# Conceptualizing and Studying Free Will Belief and Disbelief

**Kyle A. Hietala**

Advised by Thomas H. Brown, PhD

Submitted to the faculty of Cognitive Science in partial fulfillment of the requirements for the degree of Bachelor of Arts

Yale University

April 20, 2018

**Abstract**

This paper takes no position on whether free will exists, but focuses instead on how free will has been imported into the cognitive sciences from philosophy and how free will belief (FWB) is conceptualized and studied. This paper begins by discussing various conceptions of free will emerging from various disciplines. It proceeds to critique 'willusionism,' the scientific argument against the existence of free will, for overemphasizing antiquated philosophical conceptions of free will and inadequately addressing more recent philosophical and psychological refinements. The willusionist case can be strengthened by critiquing the type of free will that people believe they have, based on recent findings from experimental philosophy. Additionally, the willusionist case can be strengthened by focusing less on neuroscientific, reductionist explanations for behavior and more on external, environmental sources of influence and control. Second, this paper criticizes research on free will disbelief, especially within social psychology, for committing methodological errors and misinterpreting findings, leading to overblown and unwarranted concerns about free will skepticism and disbelief. Previous research has failed to disentangle deterministic from fatalistic beliefs and explore compatibilist accounts. Third, this paper discusses various neurocognitive explanations that can account for the observed consequences of free will disbelief. Though FWB is about feelings of self-efficacy and the ability to make choices, free will should be understood as being primarily about personally meaningful choosing.

**Contents**

# Introduction

Though philosophers have argued about free will for millennia, cognitive scientists, neuroscientists, and psychologists have recently joined the project, using empirical methods to explore major issues related to free will belief (FWB). This paper does not take a stance on whether free will exists. Rather, this paper begins with an assessment of the current debate about free will and argues that disagreements often arise from differing conceptions of free will, not from different ideas about the same conception(s). The paper proceeds to review recent research about the implications of FWB on cognitive functioning and psychological well-being, and finally presents critiques and theoretical accounts for empirical results.

I begin by discussing various conceptions of free will to illuminate the ways in which disciplines have talked past each other when studying and debating free will. I proceed to review the evidence produced by 'willusionists,' who argue that free will is an illusion. I group the willusionists into two categories: the first of which I call the exclusionists, who maintain that brains fully determine behavior without one's conscious authorship (thereby *excluding* free will), and the second of which I call the externalists, who maintain that our environments, especially our social contexts, powerfully influence or determine behavior.[1] Based on recent empirical findings from experimental philosophy, I argue that the externalists offer the more persuasive and useful case.

[1] Importantly, these approaches to disputing free will are not mutually exclusive nor exhaustive; the distinction I draw merely refers to how each group approaches disputing free will and the sorts of evidence each group finds most persuasive.

Additionally, this paper raises objections to the arguments made and methods used by many psychologists, who have argued that free will exists and we would be worse off without it. In addition to making methodological missteps, psychologists have often conflated determinism and fatalism in studying free will, leading to questionable empirical results. I argue, therefore, that we have less reason to worry about free will disbelief than the current literature would suggest.

Lastly, this paper considers theoretical accounts of FWB. First, I argue that neuroscientific evidence suggests that free will (dis)belief influences pre-conscious, pre-deliberative cognitive processes rather than conscious, deliberative processes. Second, I argue that FWB should be conceptualized as context-sensitive and variable and that the significance and personal meaningfulness of choices should inform how we study FWB.

## Free Will in Philosophy

As a part of western philosophical thought, free will traces its origin to the theologian Saint Augustine, who used it to explain why humans sinned and to answer the problem of evil. Departing from the ancient Greeks, who argued for a naturalistic and rational understanding of the human mind, Augustine elevated free will to a cosmic principle, separating the mind from natural law (Ravven, 2010). Because humans could will evil into the world, they could be held accountable for their sins (Ravven, 2010). Thomas Aquinas, in *Summa Theologica,*

refined the Augustinian conception, defining free will as "the primary mover of all the powers of the soul…and the efficient cause of motion in the body" (Stump, 1997).

From this historical tradition emerged two primary schools of thought: compatibilists, who argue that free will is still possible if determinism is true; and incompatibilists, who argue that the coexistence of free will and determinism is impossible. Determinism holds that all events are determined by prior causes. Incompatibilists are either libertarians or hard determinists; the former believe free will exists and determinism is false; the latter believe free will does not exist and determinism is true. Many scholars have proposed various modifications to this simplified taxonomy, but these distinctions remain the most common organizational features of philosophical free will.

Libertarian conceptions of free will generally fall into three accounts: noncausal, agent-causal, or event-causal. Noncausal or ownership accounts of free will hold that a person controls their will simply because their will is their own. (Ginet, 2002; Pink, 2004). Agent causal accounts describe free will as the ability of an agent to will his or her desires, thoughts, and intentions into being (Chisholm, 1966; O'Connor, 2010). Agent-causal accounts often embrace Cartesian dualism, holding that non-mental characteristics of the mind allow an agent to will his or her desires into being (Greene & Cohen, 2004). Event-causal theories of free will hold that prior events indeterministically give rise to present circumstances (Ekstrom, 2000). Free will is part of a process of deliberation; people routinely grapple with resolving competing desires and the outcomes of conflicted desires are

indeterminate. Since deliberating over competing desires occurs in multiple stages, people can choose to follow one desire at any given stage (Kane, 1996).

Compatibilism, expressed in the maxim of Arthur Schopenhauer (1839), "man can *do* what he wills but he cannot *will* what he wills," accepts determinism yet implies that self-control, rational thinking, and volitional action can sufficiently constitute free will. Most (59%) professional philosophers are compatibilists (Bourget & Chalmers, 2013), who accept that the universe is deterministic, yet hold that free will consists in a person choosing among their desires. Consequently, free will entails limited but not absolute freedom to make choices, and people are responsible for the choices they freely make (Dennett, 1984; Dennett, 2003). Hence, compatibilism aims to reconcile both free will with determinism *and* moral responsibility with determinism.

## Free Will in Psychology

Modern psychological conceptions explain free will as a type of a psychological state or capacity. Harry Frankfurt (1971) redefined free will as a second-order mental state. Though people may feel impulses which they did not willfully intend nor initiate (e.g. to eat when hungry, shout when angered, sleep when tired, etc.), people can choose whether to act on these first-order desires. Determinism can account for first-order desires (ones which simply come to us, or à la Schopenhauer, "cannot will what [we] will") but free will accounts for second-order desires (ones we initiate and carry out, à la Schopenhauer "can *do* what [we] will"). Free will, per this conception, could be understood more precisely as a power

of veto rather than as a power of initiation, since we veto some desires and act on others.

Most recently, psychologists have explained free will as a type of psychological accomplishment, arguing free will is constituted by psychological capacities related to decision making and self-control. When properly functioning, free will is a psychological accomplishment because it allows us to make decisions that lead to desirable life outcomes (Nahmias, 2017).

## Free Will in Neuroscience

Though neuroscientists who study free will tend to be willusionists, Peter Tse is a notable exception. Tse (2013) argues that free will can be explained at the synaptic level. Tse contends that free will is not found in present decision-making, but in the human ability to imagine and pursue a future self, which thereby changes the brain via rewiring synaptic connections. Free will, therefore, can be found in rapid synaptic reweighting, a process in which neurons rewire, based on a deliberately chosen set of criteria. Reweighting changes the possible physical bases for future mental events. Like strong libertarian notions of free will, Tse's conceptions requires four features:

> "We must have (a) multiple courses of physical or mental behavior open to us; (b) we must really be able to choose among them; (c) we must be or must have been able to have chosen otherwise once we have chosen a course of behavior; and (d) the choice must not be dictated by randomness alone, but by us." (pp. 133-134)

Andrea Lavazza (2016) attempts to link the psychological capacity model of free will to a neuroscientific explanation, arguing for a practical need to connect

higher order descriptions of free will to neural underpinnings. Lavazza describes free will as a type of capacity index which should represent cognitive abilities strongly linked to the concept of control. Executive functions, in this account, provide the underlying neural correlates for free will; the degree to which a person has free will can therefore be understood as the capacity of their executive functioning (psychopaths, for example, would have limited free will in this account).

