less than mine. So, in addition to assuming cardinal WB, one has to decide what is fair: decide how much my WB counts relative to yours. Unfortunately, welfare economists refuse to make fairness judgments; ordinary people make them all the time but disagree on what is fair.⁹

In my principles of economics course, I tell my students that there are no objective criteria to judge whether one distribution of WB is more or less fair than another—it is simply a matter of what you choose to believe. I have read, heard, and said this so many times I give it no thought. But this economic view that fairness is subjective is, in itself, a subjective assumption that many, including many moral philosophers, would reject.

For example, many people, but not everyone, find it unfair that the combined wealth of the world's eight wealthiest men equals the wealth of the world's bottom 50%. And they believe that a policy whose only effect on WB is to increase the WB of one of the eight would decrease societal WB even though it would increase aggregate WB and be a P.I.

For many, what is fair depends on why people do what they do, including how much behavior is chosen. For example, if you believe the poor could have been prosperous but chose, instead, leisure, while the rich are rich because they chose hard work, then you view as unethical policies that take from the rich and give to the poor—doing so would decrease societal WB.

Alternatively, if you believe whether you are rich or poor is mostly a matter of dumb luck (who bore you, where you were born, and what happened to you when you were a baby), you believe it is ethical to take from the rich and give to the poor—doing so would increase societal WB. A question that economics, psychology, and neuroscience are designed to answer is how much we control our actions and destiny. The findings, to date, suggest not as much control as we would like to think.

⁹ An economist might assert that economics is a science and specifying a criterion for fairness is unscientific. It is true that parts of economics follow the scientific method, but economic ethics isn't one of them.

Our ability, as welfare economists, to distinguish right from wrong comes down to the following. If we can agree on who is and who isn't a member of society, actions that increase the WB of some members and decrease the WB of no members are right, and those that reduce the WB of some members and increase the WB of no members are wrong. This isn't saying much.

Interpersonal WB comparisons are making a comeback amongst economic ethicists: Rawls, Sen, and Dasgupta^{iv}

The dearth of welfare-economic predictions has led a few economic ethicists to search for additional criteria to make economic ethics more useful. Their works are studied more by non-economists than economists and less by American economists than those in Europe. Rawls, Sen, and Dasgupta propose and defend specifics on what is fair. For example, Rawls believes ethical is what is contractually acceptable to society's members.

I limit my discussion to them: they are the predominant economic ethicists of the last fifty years. They are neither utilitarians nor welfare economists.¹⁰¹¹ Each is happy to make interpersonal comparisons, and none define societal WB as aggregate WB. And they care about what is fair.

Rawls

The ethic of John Rawls is based on both process and consequence; the great Harvard political economist envisioned a group of reasonable people choosing how society should be organized (the form of government, the laws, and how economic activity is organized). The group must decide how to organize society before each member knows their positions in society (wealth, intelligence, parents, abilities, etc.). This is choice behind a *veil of ignorance*. Rawls argues this is a good process because it causes good consequences—the process eliminates self-interest bias by the deciders, so Rawls is both a *contractualist* and a *process consequentialist*. The Rawlsian process will result in an ethical social contract, so adhering to it will be ethical and deviating from it unethical.

¹⁰ In contrast, [Yew-Kwang Ng](#) is a welfare economist who endorses interpersonal comparisons.

¹¹ Expanding beyond these three, consider [Samuel Bowles](#), [Jeffrey Sachs](#), and the late [Anthony Atkinson](#). Also, John Broome who might or might not consider himself an economist.

Rawls specified fair, arguing that the organization of society's institutions and the distribution of its wealth must be perceived as fair if it is to be accepted by its citizens.¹² He limits his analysis to societies with sufficient resources to meet everyone's basic needs.

