Personality and Social Psychology Review
1–31

DOI: 10.1177/10888683221087527
pspr.sagepub.com
SAGE

Manipulating Belief in Free Will and Its Downstream Consequences: A Meta-Analysis

Oliver Genschow[1*], Emiel Cracco[2*], Jana Schneider[1], John Protzko[3], David Wisniewski[2], Marcel Brass[4], and Jonathan W. Schooler[5]

Abstract

Ever since some scientists and popular media put forward the idea that free will is an illusion, the question has risen what would happen if people stopped believing in free will. Psychological research has investigated this question by testing the consequences of experimentally weakening people's free will beliefs. The results of these investigations have been mixed, with successful experiments and unsuccessful replications. This raises two fundamental questions: Can free will beliefs be manipulated, and do such manipulations have downstream consequences? In a meta-analysis including 145 experiments (95 unpublished), we show that exposing individuals to anti–free will manipulations decreases belief in free will and increases belief in determinism. However, we could not find evidence for downstream consequences. Our findings have important theoretical implications for research on free will beliefs and contribute to the discussion of whether reducing people's belief in free will has societal consequences.

Keywords

free will, determinism, belief, meta-analysis, morality, cheating, social behavior, punishment

From morality to politics, public policy, intimate relationships, and punishing behavior—"most of what is distinctly human about our life depends upon our viewing one another as autonomous persons, capable of free choice" (Harris, 2012, p. 1). Thus, unsurprisingly, most people believe that they have free will (Baumeister et al., 2009; Nahmias et al., 2005). Whether free will actually exists, however, has been a long-standing philosophical debate (e.g., Dennett, 2015; Van Inwagen, 1983). This debate has reached an extremely high level of sophistication outlining different theoretical positions that span free will skepticism to complete libertarianism (for an overview, see Dennett, 2015). However, these philosophical arguments have rarely left academic circles and therefore have had limited impact outside academia.

In the last decades, cognitive neuroscientists and psychologists entered the debate by claiming humans' perception of free will is nothing more than an illusion (e.g., Crick, 1994; Harris, 2012; Wegner, 2002) and simply arises from unconscious brain activity (Hallett, 2007; Libet et al., 1983; Soon et al., 2008). A seminal study supporting this view is the study by Libet and colleagues (1983), who measured neural activity while participants made voluntary finger movements. After each movement, participants indicated on a clock the time at which they perceived their first urge to initiate a movement. Libet and colleagues found an increase in neural activity several hundred milliseconds before participants reported being aware of this urge. These and similar findings (Libet et al., 1983, 1993) have often been used as an argument for the claim that free will does not exist.

Despite criticisms of this argument (Brass et al., 2019; Saigle et al., 2018), anti–free will viewpoints have become in vogue not only in academia (e.g., Greene & Cohen, 2004) but also, and perhaps even more so, in popular media (e.g., Chivers, 2010; Griffin, 2016; Racine et al., 2017; Wolfe, 1997). This raises the fundamental question of whether reading anti–free will viewpoints pushes people toward a deterministic worldview and lowers their belief in free will. Moreover, the question arises what impact this has on society.

Some philosophers have argued that undermining people's belief in free will saps the basis for moral behavior and

[1]University of Cologne, Germany
[2]Ghent University, Belgium
[3]Central Connecticut State University, New Britain, USA
[4]Humboldt University of Berlin, Germany
[5]University of California, Santa Barbara, USA
*These authors shared first authorship.

Corresponding Author:
Oliver Genschow, Social Cognition Center Cologne, University of Cologne, Richard-Strauss Str. 2, Cologne 50931, Germany.
Email: oliver.genschow@uni-koeln.de

would therefore have catastrophic consequences (e.g., Smilansky, 2000, 2002). In contrast, other philosophers have argued that disbelief in free will might instead have positive effects, as it would cause people to abandon retribution-based morality and illusory beliefs in a just world (e.g., Caruso, 2014; Greene & Cohen, 2004; Nadelhoffer, 2011; Pereboom, 2006). Research in social and cognitive psychology as well as in neuroscience has empirically tested these theories by experimentally manipulating belief in free will (for an overview, see Ewusi-Boisvert & Racine, 2018). These studies provided evidence for the idea that free will beliefs indeed have societal consequences. There have also been a number of failures to replicate some of these findings, however. As a result, it remains unclear whether exposing people to anti–free will viewpoints has behavioral and societal consequences. In this article, we address this controversy. First, we review the literature on belief in free will, its experimental manipulations, and its downstream consequences. Then, we test meta-analytically (a) whether beliefs related to free will can be manipulated and (b) whether these manipulations have downstream consequences.