## Free Will in Folk Understanding

Scholarly definitions aside, ordinary people generally understand free will as the ability to choose a desired course of action without restraint (Monroe, Dillon & Malle, 2014; Feldman, Wong & Baumeister, 2014; Feldman, 2017). Even if some scholars conceptualize free will in abstract, metaphysical terms (Greene & Cohen, 2004; Montague, 2008; Bargh, 2008), people tend to link free will most closely with the psychological concept of choice, not metaphysical concepts (Vonash, Baumeister, & Mele, 2018). People who make more choices tend to believe more strongly in free will, and the more strongly people believe in free will, the more they feel satisfied making choices and feel efficacious in making their decisions (Feldman, Baumeister, and Wong, 2014). Some scholars who study free will from all major approaches (philosophical traditions, psychological theories, and folk understandings) have suggested that contemporary scientific understandings of free will simply talk past ordinary notions of free will. Joshua Knobe (2014), for example, argues, "so long as we are working within a broadly scientific framework, we will never be talking about the very thing that people are getting at with their

ordinary notion of free will." The study of ordinary people's FWBs, then, may suggest the impossibility of a true scientific account of free will or perhaps provide a bridge between ordinary belief and scientific inquiry, such that insights from each could inform the other.

## Experimental Philosophy and Free Will Belief

Experimental philosophy, a nascent academic discipline that uses empirical tools such as survey questions and thought experiments to study people's philosophical beliefs, has recently explored the FWBs of ordinary people (non-specialists about free will). Different psychological processes produce different intuitions about free will, leading people to hold both incompatibilist and compatibilist beliefs, depending on the circumstances (Nichols & Knobe, 2007). Further, when people were presented with evidence that they had endorsed both compatibilist and incompatibilist intuitions over the course of an experiment (an apparent contradiction) and were asked to reconcile their inconsistency, they held to their compatibilist and incompatibilist intuitions equally (Nichols & Knobe, 2007).

This apparent contradiction may be partially explained by follow-up research which showed that peoples' judgments about whether a behavior relies on brain and body processes do not predict whether people believe those behaviors are compatible with causal determinism. (Bear & Knobe, 2005). Instead, people tend to classify behaviors as either active (effortful and spontaneous) or passive (going with the flow, being controlled by environmental or mental forces). People believe only passive behaviors are compatible with determinism. Hence, people are intuitive

compatibilists only when considering passive behaviors (Bear & Knobe, 2015). More recent research contends that people intuitively believe that the universe allows for incompatibilist free will. More strikingly, people—regardless of their personal beliefs about free will—associate a universe with incompatabilist free will with more well-being, prosocial behavior, and moral responsibility (Feldman & Chandrashekar, 2017). Across cultures, most people believe that the universe is indeterministic and that moral responsibility is not compatible with determinism when considering abstract cases of wrongdoing (Sarkissian et al., 2010). When presented with concrete cases of wrongdoing, though, most people believe moral responsibility is compatible with determinism (Nichols, 2011). Importantly, people tend to conflate not having free will with determinism and fatalism, believing they cannot change their fate (Nahmias et. al, 2005).

## Free Will's Critics – The 'Willusionist' Position

Regardless of what people believe about free will, many scholars within the brain sciences have put forth forceful critiques against the existence of free will (e.g., Bargh, 2008; Harris, 2012; Wegner, 2002). These scholars, dubbed willusionists, have attacked free will using two primary approaches. First are the exclusionists, who have critiqued free will primarily by using neuroscientific methods to show that conscious authorship is illusory and that the unconscious mind plays an outsize role in determining choices. Exclusionists present arguments at the low molecular level of organization. Second are the externalists, who have critiqued free will by primarily focusing on the relationship between the mind and

the external environment, arguing that a complex and imperceptible web of external influences determines choices. Externalists present arguments on the higher organizational levels of psychological phenomena, behavior, and complex systems (involving both the brain and its environment). Though not mutually exclusive, these approaches differ in where they find the strongest evidence against free will.

## The Exclusionists

Benjamin Libet's seminal reaction time experiments (1983, 1985) showed that people's subjective feelings of initiating volitional action came after unconscious readiness potential was detected by EEG recordings. Since electrical activity in the brain built up prior to subjects reporting they felt the conscious will to act (in the experiments, pressing a button), Libet concluded that free will could not be involved in initiating action, but only in preventing it. Libet termed this ability as 'free won't,' the ability of consciousness to inhibit (or veto) processes after having originated in the unconscious mind.

Daniel Wegner (2002) challenged Libet's 'free won't' concept when he demonstrated that asking people *not* to think about something made them preoccupied with thinking about that very thing (in Wegner's experiment, he asked people not to think of a white bear, and instead of not thinking about a white bear, people reported that their thoughts were in fact *filled* with white bears.) Additionally, Wegner's evidence extended beyond the control of thoughts to ideomotor effects; when people were asked to think about the elderly, for example,

they (unbeknownst to them) walked and moved more slowly. Free will, per Wegner, amounted to a mere illusion.

In addition to Wegner's critique, modern neuroscientific research about volitional action both corroborates and challenges Libet's initial findings. Soon et. al. (2008) showed that brain activity in the prefrontal and parietal cortices can be detected up to 10s before entering conscious awareness. One common critique of Libet's original work is that the time lag between EEG measurements of brain activity preceding awareness and reports of subjective awareness was so small that it could be explained away by simple measurement inaccuracies and participant biases in reporting events later than they happened (Joordens et al. 2002). Consequently, Soon et al. demonstrated a time lag between unconscious initiation of cognitive processes and conscious awareness longer than could be explained away by measurement inaccuracies or participant biases, lending more support to the argument that conscious free will does not initiate action. Guggisberg and Mottaz (2013) challenged Libet's interpretation, arguing that early neural markers of decision-making are not reported because they are not finalized decisions, not because they are unconscious. That is, the crucial flaw of Libet's experimental paradigm is in asking subjects when they *decided*, not when they *began to decide*. On this view, decision-making should not be considered as an abrupt event, but rather as a multi-stage process. Further meta-analysis of 48 Libet-type experiments by Saigle et al. (2018) found substantial variation in results, with most replication attempts failing.

Sam Harris, a pre-eminent public intellectual and trained neuroscientist, popularized the exclusionist case against free will in his *New York Times*-bestselling 2012 manifesto*, Free Will*. Harris argues, "our wills are simply not of our own making. Thoughts and intentions emerge from background causes of which we are unaware and over which we exert no conscious control." Harris coined the clever if not confusing maxim, "the illusion of free will is itself an illusion," to suggest that if we engage in serious enough introspection, we easily realize that our thoughts, feelings, desires, etc. spring forth spontaneously into our conscious awareness—as if we are blindly reaching into a hat, pulling things out at random.

## The Externalists

Externalists, however, tend to focus more on sources of influence and control outside the mind to build their case against free will. Additionally, the externalists tend to focus on higher organizational levels, such as psychological experiences, behavior, and complex systems. Heidi Ravven (2012) marshals a wide array of psychological and sociological evidence to suggest that free will is a myth. Ravven highlights the profound effect of social influence on choice, citing examples such as Philip Zimbardo's infamous Stanford Prison Experiment, in which ordinary college students quickly exhibited shocking levels of cruelty toward mock prisoners when put into a prison role-play scenario, to Stanley Milgram's electroshock experiment, in which participants complied with orders to administer supposedly lethal levels of shocks to unknown subjects. Ravven also discusses free will as a sociological matter, especially in relation to appalling war crimes, including the Holocaust and the

abuse of prisoners at Abu Ghraib. Ravven argues that social phenomena such as conformity and hierarchy powerfully challenge the assumption that we have free will, noting that otherwise normal and healthy people behave with abhorrent cruelty in response to a toxic context. The crucial problem of exclusionist critiques, Ravven contends, is that they study the brain independent of natural social contexts and present subjects with trivial tasks (pushing buttons, reporting when a clock hand reaches a number, identifying shapes on a screen, etc.) that have little resemblance to the types of choices people normally make.