Social cooperation will make the goods-and-services pie bigger. But to get that cooperation, the citizens must know in advance that the pie will be fairly sliced. For Rawls, it is unfair for me to get more stuff simply because I was born rich or with more than average talents. So, Rawls argues that income should be equally distributed unless distributing it unequally would be better for everyone, including the least well off. Rawls expressed this in two principles. The first takes precedence over the second.

The first principle affirms for all citizens familiar basic rights and liberties: liberty of conscience and freedom of association, freedom of speech and liberty of the person, the rights to vote, to hold public office, to be treated in accordance with the rule of law, and so on. The principle ascribes these rights and liberties to all citizens equally. Unequal rights would not benefit those who would get a lesser share of rights, so justice requires equal rights for all in all normal circumstances...

Rawls's second principle of justice has two parts. The first part, fair equality of opportunity, requires that citizens with the same talents and willingness to use them have the same educational and economic opportunities regardless of whether they were born rich or poor.

The second part of the second principle is the difference principle, which regulates the distribution of wealth and income (Wenar).¹³

The difference principle is that society can only fairly choose an unequal income distribution if everyone has more income under the unequal distribution. And, if it can choose between several unequal distributions (all preferred by everyone to an equal distribution), the fair choice is the distribution that provides the highest income for those at the bottom.

“The [second] principle does not allow the rich to get richer at the expense of the poor.” (Wenar). Ethical (unethical) economic policies are those consistent (inconsistent) with the two principles. Rawls clarifies that if a government action affects incomes, it is only ethical if it

¹² If a system is accepted as fair by all of its citizens, then behaviors that violate the rules of that system are unethical.

¹³ Quoting Rawls, “First Principle: Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all.” “Second Principle: Social and economic inequalities are to satisfy two conditions: a. They are to be attached to offices and positions open to all under conditions of *fair equality of opportunity*; b. They are to be to the greatest benefit of the least-advantaged members of society (the *difference principle*).” (Rawls 2001: 42-43)

increases the incomes of the worse off. And actions that increase their incomes are ethically preferred.

Rawls defined fairness with rights, opportunities, and income, not WB. This is added evidence that Rawls was not a utilitarian. In addition, income is easier to measure than WB, making his principle easier to apply than ethical principles based on WB.¹⁴

What do Rawls' two principles imply about living an ethical life? I speculate. Rawls said that choosing an ethical mechanism for organizing economic activity requires that the choosers choose behind a veil of ignorance. So Rawls must have felt that without the veil, individuals would choose the system personally best for them. He either believed that an individual is incapable of deviating from their highest-ranked available path, HRAP. Or thought deviating is possible but only with difficulty.

Suppose Rawls thought individuals are capable of deviating from their HRAP. Would he say each choice I make (e.g., going skiing, selling real estate, working in a food bank) is unethical unless it increases the WB of the worst off? While that would be a noble goal, most of us would not live up to it. Alternatively, would Rawls say that to the extent I am capable, I should live my life—not necessarily choice by choice—but, overall, in a way that will increase the WB of those worse off than me? The London School of Economics philosopher Alex Voorhoeve says yes, concluding,

... certain principles for individual conduct—including a principle requiring relatively advantaged individuals to sometimes make their economic choices with the aim of maximising the prospects of the least advantaged—are an integral part of a Rawlsian political conception of justice. (Voorhoeve 2005)

Rawls argued that if reasonable citizens choose behind his veil of ignorance, they will select a society based on his two principles, not utilitarian principles.¹⁵ [Rawls viewed utilitarianism as his competition. Rawls's principles are based on an agreed-upon social contract, not maximizing aggregate WB or aggregate want fulfillment, so they conflict with Benthamite and preference utilitarianism.]

¹⁴ It is much easier to identify those who are worst off in terms of income than in terms of WB.

¹⁵ His argument restrictively assumes that the citizens do not care about relative income and that citizens are risk neutral. Chapter 4 discussed data that shows people do care about their relative position.