Belief in Free Will and Determinism

Philosophical definitions regarding the concept of free will are rather diverse and complex (Carey & Paulhus, 2013). Because of this, researchers often look at laypeople's concepts of free will and to which degree laypeople believe in these concepts (e.g., Nichols, 2006). These analyses indicate that laypeople's free will beliefs are metacognitive judgments about the extent to which individuals intentionally guide their thoughts and actions (Frith, 2012). Specifically, belief in free will reflects the belief that people are responsible for their actions (Carey & Paulhus, 2013) because they can decide and control their own behavior (Paulhus & Carey, 2011).

Although conceptually related to internal locus of control (Rotter, 1966) and self-efficacy (Bandura, 1977), there are important differences between these constructs that set free will beliefs apart. Internal locus of control shares the emphasis on internal causal attribution but represents a personality dimension, whereas belief in free will reflects an attitudinal orientation (Waldman et al., 1983). Self-efficacy differs from belief in free will in that it reflects metacognitive judgments about one's specific skill or ability (e.g., "Can I execute this successfully?"). In contrast, belief in free will reflects a much broader belief about choice and freedom (e.g., "Do I have a choice? Can I freely choose to do otherwise?").

Related to the belief in free will is the belief in determinism. Previous research has defined belief in determinism in different ways. On a general level, determinism is typically interpreted as: given the past and the laws of nature, there is only one possible future at any moment in time (e.g., Van Inwagen, 1983). Some scholars differentiate between fatalistic determinism, which captures people's belief in "destiny" or "fate," and scientific determinism, which measures belief in genetic determinism (e.g., Paulhus & Carey, 2011) and reflects the belief that genes are the sole or predominant basis of personal characteristics (Keller, 2005).

Irrespective of how determinism is defined, psychologists and philosophers debate whether free will and determinism are the endpoints of the same continuum or separate constructs. For instance, incompatibilists see free will and determinism as mutually exclusive. Such a view suggests that the more a person believes in free will, the less they believe in determinism (Rakos et al., 2008; Viney et al., 1982). In contrast, compatibilists see free will and determinism as independent constructs. Based on a compatibilistic view, if determinism were true, people could nevertheless be free. Previous research has shown that belief in free will is not correlated with belief in genetic determinism (Nadelhoffer et al., 2014) and that a compatibilistic view is more widespread in the general public than philosophers and psychologists may have traditionally assumed (e.g., Monroe & Malle, 2010a; Murray & Nahmias, 2014; Nadelhoffer et al., 2014; Nahmias et al., 2006; Nichols, 2004, 2006; Nichols & Knobe, 2007; Rose & Nichols, 2013; Shepard & Reuter, 2012; Shepherd, 2012).

Research on Manipulating Belief in Free Will

To test the consequences of (dis)believing in free will, researchers have developed various approaches to experimentally manipulate people's belief in free will. The first investigations in this respect were carried out by Vohs and Schooler (2008). In one experiment, participants either read a passage from Francis Crick's (1994) book *The Astonishing Hypothesis*, arguing against the plausibility of free will, or a passage from the same book that did not mention free will. In another experiment, the authors followed a Velten-like technique (Velten, 1968), in which participants read and pondered either anti–free will or pro–free will statements. In both experiments, participants who were presented with anti–free will viewpoints reported lower beliefs in free will and were more willing to cheat on a test than control participants.

These findings inspired researchers all around the world to start investigating the consequences of experimentally reducing belief in free will by applying and adopting the manipulations introduced by Vohs and Schooler (2008). This research suggests that weakening belief in free will increases antisocial behavior, such as prejudice (Zhao et al., 2014) or aggressiveness toward others (Baumeister et al., 2009), and decreases prosocial behavior, such as helping (Baumeister et al., 2009) or cooperation (Protzko et al., 2015). At the same time, however, exposing people to anti–free will viewpoints can also lead to reduced retributive punishment (Shariff et al., 2014).

In addition, anti–free will manipulations have been found to increase conformity (Alquist et al., 2013) and feelings of alienation (Seto & Hicks, 2016), and to decrease causal attributions of other people's actions (Genschow et al., 2017a), the perceived meaningfulness of life (Crescioni et al., 2016; Moynihan et al., 2019), perceived gratitude (MacKenzie et al., 2014), counterfactual thinking (Alquist et al., 2015), and risk-taking behavior (Schrag et al., 2016). Finally, a last line of research suggests that experimentally reducing people's belief in free will influences neurocognitive processes such as intentional action preparation (Rigoni et al., 2011), deliberate motor inhibition (Lynn et al., 2013; Rigoni et al., 2012), and the processing of performance errors (Rigoni et al., 2013, 2015).