Another line of externalist critique focuses on how we misperceive information from the outside world and how we falsely assume that our decisions are rationally made via deliberation. Work on cognitive biases, begun by Nobel Laureates Daniel Kahneman and Amos Tversky, has shown the extensive ways in which our minds lead us astray and trick us into believing we make most decisions rationally and thoughtfully. For example, we make significantly different decisions in response to the exact same informational input when it is merely presented in different ways, a cognitive bias known as framing effect (Tversky & Kahneman, 1981). Further work by Michael Gazzaniga (2011) on split-brain patients has shown that our minds provide coherent yet false confabulations and interpretations to make narratives out of unrelated and irrelevant information. Gazzaniga describes the mind as an unrelenting pattern-seeker and narrative-creator, able to piece together meaning out of otherwise unrelated objects and events. Free will is an illusion created, in part, to imbue meaning.

## But which Free Will is an Illusion?

Other cognitive scientists remain skeptical that the preceding neuroscientific, psychological, and sociological evidence amount to a death blow to free will. Eddy Nahmias (2009) contends that the willusionists (Nahmias is focused primarily on exclusionists) argue primarily against the libertarian conception of free will, dismissing compatibilism. Some of free will's most strident critics only contest an immaterial, libertarian, ghost-in-the-machine version of free will (Ryle, 2009). Read Montague says, "the idea of free will is not even in principle within reach of scientific description" (2008) and Joshua Greene and Jonathan Cohen add that "free will…requires…commitment to some magical mental causation" (2004).

Indeed, Sam Harris dismisses compatibilism in his tract against free will, positing, "the free will that compatibilists defend is not the free will that most people feel they have" (2012). Daniel Dennett provides an extensive counterargument, expounded in *Elbow Room* (1984) and *Freedom Evolves* (2003). Dennett argues that conscious authorship is irrelevant to notions of free will and that it doesn't matter (both as a philosophical and practical matter) if thoughts arise from the unconscious. Dennett might be correct; people do not believe that the possibility of perfect neuroscientific prediction conflicts with free will. That is, even if 'my brain made me do it,' people don't tend to therefore believe, 'I don't have free will.' (Nahmias, Shepard & Reuter, 2014). Free will, per Dennett, is an evolved human ability that comes in degrees (in some cases, it is highly constrained by environmental factors). Free will is the ability of humans to control their actions in

response to environmental stimuli to pursue rational and desirable courses of action. This ability is like other psychological features.

Consequently, Dennett and fellow compatibilists are not just disagreeing with willusionists over the existence of free will; they are further disagreeing over what free will means. Whereas many willusionists assume indeterminism is correct, many compatibilists in fact argue that free will *requires* determinism to be true. John Baer summarizes the modern compatibilist stance on free will as follows:

> "Each of us has many courses of action that are possible in the sense that they are within our power—we could do them if we choose to do so—but we act only in ways that accord with our natures....if who a person is (her personality, cognitive abilities, beliefs, ideas, emotions, memories, wishes, thinking styles, etc.) is to have power over what she does—and isn't this what we really mean by free will?—then the only kind of free will that is coherent is deterministic free will." (Baer, 2008)

Until the disagreeing parties settle on what they're disagreeing about, though, we should be skeptical of any neuroscientific or psychological claim either against *or* in favor of free will. At minimum, we should consider what conception of free will is being defended or attacked. The interdisciplinary nature of cognitive science offers insights into how to fix this definitional morass. Using the tools of each discipline to inform and align research on free will should refocus debates around common understandings and produce more definitive results. By engaging more directly with philosophical nuances, calling into question accounts of free will that focus on physicalist notions of self-control, choice, etc., the willusionists can strengthen their case against free will.

Across three experiments in which people were presented with evidence that decisions were made entirely by the brain, 85% of people still maintained belief in free will (Nahmias, Shepherd & Reuter, 2014). Given that belief in free will is not threatened by neuroscientific prediction and disproving conscious authorship (two major arguments employed by exclusionists) (Nahmias, Shepard & Reuter, 2014), the externalist case against free will looks to be a more persuasive approach for the willusionists. This is corroborated by research showing that people judge human actions to be caused by complex, dynamic systems involving psychological, neurological, and social events (Turri, in press). Additional research should examine the extent to which externalist critiques of free will effectively undermine FWB, especially in comparison to exclusionist critiques. This is *not* to say that the externalist critique necessarily provides better objective evidence against free will than the exclusionist critique, but that it is more practical because it is better received. This is also not to say that exclusionist critiques of free will have no value, but that externalist critiques of free will have comparatively more value in persuading ordinary people to give up on free will, an ultimate goal held by many willusionists.

I have wagered in the previous section that analyzing the problems of the free will debate is, for the time being, a more useful project than engaging in the debate itself. I take no ultimate stance in this paper on whether free will exists or doesn't exist. I turn instead to consequences of and theoretical accounts for FWB.

## Consequences of Free Will Disbelief

Not believing in free will has been shown to have a wide range of implications on cognitive functioning, psychological well-being, and social behavior. Baumeister & Brewer list numerous negatives in their 2012 review:

> "Disbelief in free will has been shown to cause dishonest, selfish, aggressive, and conforming behavior, and to reduce helpfulness, learning from one's misdeeds, thinking for oneself, recycling, expectations for occupational success, and actual quality of performance on the job…Belief in free will has correlated positively with life satisfaction and finding life meaningful, with self-efficacy and self-control, with low levels of stress."

The severity of the harms of free will disbelief has prompted some scholars to argue for illusionism, the position that although free will is an illusion, we must still believe in it (Smilansky, 2001). Others have asked whether the harms of free will disbelief suggest a neuroethical responsibility to contain the claims of willusionists such that they do not become widely publicized (Nahmias, 2009).[2] The following section assesses the evidence cited by Baumeister and Brewer (2012) in addition to a significant body of newer empirical work, most of which points to similar negative outcomes. In reviewing this evidence, I offer a variety of methodological criticisms, especially in terms of how researchers manipulate and measure FWB. I repeatedly highlight problems with using the term 'free will' in experimental contexts and suggest alternative cognitive explanations for why people seem to be worse off when they don't believe in it.

[2] Eddy Nahmias discusses the containment argument, but concludes that it would, in fact, be ethically *ir*responsible *not* to tell people that free will is an illusion.

## Free Will Disbelief and Learning

Belief in free will has been shown to promote increased counterfactual thinking, which has been shown to contribute to increased learning, especially in helping people to re-direct goal-oriented behaviors considering their past mistakes. Counterfactual thinking refers to the tendency to create alternative possibilities for life events that occurred in the past. Alquist et al. (2015) found that priming subjects with FWB produced more unique counterfactual thoughts about a past situation, whereas priming subjects with free will disbelief produced fewer counterfactual thoughts (Cohen's $d = 0.60$). Additionally, subjects who self-reported stronger belief in free will (independent of priming) produced a greater number of counterfactual thoughts.

One major problem with this research is how the authors conceptualize determinism, which they describe as follows:

> "Determinism and other views that oppose free will generally assert that each human action is *inevitable* and that a given person in a given situation could only possibly perform one particular action. With such a belief, there are no legitimate, plausible counterfactuals."

The problem with this framing is the notion that determinism necessarily opposes free will. Further, this framing is inconsistent with the FAD-Plus (Paulhus & Carey, 2011) scale of FWB used in the study. Alquist et al. do not differentiate between participants who endorsed both deterministic statements *and* free will statements versus participants who endorsed *only* free will statements. Consequently, Alquist et al. are only probing incompatibilist notions of free will by failing to include experimental conditions that prime people to believe in *both*

determinism and free will (the compatibilist position). Insofar as there is evidence from folk psychology (e.g. Knobe & Nahmias, 2007) that people hold compatibilist intuitions, this is a significant limitation of the study. Hence, even if Alquist et al. can conclude that belief in free will is closely associated with heightened counterfactual thinking, they cannot simultaneously conclude that belief in determinism is *not*.

Even if my preceding objection is irrelevant, the study only examines the effect of *salient* FWB on counterfactual thinking. It might be the case that when people are actively thinking about free will that counterfactual thinking is impacted; however, this neglects the possibility that most of us don't often think on our FWBs in the same way that we don't often think on our beliefs about dualism, consciousness, the existence of God, etc. Consequently, a committed determinist might be no worse off in counterfactual thinking, if their deterministic beliefs don't frequently occupy their conscious awareness.