Sen

Sen argues that the 19th and 20th-century purge of interpersonal comparisons from economics was misguided. And to be useful, economics must judge the gains of some against the losses of others. He argues that economics can and should propose criteria for interpersonal judgments. Like Partha Dasgupta (1942-), Sen is particularly interested in the WB of the poor in poor countries. Both offer a quality-of-life approach for identifying WB. Sen equates WB with what you do and achieve (are capable of and choose to do) and your unrealized capabilities (what you could do but choose not to do).¹⁶ More capabilities are ethically preferred—you have more freedom (more options).

Feeling happy is of value, but it's not all of WB. Other aspects are achieving or having the capacity to achieve a long life, enough to eat, knowing there will be enough to eat, and literacy—all these expand what one is capable of. These, for Sen, are components of WB (intrinsically good) even if they never increase your emotional or life-satisfaction WB. And emotional WB does not contribute to WB if it is *inauthentic*, such as living in dire circumstances but being happy because you were born with happy brain chemistry. Of course, having happy brain chemistry is better than having depressed brain chemistry, but happy neurons do not imply more WB as Sen defines WB. And if you are in dire straits, happiness that results because something worse did not happen when you expected it would, is not increased WB.

For Sen, many components of WB are objective and easily measured, or, at least, are easier to measure than emotional or life-satisfaction WB. Sen also asserts that fulfilling desires does not necessarily increase WB. He puts forth two reasons: we don't always like what we desire, and the oppressed have such limited desires that fulfilling them will not increase their WB. Sen admits flawed choosing, rejecting the notion that choosing Path *k* over Path *j* necessarily means path *k* gets you more WB. Sen's conclusions are in step with recent developments in neuroscience on wanting versus liking discussed in Chapter 4. Sen is neither a Benthamite nor a preference utilitarian.

Recently, Sen (2009), following in the step of John Rawls (Sen was a student of Rawls), proposed a philosophy of social justice (*The Idea of Justice*). While Rawls laid out his conception of a just society—*perfectly just* (what Sen refers to as *transcendental justice*), Sen

¹⁶ This view of well-being goes back to Adam Smith and Karl Marx (Sen 1988)

finds Rawls's tack impractical for actual judgments in the real world. Instead, he views justice as a continuum (*comparative justice*). The goal is not to make the world just, only less unjust.

Sen equates justice with fairness, and while indicating there are different notions of fairness, he wants to judge fairness from an impartial perspective, a global observer who takes account of all points of view. In practical terms, gross injustices such as slavery and unnecessary starvation are unjust. But the implications are less clear if one must choose between two states of injustice, neither grossly unjust. But that, in a way, is his point; often, there is not a clear answer. Sen is outlining a process for deciding, not the answer.

Dasgupta

For Dasgupta (*Human Well-Being and the Natural Environment*), a better society is a wealthier society: wealth is broadly defined to include not just manufactured capital but also the natural-resource stock, the environment, education, knowledge, freedoms, and society's institutions. An action is right if it adds wealth because greater wealth leads to greater societal WB.¹⁷ For Dasgupta, societal WB depends on both average WB and how WB is distributed, and the poor have a greater claim to additions to the pie.

Dasgupta equates personal WB with a "well-lived life" (to achieve an admirable character, a successful family, warm friendships, a meaningful job, and a life free of coercion). [As Aristotle pointed out, you can't have a well-lived life without wealth.] Components of WB include human rights and that humans should not be treated differently based on their "names, castes, or political or religious or ethnic affiliations." Dasgupta is not a pure rights theorist. He feels that these rights need to be justified in terms of human flourishing, so they are not absolute rights.

A problem is how to measure and weigh these different components of wealth. Doing this starts with identifying the properties of an acceptable index, the topic of his book. For Dasgupta, societal WB is an aggregate of the WB of individuals, and an individual's WB depends on both his material WB and his "ability to exercise various kinds of freedoms". Democracy, civil

¹⁷ Dasgupta was an early arguer that income measures of wealth ignore our dependence on natural resources. He is an eminent natural-resource economist. He notes that he ignores animal welfare.

liberties, and sustainability enhance welfare. Intergenerational WB is important but threatened because we are using our natural resources too quickly.