In sum, there is a large body of research suggesting that manipulating belief in free will affects societally relevant behaviors such as cheating (Vohs & Schooler, 2008), retributive punishment (Shariff & Vohs, 2014), and antisocial behavior (Baumeister et al., 2009; Zhao et al., 2014), as well as basal neurocognitive mechanisms (Rigoni & Brass, 2014). Based on this research, it has been argued that one should be careful in how anti–free will and deterministic viewpoints are presented to society because it may change the way people interact with each other. For example, some scholars suggested that encountering anti–free will viewpoints in the popular press may "move judges and jurors toward being less punitive and less retributive in general" (p. 1569, Shariff et al., 2014) or "provide the ultimate excuse to behave as one likes" (p. 54, Vohs & Schooler, 2008).

Failed Replications

Despite the mounting evidence that manipulating belief in free will influences behavior, a number of studies have reported difficulties in replicating some key results (Crone & Levy, 2019; Eben et al., 2020; Genschow et al., 2020; Giner-Sorolla et al., 2016; Monroe et al., 2017; Nadelhoffer et al., 2020; Open Science Collaboration, 2015; Schooler et al., 2014; Shariff & Vohs, 2014; Zwaan, 2014). For example, Monroe et al. (2017) found no effect of diminishing participants' belief in free will on moral behavior, judgments of blame, or punishment decisions. Similarly, Nadelhoffer et al. (2020) found that manipulating free will beliefs in a robust way is more difficult than implied by previous work and that the proposed link with immoral behavior, such as cheating, for instance, might be similarly tenuous (for similar findings, see Crone & Levy, 2019; Giner-Sorolla et al., 2016; Zwaan, 2014).

Although these failed replications call into question the societal relevance of belief in free will, it is not yet clear what caused them. Before we can draw conclusions about the role of free will beliefs in society, it is imperative to understand why some findings failed to replicate as well as which mechanisms underlie free will belief manipulations. In principle, three explanations could account for the replication failures reported in the literature. First, it could be that the failed replications are false negatives. That is, they were not able to detect an effect that is actually real. Second, it might be that free will beliefs cannot be manipulated and that successful studies in the literature are therefore false positives. Third, it could be that manipulations of belief in free will successfully affect free will beliefs, but that these manipulations are not causally related to other behaviors, and thus have no downstream consequences. In this article, we investigate these explanations by analyzing all available evidence in a meta-analysis.

Previous Reviews and Meta-Analyses

To the best of our knowledge, this is the first comprehensive meta-analysis on the effectiveness of manipulations related to free will beliefs. There are, however, two articles related to our analysis. First, Ewusi-Boisvert and Racine (2018) published a qualitative review of the literature on free will belief manipulations. Overall, the authors report a substantial amount of methodological diversity and a lack of replication studies in the published literature. Moreover, the review suggests that the studied samples are heavily constituted of women, students, and younger participants and contain little information about the representation of ethnic minorities. This review did not allow the accurate estimation of the effect size of free will belief manipulations and its downstream consequences, however, as the researchers provided a descriptive review of the literature and did not include unpublished data.

Second, Genschow and colleagues (2017a) tested the effectiveness of one specific manipulation (i.e., the Crick manipulation) in a mini meta-analysis involving nine published and unpublished experiments that were conducted in their own research group (N = 625 participants). Overall, the authors found that participants who read the anti–free will text had a significantly lower belief in free will than participants who read the control text. However, this analysis did not allow strong conclusions regarding the general effectiveness of free will belief manipulations because only one manipulation type, a small number of experiments, one measure, and data from only one research team were investigated. In addition, this analysis did not explore potential reasons for why the manipulation sometimes works and sometimes fails, nor did it investigate whether the manipulation has any downstream consequences.

The Present Meta-Analysis

The present meta-analysis aims to build on and considerably extend previous work by including both published and unpublished evidence and addressing two main research questions:

Research Question 1: Can belief in free will be experimentally manipulated?
Research Question 2: Does this have any downstream consequences?

Research Question 1: Can Belief in Free Will Be Manipulated?