One final objection to Alquist et al. is that their work was funded by the pro-conservative, pro-religion John Templeton Foundation. John Templeton Jr., the leader of the foundation, has funded right-wing organizations such as Freedom Watch, a major proponent of the War in Iraq, and Let Freedom Ring, a group tasked with increasing evangelical voter turnout for George W. Bush (Ehrenreich, 2007). [3] The Templeton Foundation explicitly states that it does not fund "projects aimed at

[3] For a detailed discussion, see: https://www.thenation.com/article/john-templetons-universe/

hostility towards religion, or that promote reductionist materialism."[4] Of course, work funded by ideologically-motivated benefactors can still amount to good science. But given the religious significance of free will and the unapologetically reductionist and materialistic views of most free will skeptics (esp. among neuroscientists), it should not be surprising that a study funded by Templeton produced favorable results for FWB.

## Free Will Disbelief and Subjective Well-Being (SWB)

Free will disbelief has been shown to be negatively correlated with subjective well-being, a well-validated psychological measure of people's present happiness and general levels of life satisfaction (Diener et al. 2003). Bergner and Ramon (2013) found positive associations between FWB and non-reductionism, and levels of life satisfaction, though without any clear causal links. Additionally, the study treated free will and determinism as contrasting terms (that is, it framed FWB as incompatibilist). The study used Monroe and Malle's (2010) folk conception of FWB *(e.g. "the capacity to choose based on one's desires and free from constraint").* This conception, however, seems more congruent with compatibilist than incompatibilist accounts: note the emphasis on *choosing among* one's desires (which may still spring from out of the unknown into consciousness); and note there is no mention of any magical or mythical property of free will, as many critics of libertarianism mention (e.g. Cohen & Greene; Harris; Montague). Crescioni et al. (2016) replicated

[4] For information on what Templeton funds and does not fund, see: https://www.templetonworldcharity.org/what-we-fund/what-we-do-and-do-not-fund

findings showing positive correlation between FWB and SWB, ($r(78) = 0.3$) finding that FWBs better predicted subjective well-being than other related psychological constructs, such as having an internal versus external locus of control. The study, however, used the same definition of free will as Bergner and Ramone (2013) and Monroe and Malle (2010), and was consequently probing *one* (compatibilist) version of free will.

Li et al. (2017) strengthened the case for a relationship between FWB and life satisfaction, studying a population of 1,660 Chinese adolescents. Notably, this was the first study to examine the relationship between FWB and SWB in a collectivist, rather than individualist culture. In line with previous findings, FWB was consistently positively associated with higher life satisfaction and positive affect. Cohen's effect size value ($d = 0.28$), however, suggests small practical significance. Additionally, Li et al. (2018) found that belief in free will, even in collectivist cultures, predicts perseverance, especially in terms of adhering to long-term goals. Once again, Cohen's effect size value ($d = 0.38$), suggests small to moderate practical significance.

Like previous research, these studies also used a narrow framing of free will: participants were presented with a two-alternative forced choice paradigm to classify them as either free will believers or determinism believers, which raises the question: what would a compatibilist choose? Furthermore, neither study by Li et al. established evidence for a causal relationship between FWB and correlates of SWB, including moment-to-moment affect, life satisfaction, and perseverance.

Crescioni et al. (2016) established limited evidence suggesting a causal relationship between FWB and perceived meaningfulness in life, which has been shown to promote feelings of social belongingness (Moynihan, Igou, and Tilburg, 2017). Crescioni et al. (2016) primed people to disbelieve in free will by having them rewrite statements exclusively in favor of determinism. As a result, people primed to believe in determinism showed significantly higher scores ($r(67) = 0.31$) on the Kunzendorf no Meaning Scale, a measure of perceived meaninglessness in life (Simon et al., 1998). However, these results cannot be readily generalized given the sample (all college-aged students, 78.6% female) was unrepresentative of the general population.

Gooding, Callan, and Hughes (in press) found that trait-level self-control beliefs predicted variation in subjective well-being better than manipulated free will beliefs. Consequently, it may just be the case that people who feel a greater sense of self-control over their lives experience higher levels of subjective well-being, independent of free will beliefs.

## Free Will Disbelief and Academic Performance

Feldman, Chandrashekar, and Wong (2016) conducted the first direct empirical investigation of the relationship between student's FWB and academic performance, studying the FWB and academic performance of a group of 614 undergraduate students in Hong Kong. Students who reported a strong personal belief in free will correctly identified more mistakes on a spelling task than students who reported a weaker personal belief in free will or no belief in free will ($r(112) =$

0.20). The students' FWBs were measured using a slider scale response score to, "do you have free will?" It is unclear if the students held similar and consistent concepts of free will. However, when the students' course grades and overall GPAs were correlated with their scores on an extended 8-item subscale measuring FWB, a positive correlation was observed between FWB and academic performance ($r(608) = 0.24$). Importantly, FWB predicted academic performance better than other related and well-established predictors, including trait self-control.

Both studies established correlational support for a link between FWB and academic performance but not causal evidence. Further, neither study examined nuances of FWB, including the possibility that some students may have held compatibilist conceptions of free will and others may have held indeterminist libertarian conceptions. Media outlets were quick to exaggerate the findings, though, with one British paper proclaiming in its headlines in 2016, "a belief in free will found to be the key to exam and academic success."

Though not directly related to academic performance, Vohs' and Schooler's (2008) seminal experiment on the relationship between free will and moral behavior suggests that encouraging a belief in determinism could increase the likelihood of students committing academic dishonesty. Vohs and Schooler asked a group of 30 undergraduate students to read a passage from *The Astonishing Hypothesis,* a book by Francis Crick, claiming that free will is an illusion. Participants who read the Crick passage reported weaker FWB than participants who read a control passage. Participants then completed an arithmetic task and were given the opportunity to

cheat; those who read the Crick passage about determinism cheated more frequently ($r(30) = -0.53$).

In a follow-up experiment, 122 undergraduate students were primed with passages endorsing various positions on free will (e.g. free will exists, free will is an illusion, etc.) and then asked to complete an academic test and score their own work. Participants were then told they could reward themselves with a certain amount of money per each correct answer; those primed with determinism were more likely to cheat by over-rewarding themselves with money ($r(71) = 0.47$). A key limitation of the study, however, is that no long-term follow-up tests of the effect of FWB on academic dishonesty were conducted. It is unclear, therefore, if students who rejected free will would continue to cheat more than students who did not.

Additional studies examining the effect of FWB on various behaviors related to academic performance suggest that FWB is positively associated with academic achievement. For example, people who believe in free will have been shown to be more likely to set and work toward long-term goals, a function linked to academic achievement (Seligman et. al., 2013). Additionally, FWB has been shown to help people learn more from their mistakes, possibly because a lack of belief in free will nudges people toward fatalistic assumptions about their abilities to improve, thereby discouraging them from undertaking necessary efforts to improve future performance (Alquist et al., 2015). People who read scientific texts denying free will have been shown to be less attuned to their own errors in performance tasks and consequently less able to alter their performance to minimize the likelihood of

committing future errors. Weakened FWB via priming has been shown to be strongly correlated ($r(20) = 0.73$) with post-error slowing (Rigoni, Wilquin, Brass, & Burle, 2013). Like previously discussed studies, though, these studies also do not examine the long-term relationship between FWB and behaviors related to academic performance.

## Free Will Disbelief and Social Behavior & Attitudes

Priming people to disbelieve in free has been shown to reduce helpful behavior and increase aggression. In one experiment, participants who first read sentences endorsing determinism reported significantly less willingness to help others in hypothetical scenarios than participants who read neutral sentences (i.e. irrelevant to determinism and free will) or sentences endorsing free will (Cohen's $d = 0.64$). Participants in another experiment were asked to prepare food samples for an experimental partner, after reading a taste preference form filled out by their partner. Participants were assigned to two conditions, in which they read sentences about either "human freedom of action or deterministic inevitability." Those who read deterministic inevitability sentences gave significantly more of an unwanted food (a particularly spicy hot sauce) to their experimental partners than participants in the free will condition (Cohen's $d = 0.59$) (Baumeister, Masciampo, and Dewall, 2009).