Rawls' principles are rights, opportunities, and income; Sen's and Dasgupta's principle is WB, defining WB more broadly than emotional and life-satisfaction WB, including rights and capabilities. A difference is Rawls uses income directly rather than indirectly through its influence on WB. All three argue that bettering the poor should have more weight than bettering everyone else. Sen and Dasgupta are more specific about what constitutes WB. Sen and Dasgupta consider what is ethical in our world of poor societies with immense suffering—agreeing that more command over goods and services, either individually or collectively, isn't always the best way to increase WB. Rawls is more abstract and limits his two principles to societies with enough resources to provide everyone with food and shelter.

Nevertheless, their ethics have a lot in common: (1) emphasizing making the worst off better off. As a Buddhist would say, reduce suffering. (2) Each requires that we all have the same rights and freedoms. Each requires that our educational opportunities and ability to live a long and healthy life should not depend on our wealth, gender, ethnicity, or race. And (3) none directly addresses how an individual could live a more ethical life or even whether an individual is capable of choosing to live more ethically.

ⁱ Sources:

1. Lionel Robbins (1938) Interpersonal Comparisons of Utility: A Comment, *The Econ. J.* 48(192): 635-641.

ⁱⁱ Sources:

1. John Broome (1988) What's wrong with poverty? *London Rev. of Books* 10(10):16-17.
2. *The Concise Encyclopedia of Economics* (2008) Vilfredo Pareto. Library of Economics and Liberty. At <http://www.econlib.org/library/Enc/bios/Pareto.html>.
3. Robbins (1938) *ibid*

ⁱⁱⁱ Sources:

1. John Broome (1998) *ibid*.
2. Gerry Mullany (Jan. 16, 2017) World's 8 richest have as much wealth as bottom half, Oxfam says, *N.Y. Times*. At <https://www.nytimes.com/2017/01/16/world/eight-richest-wealth-oxfam.html>

^{iv} Sources:

4. Partha Dasgupta (2001) *Human Well-Being and the Natural Environment*, Oxford, Oxford University Press. Chapter summaries at <http://www.oxfordscholarship.com/view/10.1093/0199247889.001.0001/acprof-9780199247882-chapter-2>
5. Kew-Kwang Ng (1997) A case for happiness, cardinalism, and interpersonal utility comparability, *The Econ. J.* 107: 1848-58.
6. John Rawls (1999) *A Theory of Justice, revised edition*, Cambridge, MA: Harvard University Press.

-
7. John Rawls (2001) *Justice as Fairness: A Restatement*, Erin Kelly (ed.), Cambridge, MA: Harvard University Press. Quoting from the publisher, “This book originated as lectures for a course on political philosophy that Rawls taught regularly at Harvard in the 1980s. In time the lectures became a restatement of his theory of justice as fairness, ... As Rawls writes in the preface, the restatement presents ‘in one place an account of justice as fairness as I now see it, drawing on all [my previous] works.’”
 8. Amartya Sen (1987) *On Ethics and Economics*, Oxford, Blackwell Publishing.
 9. Amartya Sen (1988) *The Standard of Living (Tanner Lectures Human Values at Cambridge)*, edited by Geoffrey Hawthorne, Cambridge, Cambridge University Press. Earlier version at [http://www.akira.ruc.dk/~fkt/filosofi/Artikler_m.m/Sen – The Standard of Living.pdf](http://www.akira.ruc.dk/~fkt/filosofi/Artikler_m.m/Sen%20-%20The%20Standard%20of%20Living.pdf)
 10. Amartya Sen (2009) *The Idea of Justice*, Cambridge, Harvard University Press.
 11. Alex Voorhoeve (2005) Incentives and principles for individuals in Rawls's theory of justice. *Éthique et Économique* 3 (1): 1-7.
 12. Leif, Wenar (2014) SEP on *John Rawls*. At <http://plato.stanford.edu/entries/rawls/>
 13. Jonathan Wolff (2010) On John Rawls' *A Theory of Justice*, Philosophy Bites. <http://philosophybites.com/2010/02/jonathan-wolff-on-john-rawls-a-theory-of-justice.html>