In the first part of the meta-analysis, we investigate whether it is possible to experimentally manipulate beliefs related to free will and explore the conditions under which the manipulations are effective. To this end, we investigated different moderators.

Beliefs. The two most often used measures to test the effectiveness of free will belief manipulations are belief in free will scales and belief in determinism scales (Ewusi-Boisvert & Racine, 2018). It is often assumed (without giving explicit explanations) that anti–free will manipulations should not only decrease free will beliefs but should also increase beliefs in determinism (for an overview, see Ewusi-Boisvert & Racine, 2018). This assumption, however, has never been systematically tested. Therefore, in the present meta-analysis, we tested whether anti-free will manipulations decrease the belief in free will, increase the belief in determinism, or both.

Scales. Several validated scales have been developed to measure belief in free will and determinism. The most frequently used scales are the FWD scale (Rakos et al., 2008) and the FAD scale (Paulhus & Carey, 2011), which both stand for Free Will and Determinism Scale, and the Free Will Inventory (FWI; Nadelhoffer et al., 2014). In addition, researchers have sometimes also used self-made rating scales (e.g., Baumeister et al., 2009; Moynihan et al., 2019). Less frequently used scales measuring beliefs related to free will and determinism include the Belief in Genetic Determinism scale (BGD; Keller, 2005), the Belief in Social Determinism scale (BSD; Rangel & Keller, 2011), and the WiF scale (Melcher, 2019). However, while a wide range of scales has been used, it remains unknown whether certain scales are more sensitive in picking up the effects of the manipulation than others. To address this question, the present meta-analysis tests whether the effects of free will belief manipulations differ across scales. We focus on the FWI and the FAD, as these are the only two scales that tease apart belief in free will and belief in determinism. Moreover, they are also the two most commonly used scales and the only scales for which sufficient data are available to reliably compare them with each other.

In addition to the sensitivity of different scales, little is known about whether free will belief manipulations specifically affect beliefs in free will and determinism or also influence other, related beliefs. Therefore, we investigated whether the influence of free will belief manipulations extends to other beliefs that are related to free will and are measured in free will questionnaires, such as belief in dualism, fatalistic determinism, and unpredictability.

Type of manipulation. Not only the scale but also the type of manipulation differs across studies. Specifically, four types of manipulations can be distinguished. The first method is to let participants read a text—either a control text or a text arguing against the plausibility of free will. Frequently used anti–free will texts include a passage of Francis Crick's (1994) book *The Astonishing Hypothesis* (e.g., MacKenzie et al., 2014; Rigoni et al., 2011; Shariff et al., 2014; Vohs & Schooler, 2008) or articles featuring neuroscientific experiments (e.g., Harms et al., 2017; Protzko et al., 2016; Shariff et al., 2014), among others. The second method is to present participants with several statements using a Velten-like technique (Velten, 1968). In this method, participants are presented either with a variety of anti–free will statements or with a variety of control statements (e.g., Rigoni et al., 2012; Stillman et al., 2010; Vohs & Schooler, 2008). The third method uses a combination of text and statements (e.g., Seto & Hicks, 2016). Finally, the fourth method is to let participants watch videos related to (anti–)free will viewpoints (e.g., Highhouse & Rada, 2015). To reliably manipulate belief in free will, it is important to know which of these manipulations are most effective. Therefore, we directly compared them in the present meta-analysis.

Participant involvement. Whether the manipulation is effective may, however, not only depend on the type of manipulation but also on participants' involvement in the manipulation. While some researchers merely presented participants with anti–free will viewpoints (e.g., Baumeister et al., 2009; Goodyear et al., 2016; Rigoni et al., 2012; Shariff et al., 2014; Stillman et al., 2010), others engaged participants more strongly by letting them, for instance, summarize or rewrite the presented messages immediately after reading them (Harms et al., 2017; Highhouse & Rada, 2015; Moynihan et al., 2019; Rigoni et al., 2011; Seto & Hicks, 2016; Vonasch et al., 2017). Yet other researchers told participants they had to attend the presented messages carefully so they could summarize them at the end of the experiment (e.g., Genschow et al., 2017a). An interesting question is how deeply participants have to process the provided information for the manipulation to succeed. To answer this question, we investigated if the effectiveness of the manipulation depends on whether and when participants have to summarize or rewrite the presented messages.