One key limitation (and not unique to this experiment) is how the study equates determinism with inevitability. Participants may not have been primed with deterministic beliefs, but rather with fatalistic beliefs. "Determinism means

that every action in the quasi-classical universe has a cause," notes J.B. Miles (2011). "In contrast," he adds, "fatalism is a resignation to events that suggests that as everything is determined it is pointless to act because of a belief that no matter what one does one's future will not change." One additional drawback of the study is its funding from the Templeton Foundation, an organization known for pushing a pro-individualist, pro-free will, right-wing agenda, as previously discussed.

Harms et. al (2017) found a more complex relationship between FWB and helping behavior. After reading a popular science article presenting neurological evidence suggesting free will is an illusion, subjects conducted binary dictator games where they could distribute money between themselves a charity. For the total sample, no significant effect was found between induced free will disbelief and generosity. But when the sample was divided by religiosity, participants who self-identified as non-religious showed a significant reduction in charitable giving compared to a control group ($r(108) = -0.54$). In another experiment, participants who read anti-free will statements behaved more impulsively and selfishly, showing reduced intuitive cooperation with other experimental partners, as measured by contributing money to a public pool (Cohen's $d = 0.67$). This effect, however, deteriorated over time (Protzko, Ouimette & Schooler, 2016).

Zhao et al. (2014) studied the relationship between FWB and prejudice, a social attitude strongly linked toward negative treatment of groups of people. Han Chinese people who held stronger FWBs showed less prejudicial attitudes toward Tibetan Chinese ($r(70) = -0.32$). Additionally, white people who were manipulated

to believe in free will showed stronger pro-black attitudes (a proxy for less prejudice toward black people) than white people who were manipulated to disbelieve in free will (Cohen's $d = 0.82$). The study, however, used statements to prime determinism (or disbelief in free will) that could easily have been interpreted as fatalism. For example, "fate determines one's success and failure" and "the future has already been determined by fate" were used. Additionally, the study did not include a neutral control condition. It may be more accurate to say that people who are primed to believe in fatalism, not to disbelieve in free will, show more prejudicial attitudes.

In another Templeton-backed study, Mackenzie, Vohs, and Baumeister (2014) found that weaker belief in free will was associated with feeling less gratitude. People with lower levels of trait free will (as measured by the FAD-Plus) also had lower levels of trait gratitude ($r(91) = 0.22$), as measured by The Gratitude Questionnaire (GQ-6) (McCullough, Emmons, & Tsang, 2002). In an experimental manipulation, participants were primed with pro-free will statements, anti-free will statements, or neutral statements, and then asked to recall three prior autobiographical events for which they felt grateful and to rate their gratitude for each event. Though participants in the pro-free will condition felt significantly higher levels of gratitude when recalling and rating their first autobiographical memory (Cohen's $d = 0.71$), all three groups showed similar levels of gratitude when recalling and rating their second and third autobiographical memories. In a follow-up study, participants received an actual favor during their laboratory session.

Participants in the anti-free will condition felt less gratitude (Cohen's $d = 0.76$), which was mediated by their feeling that the benefactor had less free will and was less sincerely motivated.

The study does not provide direct evidence that people would *express* less gratitude because of free will disbelief, only that people might *feel* less gratitude. Additionally, the study points to unstudied potential upshots of free will disbelief if it is true that people who believe less strongly in free will more strongly doubt the sincerity of other people's motives. For example, free will disbelief might also lead people to attribute less blame or to be more forgiving of transgressions. Additionally, free will disbelief might also help people to discriminate between sincerely-given favors and gifts and insincerely-given favors and gifts (e.g. with strings attached, expectations for reciprocity).

## Free Will Disbelief and Attitudes toward Punishment

A desire to hold people accountable for wrongful behaviors (e.g. breaking laws) may underlie and motivate FWBs. When people consider immoral acts, they report higher FWBs than when they consider morally neutral acts. (Clark et al., 2014). Additionally, people tend to discount the legitimacy of anti-free will research when confronted with evidence about others' immoral behaviors. Heightened punitive motivations drive these belief patterns (Clark et al., 2014). Believing in free will helps to alleviate punitive distress by helping to justify punishment. Free will skeptics tend to show higher levels of anxiety when asked to punish someone, and people who are primed with free will disbelief tend to show higher levels of

anxiety associated with their punitive desires. Additionally, when punishing a person who had no choice in committing their wrongdoing (e.g. a person shown to have diminished free will), punishers tend to feel more anxious (Cohen's $d = 0.68$) (Clark, Baumeister & Ditto, 2017). Consequently, in legal systems, FWBs may lessen the psychological costs of punishment for punishers (e.g. law enforcement, jurors, judges, etc.) at the expense of doling out harsher outcomes for people who are convicted (and wrongfully convicted) of crimes. Priming people with neural determinism may produce an opposite effect; compared to a control group, females (but not males) who read a text defending neural determinism gave fewer electric shocks toward other experimental participants who were described to shock deliverers as deserving of punishment (Cohen's $d = 6.55$) (Caspar, et al, 2017).

Besides altering punitive attitudes, FWBs also influence what types of punishment people favor. People with weaker FWBs tend to favor consequentialist (e.g. restorative justice), not retributive justice (e.g. doing 'hard time'). Additionally, when people learn biological and neural explanations for human behavior, they tend to support retributive punishment less (Cohen's $d = 0.44$). This effect is mediated by changed perceptions of blameworthiness (Shariff et al., 2014). In an analysis of the World Values survey from 65,111 respondents across 46 countries, Martin, Rigoni & Vohs (2017) found that high FWBs predicted support for severe forms of criminal punishment, especially in countries with legal institutions of high integrity (e.g. high political stability, low levels of corruption, government effectiveness).

Free will skeptics and opponents of retributive justice may view the implications of decreased FWB on the legal system as positive, whereas free will believers and supporters of retributive justice may worry that undermining FWBs may also undermine moral norms and social order. Evans (2013) suggests that both consequences are plausible, and that stronger beliefs in determinism are "likely to mollify one of our species' least admirable tendencies involving retributive moral anger, while concomitantly exacerbating one of our worst, namely our tendency toward moral apathy."

If willusionists are correct and/or more people accept the willusionist position, then we must grapple with the downside discussed by Evans. There are many reasons, though, to doubt that the downside of free will disbelief on social morality is insurmountable (or grounds for accepting Smilansky's illusionism). First, free will disbelief has not been shown to prevent people from carrying out justice altogether; rather it has changed how justice is carried out (e.g. less retributive and more consequentialist punishments) (Shariff et al., 2014). Additionally, it is possible that justice systems more strongly rooted in consequentialism might produce better results (e.g. reducing recidivism, preventing victimization, lowering costs). Second, given that FWBs are widespread, especially in WEIRD (western, educated, industrialized, rich, democratic) countries, it is possible there might be a transition phase in which people will adjust to adopting deterministic beliefs, rationalizing punishment and systems of law and order without using free will as a core justification and motivation. Future research could

explore the links between FWB and attitudes toward punishment in cultures that tend to endorse free will less strongly than the West. Third, it is possible (as in other previously-discussed experiments on free will disbelief) that people are simply conflating determinism with fatalism (Strawson, 1986), and are therefore more likely to succumb to moral apathy when presented with evidence in favor of determinism. This possibility further highlights the need for researchers to clarify the methods they use to study FWB and disbelief and to compare how determinism versus fatalism *(rather than just determinism versus free will)* affects attitudes.

## How We Conceptualize and Study Free Will Belief Matters

Studies on the consequences of free will disbelief have generally used three different tools to prime and measure FWB.[5] The original Free Will and Determinism scale (FWD) by Viney et al. (1982) directly contrasts determinism with free will, such that it is impossible to differentiate between people who hold incompatabilist versus compatibilist beliefs. The more recent Free Will and Determinism scale by Rakos et al. (2008) is also predicated on indeterminism. Additionally, validation of the scale was likely statistically underpowered since only 100 participants were used (Kline, 2011). The most commonly used scale, the FAD-Plus (Paulhus & Carey, 2011), presents further problems. For example, scores on its fatalistic determinism subscale correlate with scores on its unpredictability subscale (e.g. how much people believe in randomness), and surprisingly, the

[5] For an excellent and comprehensive critique of the methodological approaches used to study FWB, see "The free will inventory: Measuring beliefs about agency and responsibility" by Nadelhoffer et al. (2014).

fatalistic determinism subscale is not inversely correlated with the free will subscale. Additionally, the FAD-Plus was validated using only a population of predominantly white college students, so it may lack cross-cultural validity. The newest psychometric tool, the Free Will Inventory (FWI), does not show the extent of statistical problems observed in the other tools and was validated by a larger and more representative sample (Nadelhoffer et al., 2014). The FWI, however, has not yet displaced the FAD-Plus as the default tool for measuring and manipulating FWB. It is also possible that further adoption of the FWI will reveal problems and limitations.