Epilogue: Do you make choices, and if so, has your choosing made you happy?ⁱ

08172022

You have gotten to the end, suggesting you either have sympathy for the research and literature discussed, or you can't not finish a book. Either way—bravo and thanks.

To summarize, this book is an exercise in wondering why you and I do what we do. I distinguish between behavior and choice (chosen behaviors) because the economic default is to assume that much of what we do, we first chose to do. Economists call their behavior model “choice theory”, not “behavior theory”. And everyone desperately wants to believe that if they have an experience of choosing, it was that conscious, cognitive experience that determined why they did what they then did—we did it because we consciously chose to do it.

Effort was direct at defining what is meant by a “choice”. If choice requires the ability to select differently than one selected—the lay definition—economici don't make choices. For economists, there is choice unless external constraints result in only one available path. Humans, mollusks, and plants all make choices in this weak sense of choice.

I proposed a definition of choice. The whole philosophical debate about free will comes down to two things: how to define a choice and, given that definition, whether humans and other animals are capable of choice.

My objective has been to get you interested in the psychology, neuroscience, and philosophy of behavior and choice and their implications. I did this against the backdrop of how economists think about behavior, provoking you to think about the premises of neoclassical choice theory, NCT, and where they came from. I asked you to question the theory's ability to predict your behavior and the behaviors of others (humans and other animals).¹ And to cause you to contemplate the practical relevance of welfare economics, the predominant ethic of economics.

¹My conjecture is that economicus, the imagined individual in neoclassical choice theory, was born out of the religious belief that man was created in the image of God, so rational and non-erring: its creators were Christians living in late 19th Century and early 20th Century England. Economicus is not the product of evolution.

One aspect of NCT has excellent predictive power—you can't do what you can't afford; you cannot have what is not available. My Principles of Microeconomics course discusses the budget constraint—you can't consume a path of goods that costs more than you have to spend. This simple assumption generates the big prediction of NCT: when a price goes up, demand goes down—except when it doesn't. If a price increases or your income falls, you can no longer afford your consumption path, so you need to cut back. Unfortunately, economists don't explicitly consider other equally important constraints imposed by age, health, culture, skills, and obligations to friends and family. I suspect economists do this because these constraints are fuzzier and difficult to mathematically model. So, rather than tackling the difficult job of modeling them, they get lumped into a hodgepodge called *preferences*. NCT then assumes that this ordering of paths is given and fixed, making it your ultimate constraint. But when I assert this, colleagues look at me funny. Modern preferences are a hodgepodge of likes, desires, and constraints.

When I behave differently from you, and the difference can't be explained by us having different constraints, the difference is simply attributed to different preferences. But as George Stigler and Gary Becker noted in 1977, “No significant behavior has been illuminated by assumptions of differences in tastes.”—something that can explain everything explains nothing. The full range of constraints needs to be explicitly modeled, and to the extent tastes do differ, it is essential to ask why and try to model why.

As I was taught, NCT assumes that you will choose your most-preferred path subject to your constraints or, said another way, maximize your utility subject to your constraints. But, as noted, what “more preferred” means is hard to say; it isn't much discussed. One conjecture is that if path t is preferred to path d , you associate more WB with t (more emotional WB or life-satisfaction WB). I discussed emotional WB (happiness) in depth, how our conception of it has changed over the ages, and what makes you happy and unhappy (and for how long). A significant raise will make you happier, but not forever, or as long as you predict it will—unless it permanently raises your income relative to your peers.