One possible methodological consideration in studying FWB is that people may simply not tend to hold coherent and consistent beliefs about free will and determinism. If this is true, then seemingly contradictory correlations within subscales (e.g. fatalistic determinism not being inversely correlated with free will) might reveal an important fact about how people think about free will, not a problem with measurement tools. Free will, fatalism, determinism, randomness, etc. are ideas that have long been debated, refined, and conceptualized. What is intuitive to free will specialists may not correspond to what is intuitive to non-specialists. It could be the case, for example, that many people are intuitive illusionists (they believe in determinism, deep down, yet maintain an illusory belief in free will) to preserve some psychologically important sense of control, moral responsibility, etc., necessary for optimal functioning.

Given that people's beliefs about compatibilism have been shown to be context-sensitive (Bear & Knobe, 2015), it could be the case that people's beliefs about free will, determinism, and similar ideas are also highly context-sensitive—coherent and consistent in one context, incoherent and inconsistent in another. A meta-analysis by Ewusi-Boisvert and Racine (2018) of 52 studies on the implications of FWB found that belief in free will is a dynamic, not static phenomenon, fluctuating in strength and producing different effects in response to internal and external conditions. This should be kept in mind as new psychometric tools to measure FWB are developed and improved.

Additionally, it is important to study free will belief in diverse populations, since beliefs are often influenced by social factors such as culture and religion. The same meta-analysis by Ewusi-Boisvert and Racine (2018) found that samples were disproportionately made up of women, college students and younger aged people, and lacking in ethnic diversity. Consequently, the generalizability of most studies of free will belief is limited.

## Cognitive Accounts of Free Will Belief

Suppose the objections I have raised to the research showing myriad harms of free will disbelief are wrong. Of course, if free will exists and eventually enjoys broad empirical support, these worries are only short-lived. But if free will does not exist (and the debate over its existence settles), or people simply stop believing in free will, then we must consider what cognitive explanations might account for the

effects of free will disbelief. This section discusses theoretical accounts (found within existing literature) for why free will disbelief produces generally undesirable behaviors and feelings. I add my own hypothetical account which focuses on the meaningfulness of choices as a potentially important and neglected component of FWB.

## The Bignetti Model

The Bignetti Model (TBM) provides a functional theoretical account for FWB, explaining how (the alleged illusion of) free will helps people to feel a sense of agency and responsibility associated with their actions and to then appropriately attribute praise and blame, re-enforcing adaptive and functional behavior to improve future action tendencies. TBM posits that the unconscious mind makes choices, of which the conscious mind belatedly becomes aware, and then the conscious mind incorrectly believes that choices were freely decided (Bignetti, 2014). TBM holds that challenging FWBs should interrupt the deliberate and reflective step between the conscious feeling of having made a choice and the objectively false but subjectively true association of that choice with one's sense of agency and sense of responsibility.

## Are Unconscious or Conscious Processes Impacted?

Findings from Rigoni (2013, 2014) and Lynn et al. (2014) challenge this theoretical account, insofar as FWBs affect intrinsic, automatic, pre-reflective cognitive processes, not explicit, deliberate, and reflective cognitive processes. An

alternative theoretical explanation supported by neurocognitive evidence for how FWBs affect behavior is by influencing perceived intentional control and intentional action. Rigoni et al. (2011) presented participants with a text by Francis Crick endorsing determinism, arguing that free will was an illusion. Participants then performed the Libet task and EEG recordings measured readiness potential. When primed with the passage designed to undermine FWBs, participants showed significantly reduced readiness potential, evident more than 1s before the participants consciously decided to move. Rigoni et al. (2013) found that participants who read a text denying free will showed reduced behavioral adjustment after making an error. That is, weakening belief in free will (which is directly associated with perceived intentional control) weakens behavioral control. Additionally, Rigoni et al. (2014) found that in people who were induced to disbelieve in free will, error-related negativity was reduced. Changing beliefs about free will can impair error detection mechanisms, which are basic and automatic processes. Consequently, free will disbelief does not necessarily make people behave badly on purpose, but with less automatic behavioral control.

Free will disbelief affects automatic cognitive processes related to control. People's sense of agency, which is closely related to control, can be diminished because of weakened FWBs. When people are primed to disbelieve in free will, explicit components of agency are not affected but implicit components are reduced (Lynn et al., 2014). Weakened FWBs do not seem to affect deliberate and reflective cognitive processes. Instead, weakened FWBs affect basic, implicit processes central

to how we perceive ourselves as free agents. Pre-reflective cognitive processes, once impaired, may produce most of the suboptimal behavioral changes observed in experiments that manipulate FWB (Lynn et al., 2014). Consequently, and contra TBM, FWBs impair cognition *prior* to conscious deliberation, not during it.

One limitation of these explanations, however, is that they are based almost exclusively on priming experiments. What if cognitive processes differ when people have not been recently primed to think about free will? It is possible, for example, that cognitive adaptation occurs and that implicit processes affected by primed and salient free will disbelief return to normal functioning once beliefs about free will have receded from attention. Future research could conduct follow-up tests to test durability of effects in addition to conducting experiments more cleverly blinded such that participants are not likely to bring FWBs into consciousness and attention.

## Free Will, Choice, and Personal Meaning

Current theoretical accounts of FWB do not adequately differentiate between types of choices and actions; they treat choice as a homogeneous psychological phenomenon (e.g. Feldman, Baumeister, and Wong, 2014). I propose that believing in free will holds more subjective significance for personally meaningful choices and actions than personally non-meaningful ones. Meaningful choices and actions relate intimately to a person's sense of self, aspirations, goals, etc.—they are choices which usually carry significant emotional weight, have significant implications for our future self (or loved ones), and often involve a moral dimension. Choosing whom to

marry, where to live, what to vote for, what to study, what job to take, and how to raise children are examples of personally meaningful choices. Choosing what brand of toothpaste to purchase, how bright to set the dimmer light-switch, what color pen to use, whether to sneeze into one's elbow or hands, and which side of bed to get out of in the morning are examples of non-meaningful choices. If we were to learn that we lack free will when making non-meaningful choices, I hypothesize we would be less negatively affected than we would be if we were to learn that we lack free will when making meaningful choices.

This is not, however, how free will disbelief is instilled and presented in experiments that provide the basis for theoretical accounts. Experiments prime free will disbelief using a passage from *The Astonishing Hypothesis* by Francis Crick, reading, "although we appear to have free will, in fact, our choices have already been predetermined for us and we cannot change that." Or experiments use items from the FAD-Plus, reading, "no matter how hard you try, you can't change your destiny," or "whatever will be, will be – there's not much you can do about it," or "your genes determine your future" (Paulhus, 2011). Consequently, experiments do not adequately prime people to believe or disbelieve in free will on a spectrum.

Therefore, given the existing literature, it is not reasonably possible to determine *what* choices FWB matters for, or *in what contexts* FWB is about agency and responsibility. This is an unexplored frontier in the study of FWB. It might be the case, for example, that people would be completely unaffected if they learned that most choices and decisions we make are not free in any meaningful sense, but

are made unconsciously and implicitly. As far as FWB is concerned, all choices may not be equal if free will is in fact about choosing. On the contrary, a disproportionately small number of choices (personally meaningful choices) may account for the functional importance of FWB.

## Conclusion

Throughout this paper, I have tried to show that free will, regardless of whether it is an illusion, is part of most people's belief systems. I have also tried to show how academic conceptions and studies of free will belief have often treated free will belief as something it is not, removing it too far from ecologically valid contexts and failing to do justice to the complex, dynamic nature of the mind.

Willusionists arguing against the existence of free will, social psychologists arguing for the harms of willusionism, and theorists explaining free will in terms of agency, control, and choice, have tended to make similar missteps. Perhaps the most common and problematic misstep is how free will is conceptualized, defined, and instrumentalized. Many willusionists argue against a type of free will (libertarian, agent-causation) which has fallen out of intellectual favor and doesn't reflect how ordinary people generally think about free will. Many social psychologists conflate determinism with fatalism and tend to study hypotheses that focus on the harms of free will disbelief rather than the plausible benefits. Theorists frequently treat choice as a homogeneous psychological phenomenon, building theories from empirical results found using problematic methodologies. These

problems likely arise because free will has long been defined philosophically, and has only recently been considered as a scientific concept (Brembs, 2011).

Clarifying definitions, examining the nuances of FWB, and building more precise theoretical accounts carry extensive social implications. Arguably, findings within the cognitive sciences have progressively cast more doubt on the extent and existence of free will. Scientists have continued to uncover more evidence that points to the genetic and environmental determinants of behavior and the outsize role of the unconscious mind. Consequently, the prevalence of free will skepticism and disbelief seems likely to rise. Though I have attempted to show that the current literature describing harms of free will disbelief is unpersuasive, it is possible that once better methodologies are used and concepts are disentangled, researchers will find that free will disbelief does threaten individual and social well-being. If such findings prove to be robust, then we must reconcile, both individually and as a society, with how to preserve social order, foster psychological well-being, and promote positive behaviors while accepting the nonexistence of free will, assuming we cannot maintain the illusion considering overwhelming evidence against it. Understanding the underlying cognitive mechanisms of FWB and disbelief offers a path toward reconciliation and it will best be accomplished by clarifying concepts, improving empirical tools, and refining theories.

## Author Contributions

THB suggested to review social psychological research on free will disbelief and to look for methodological flaws. KH conducted a literature review. After further discussion, THB suggested to include statistical quantifications of correlation strengths and effect sizes.

## Acknowledgments

I am grateful to my adviser, Thomas H. Brown, for providing guidance and feedback throughout the writing process. I am also grateful to Mark Sheskin for quickly and thoughtfully answering my myriad logistical questions. Additionally, my peer reviewers provided generous, candid, and useful advice. My friends Zhen Huang, Marina Tan, and Mojmír Stehlik were instrumental in improving the quality of my writing. I would also like to thank Joshua Knobe for inspiring my interest in Cognitive Science, especially because of his ability to draw on modern science and timeless philosophical traditions to provide clear but profound insights into human nature. Finally, I thank my mom, dad, and extended family for making my education possible and for showing me how to be a good student and person.

## References

Ali, A. (2016). A belief in free will found to be the key to exam and academic success for students, claims new study. Retrieved from https://www.independent.co.uk/student/student-life/Studies/a-belief-in-free-will-found-to-be-the-key-to-exam-and-academic-success-for-students-claims-new-study-a6799436.html

Alquist, J. L., Ainsworth, S. E., Baumeister, R. F., Daly, M., Stillman, T. F., & Will, F. (2015). The Making of Might-Have-Beens: Effects of FWB on Counterfactual Thinking. *Personality and Social Psychology Bulletin* Vol 41, Issue 2, pp. 268 – 283 https://doi.org/10.1177/0146167214563673

Baer, J. (2008). Free will requires determinism. In J. Baer, J. C. Kaufman, & R. F. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 308–09). Oxford: Oxford University Press.

Baumeister, R. F., & Brewer, L. E. (2012). Believing versus Disbelieving in Free Will: Correlates and Consequences, *Social and Personality Psych. Compass*, *10*, 736–745.

Baumeister, R. F., Masciampo, E. J., & Dewall, C. N. (2009). Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness, *35*(2), 260–268. https://doi.org/10.1177/0146167208327217

Bargh, J. (2008). Free will is un-natural. In *Are we free? Psychology and free will, ed*. J. Baer, J. Kaufmann, and R. Baumeister, 128–154. New York: Oxford University Press.

Bear, A.; Knobe, J. (2015). What Do People Find Incompatible with Causal Determinism?. Cognitive Science 40(8): 2025-2049

Brembs, B. (2011). Towards a scientific concept of free will as a biological trait: spontaneous actions and decision-making in invertebrates. *Proceedings of the Royal Society B*, 278, 930e939.

Bourget, David & Chalmers, David J. (2014). What do philosophers believe? *Philosophical Studies* 170 (3):465-500.PhilArchive copy v6: https://philarchive.org/archive/BOUWDPv6

Caspar, E. A., Vuillaume, L., Magalhães De Saldanha Da Gama, P. A., Cleeremans, A. (2017). The influence of (dis)belief in free will on immoral behaviour. *Frontiers in Psychology*, 8, Article 20.

Chisholm, R. (1966) "Freedom and Action," in *Freedom and Determinism*, K. Lehrer (ed.), New York: Random House, pp. 11–44; reprinted in Myles Brand (ed.), *The Nature of Human Action*, Glenview, IL: Scott Foresman and Company, 1970, pp. 283–292

Clark, C. J., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: A motivated account of FWB. *Journal of Personality and Social Psychology, 106*(4), 501-513. http://dx.doi.org/10.1037/a0035880

Crick, F. (1994) *The astonishing hypothesis*, Scribner's.

Dennett, D. C. (1984). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: MIT Press

Dennett, D. C. (2003). *Freedom Evolves*. New York: Penguin Group (Viking).

Diener, E., Oishi, S., and Lucas, R. E. (2003). Personality, culture, and subjective well-being: emotional and cognitive evaluations of life. *Annu. Rev. Psychol.* 54, 403–425. doi: 10.1146/annurev.psych.54.101601.145056

Ehrenreich, B. (2015). John Templeton's Universe. Retrieved from https://www.thenation.com/article/john-templetons-universe/

Ekstrom, L.W. (2000). *Free Will: A Philosophical Study*, Boulder: Westview Press.

Ewusi-Boisvert, E. & Racine, E (2018). A Critical Review of Methodologies and Results in Recent Research on Belief in Free Will. *Neuroethics*. 11: 97. https://doi.org/10.1007/s12152-017-9346-3

Evans, J., & Evans, J. (2017). The moral psychology of determinism. *5089*(October). https://doi.org/10.1080/09515089.2012.713163

Feldman, G., Baumeister, R. F., & Wong, K. F. E. (2014). Free will is about choosing: The link between choice and the belief in free will. *Journal of Experimental Social Psychology*, 55, 239–245.

Feldman, G., Chandrashekar, S. P., Wong, K. F. E. (2016). The freedom to excel: Belief in free will predicts better academic performance. *Personality and Individual Differences*, 90, 377-383.

Feldman, G. (2017). Making sense of agency: Belief in free will as a unique and important construct. *Social and Personality Psychology Compass* 11. https://doi.org/10.1111/spc3.12293

Feldman, G. and Chandrashekar, S.P. (2017). Laypersons' Beliefs and Intuitions About Free Will and Determinism: New Insights Linking the Social Psychology and ... Laypersons' Beliefs and Intuitions About Free Will and Determinism. *Soc. Psych. and Personality Sci.,* (May). https://doi.org/10.1177/1948550617713254

Feltz, A. (2015). Experimental philosophy of actual and counterfactual free will intuitions. *Consciousness and Cognition, 36*, 113–130

Frankfurt, H. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68, 5–20.

Gazzaniga MS. (2011). *Who's In Charge?* New York: HarperColins.

Ginet, C. (2002) Reasons Explanations of Action: Causalist versus Noncausalist Accounts, in *The Oxford Handbook of Free Will*, Robert Kane (ed.), New York: Oxford University Press, 386–405.

Guggisberg AG., Mottaz A. (2013) Timing and awareness of movement decisions: Does consciousness really come too late? *Front Hum. Neurosci.* 7:385

Gooding, PLT., Callan, M., Hughes, G. (In press). The association between believing in free will and subjective well-being is confounded by a sense of personal control. *Frontiers in Psychology*.