I discussed the diverse literature on the relationship between income, consumption, emotional WB, and life-satisfaction WB—it's complicated. But, generally speaking, more

money does not necessarily increase emotional WB in the long run unless you started critically poor.

It's tough to conceptually separate NCT and economic ethics. NCT assumes you know what is best for you and, subject to your constraints, do it. This, combined with the welfare-economic assumption that right actions are actions that increase the WB of individuals (betters them), implies that you should be free to pursue your interests, as long as doing so does not hurt others. In contrast, welfare consequentialists who believe that humans often get it wrong advocate strong or weak paternalism. [The sexy term for weak paternalism is *nudging*.²] While most people believe the effect on WB is a factor in determining right from wrong, few believe it should be the only factor.

Most moral philosophers are not even welfare consequentialists—a surprise to economists. Many modern moral philosophers adhere to a duty-based ethic or a process ethic. The book included a primer on ethics. It is common for economists to conflate WC and liberalism (more freedom is better). But more freedom being better than less only follows from WC if we all behave in our own best interests and doing so does not hurt others. Liberalism, in contrast, argues that freedom is good even if it allows you to make yourself worse off.

A century ago, NCT abandoned the need for cardinal utility, and welfare economics forsook comparing utility across individuals. So welfare economics now provides little guidance for those trying to parse right from wrong behaviors and policies. In desperation, welfare economics default to “More efficient is better than less efficient”, but as I tell my students, “Efficiency is like good sex; more is better except when it's not.” Economic philosophers such as Amartya Sen and Partha Dasgupta argue that if economic ethics is to be more prescriptive it will need to embrace more criteria for judging than each individual's perception of their own ordinal WB.

Alternatively, economists could exit the business of claiming to be able to judge behaviors and policies, admitting that modern welfare economics has little ability to parse right from wrong.

² “Nudge” as in you are nudged towards the alternative in your best interest. Thaler and Cass Sunstein have a well-known book on the topic, *Nudge*. Thaler won the 2017 Nobel Prize in Economics.

Research in psychology and behavioral economics (if you believe the research) indicates your choosing is often flawed. The good news is your flawed choosing is often predictable and avoidable. For example, when something significant happens, there is a pleasure or misery deviation from your norm, but it wanes over time. Unfortunately, we don't anticipate the waning. Instead, we incorrectly believe that our current feeling will maintain, be it happy or miserable. For example, the new freshman (my kid, years ago) drops out of college after a week because she is anxious and incorrectly believes the anxiety will never diminish. Or you buy a new car because you imagine the pleasure of having a new car will last, forgetting that it will soon be your old used car. Systematic mis-assessment of an action's benefits and costs causes flawed choosing. Being aware that you suffer from duration bias makes you less likely to err but does not guarantee it. You might want your friends, family, or government, to nudge you away from flawed choosing. Duration bias should be taught in elementary school.

Most people, including most economists, believe that which alternative you experience is determined by your conscious experience of choosing that alternative. But this belief is not required to explain behavior. NBT is consistent with everything determined by your unconscious. And that the choosing experience is only window dressing. One could add the assumption that what you experience is always driven by conscious choosing (Assumption 12), but the evidence does not support this.

Substantial research in neuroscience and psychology indicates the choosing experience is generated by the unconscious to make you feel responsible for and accept what you then do (buy the new car, divorce the old guy, and go skiing). If the hypothesis is correct, the choosing experience does not determine what you next do, but this does not mean conscious thought does not influence behavior. A choosing experience, while it does not affect what you already unconsciously selected, can influence how your unconscious will select next time. More generally, conscious thoughts affect what the unconscious will later decide. The experience of conscious choosing is a critical component of our sense of self. And it makes you believe that other people are responsible for what they do and should be treated accordingly—execute the killer, reward the humanitarian.

Even though I intellectually believe my choosing experiences have little influence on what I do next, I stumble through life imagining/fantasizing that they do. For example, I