Harms J, Liket K, Protzko J, Scholmerich V (2017) Free to help? An experiment on FWB and altruism. *PLoS ONE* 12(3): e0173193. doi:10.1371/journal.pone.0173193

Joordens, S., van Duijn, M. & Spalek, T.M. (2002). When timing the mind one should also mind the timing: Biases in the measurement of voluntary actions. *Conscious. Cogn.* 11, 231–240 (2002).

Kane, R. (1996). *The Significance of Free Will*, New York: Oxford University Press.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press

Knobe, J. (2014) Free Will and the Scientific Vision. *Current Controversies in Experimental Philosophy*. Routledge.

Lavazza, A. (2016). Free will and neuroscience: from explaining freedom away to new ways of operationalizing and measuring it. *Front. Hum. Neurosci.*10:262. doi: 10.3389/fnhum.2016.00262

Li, C., Wang, S., Zhao, Y., Kong, F., Li, J., Webb, T. L., & Crescioni, A. W. (2017). The Freedom to Pursue Happiness: Belief in Free Will Predicts Life Satisfaction and Positive Affect among Chinese Adolescents, *7*(January), 1–8. https://doi.org/10.3389/fpsyg.2016.02027

Li, J., Zhao, Y., Lin, L., Chen, J., & Wang, S. (2018). Personality and Individual Differences The freedom to persist: Belief in free will predicts perseverance for long-term goals among Chinese adolescents. *Personality and Individual Differences*, *121*(August 2017), 7–10. https://doi.org/10.1016/j.paid.2017.09.011

Libet, Benjamin; Gleason, Curtis A.; Wright, Elwood W.; Pearl, Dennis K. (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. doi:10.1093/brain/106.3.623. PMID 6640273.

Libet, Benjamin (1985). Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. *The Behavioral and Brain Sciences*. **8**: 529–566. doi:10.1017/s0140525x00044903

Lynn, M. T., Muhle-Karbe, P. S., Aarts, H., and Brass, M. (2014). Priming determinist beliefs diminishes implicit (but not explicit) components of self-agency. *Front. Psychol.* 5:1483. doi:10.3389/fpsyg.2014.01483

MacKenzie, M. J., Vohs, K. D., & Baumeister, R. F. (2014). You didn't have to do that: Belief in free will promotes gratitude. *Personality and Social Psychology Bulletin*, *40*(11), 1423–1434. doi:10.1177/0146167214549322

Martin, N. D., Rigoni, D., Vohs, K. D. (2017). FWBs predict attitudes toward unethical behavior and criminal punishment. *Proceedings of the National Academy of Sciences*, 114, 7325-7330.

McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, 82, 112-127.

Miles, J. B. (2013). 'Irresponsible and a Disservice': the integrity of social psychology turns on the free will dilemma. *Br. J. Soc. Psychol.* 52, 205–218. doi: 10.1111/j.2044-8309.2011.02077.x

Morris, S. (2017). The implications of rejecting free will: An empirical analysis. *Philosophical Psychology* 75, 1-23.

Montague, R. (2008). Free will. *Current Biology* 18: R584–585

Moynihan, A. B., Igou, E. R., van Tilburg, W. A. (2017). Free, connected, and meaningful: FWBs promote meaningfulness through belongingness. *Personality and Individual Differences*, 107, 54–65.

Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., Ross, L. T. (2014). The free will inventory: Measuring beliefs about agency and responsibility. *Consciousness and Cognition*, 25, 27–41. doi:10.1016/j.concog.2014.01.006

Nahmias, E. (2011). Why 'Willusionism' Leads to 'Bad Results': Comments on Baumeister, Crescioni, and Alquist. *Neuroethics*. 4. 17-24. 10.1007/s12152-009-9047-7.

Nahmias E, Shepard J, Reuter S. (2014) It's OK if "my brain made me do it": People's intuitions about free will and neuroscientific prediction. *Cognition*. 133(2):502–16. doi: 10.1016/j.cognition.2014.07. 009 PMID: 2519507716.

Nahmias, E. (2017). Free will as a psychological accomplishment. In D. Schmidtz & C. Pavel (Eds.), *Oxford handbook on freedom*. Oxford University Press.

Nichols, S. and Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs,*, 41: 663–685. doi:10.1111/j.1468-0068.2007.00666.x

Nichols, S. (2011). Experimental philosophy and the problem of free will. *Science.* 331 (6023): 1401 – 1403.

O'Connor, T. (2010). Agent-Causal Theories of Freedom, in Robert Kane (ed.) *Oxford Handbook on Free Will*, 2nd edition. New York: Oxford University Press.

Paulhus, D. L., Carey, J. M. (2011). The FAD-Plus: Measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment*, 93, 96–104

Pink, T. (2004). *Free Will: A Very Short Introduction*, Oxford: Oxford University Press.

Protzko J, Ouimette B, Schooler J (2016) Believing there is no free will corrupts intuitive cooperation. *Cognition* 151:6–9.

Rakos, R. F., Laurene, K. R., Skala, S., & Slane, S. (2008). Belief in free will: Measurement and conceptualization innovations. *Behavior and Social Issues*, 17, 20–39.

Ravven, H. M. (2013). *The Self Beyond Itself: An Alternative History of Ethics, The New Brain Sciences, and the Myth of Free Will*. New York. New Press.

Rigoni, D., Pourtois, G., & Brass, M. (2014). 'Why should I care?' Challenging free will attenuates neural reaction to errors. *Social Cognitive and Affective Neuroscience*. 262–268. https://doi.org/10.1093/scan/nsu068

Rigoni, D., Kühn, S., Sartori, G., and Brass, M. (2011). Inducing disbelief in free will alters brain correlates of preconscious motor preparation. *Psychol. Sci.* 22, 613–618. doi: 10.1177/0956797611405680

Rigoni, D., Wilquin, H., Brass, M., and Burle, B. (2013). When errors do not matter: weakening belief in intentional control impairs cognitive reaction to errors. *Cognition* 127, 264–269. doi: 10.1016/j.cognition.2013.01.009

Ryle, G. (2009). *The Concept of Mind*. London: Routledge

Saigle, V., Dubljević, V., & Racine, E. (2018) The Impact of a Landmark Neuroscience Study on Free Will: A Qualitative Analysis of Articles Using Libet and Colleagues' Methods*, AJOB Neuroscience*, 9:1, 29-41, DOI: 10.1080/21507740.2018.1425756

Sarkissian, H., Chatterjee, A., De brigard, F., Knobe, J., Nichols, S., & Sirker, S. (2010). Is belief in free will a cultural universal? *Mind and Language*, *25*(3), 346-358. DOI: 10.1111/j.1468-0017.2010.01393.x

Schopenhauer, A. (1839). *Prize Essay on the Freedom of the Will.* Edited by Gunther Zoller, Translated by E.F. J. Payne. Cambridge: Cambridge University Press, 1999.

Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J., Clark, C., Schooler, J. W.,. Vohs, K. D. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. Psychological Science, 25, 1563-1570

Shepard, J. & O'Grady, A. (2017). What kinds of alternate possibilities are required for the folk concept(s) of choice? *Consciousness and Cognition*, 48, 138-148.

Soon, C. S. *et al.* (2008). Unconscious determinants of free decisions in the human brain. *Nature Neurosci.* 11, 543–545

Smilansky, S. (2001) Free will: from nature to illusion. *Proceedings of The Aristotelian Society*, 101, 71–95.

Strawson, G. (1986). Freedom and belief. Oxford: Clarendon Press.

Stump, E. (1997). Aquinas's Account of Freedom: Intellect and Will. *Monis*t 80: 576 –97

Tse P.U. (2013) *The neural basis of free will: Criterial causation.* Cambridge, MA: MIT Press.

Turri, J. (in press). Exceptionalist naturalism: Human agency and the causal order. *Quarterly Journal of Experimental Psychology*.

Tversky, A. &. Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science* 211:453–58.

Vohs, K.D., Schooler, J.W. (2008). The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science*, 19, 49–54.

Vonasch, A.J., Baumeister, R.F., Mele, A.R. (2018) Ordinary people think free will is a lack of constraint, not the presence of a soul. *Consciousness and Cognition*, 60: 133-151.

Wegner, D. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Zhao, X., Liu, L., Zhang, X. X., Shi, J. X., Huang, Z. W. (2014). The effect of belief in free will on prejudice. *PLoS One,* 9(3), e91572.