Baseline condition. Another factor that may influence the manipulations' effectiveness is the baseline condition. Past research has shown that most people believe in free will (Baumeister et al., 2009; Nahmias et al., 2005). As a result, a tacit assumption in the literature is that beliefs in free will can only be diminished but not increased. Yet, some experiments have nevertheless used not only anti–free will and neutral messages but also pro–free will messages (e.g., Baumeister et al., 2009; Clark et al., 2017; Highhouse & Rada, 2015; Moynihan et al., 2019; Schrag et al., 2016; Seto & Hicks, 2016; Vohs & Schooler, 2008). This allowed us to test the effectiveness of anti–free will messages both with respect to control messages and with respect to pro–free will messages and thereby allowed us to investigate whether belief in free will can also be experimentally increased, in addition to being decreased.

Measurement moment. A final moderator that may influence the effectiveness of the manipulation is the moment at which free will beliefs are measured. While some experiments measured participants' beliefs directly after the manipulation (Baumeister et al., 2009; Monroe et al., 2017; Seto & Hicks, 2016; Shariff et al., 2014; Vohs & Schooler, 2008; Vonasch et al., 2017), others measured them after participants performed a secondary task, often included to explore potential downstream consequences (Clark et al., 2017; Genschow et al., 2017a; Harms et al., 2017; Highhouse & Rada, 2015; Protzko et al., 2016; Rigoni et al., 2012; Schrag et al., 2016). Measuring beliefs at the end of the experiment assumes the manipulation has a relatively long-lasting effect. This may, however, not necessarily be the case. For instance, it could also be that free will belief manipulations do not fully change people's beliefs but rather prime an anti–free will mindset. Such priming effects may not be as long-lasting, and measuring free will beliefs at the end of the experiment may therefore make it more difficult to confirm that the manipulation worked. To test this hypothesis, we investigated if the effect of the manipulation differs depending on whether beliefs are measured immediately after the manipulation or at the end of the experiment (i.e., after completing another task).

Secondary moderator analyses. In addition to the above-mentioned moderators, we also considered the influence of participant age, sex, the continent in which the study was conducted (United States vs. Europe), test location (i.e., online vs. laboratory), and sample type (students vs. nonstudents) as secondary moderators.

Research Question 2: Do Free Will Belief Manipulations Have Downstream Consequences?

In the second part of the meta-analysis, we investigate whether experimentally reducing free will beliefs influences attitudes, behavior, and cognition. An interesting characteristic of the literature is that a large variety of dependent variables have been studied (for an overview, see Table 1). While this illustrates the breadth of the field, it also makes it difficult to quantify the evidence for downstream consequences, because for some dependent variables only one or a few experiments exist. In this respect, one solution could be to group the variables into broad categories such as "behavior" or "attitudes." However, this would involve pooling together studies with vastly different dependent variables under the same denominator and would therefore run the risk of making the meta-analysis uninterpretable, a problem that is well known as the "apples and oranges" critique of meta-analysis (Carpenter, 2020). To deal with this issue, we proceeded in three steps.

In the first step, we ran a *p*-curve analysis across all dependent variables. While the aim of estimating a population effect size makes a meta-analysis unsuited to evaluate diverse sets of dependent variables, this is not the case for *p*-curve. Rather than estimating a population effect size, *p*-curve investigates whether a set of statistically significant findings contains evidential value by testing whether the distribution of *p*-values is consistent with the existence of a true effect (Simonsohn et al., 2014). Importantly, if confirmed, this does not mean that all included studies show a true effect. Instead, it merely implies that at least one study does (Simonsohn et al., 2014). As such, *p*-curve can be applied to diverse findings as long as they form a meaningful whole (Simonsohn et al., 2015).

In a second step, we ran meta-analyses on internally coherent sets of dependent variables. Upon reviewing the literature, one clear set arose—namely, antisocial versus prosocial behavior (for an overview, see Table 1). Hence, we pooled together the studies in this set and subjected them to a meta-analysis testing whether manipulating belief in free will influences social behavior. However, pro- and antisocial behavior is still a relatively broad and unspecific dependent variable. Therefore, in a third and final step, we also ran meta-analyses on three specific dependent variables that have been used in at least five experiments: conformity, punishment, and cheating.

Method

Search Strategy

The literature search for published articles was initiated in July 2018 and includes studies published between January 2008—based on the publication date of the first study that included an experimental belief in free will manipulation (Vohs & Schooler, 2008)—and July 2019. To collect published studies, we entered the following search terms in Web of Science, PubMed, and PsycINFO: ("Free will" AND "belie*") OR ("Free will" AND "manipulat*") OR ("Free will" AND "experiment*") OR ("Free will" AND "group") OR ("Free will" AND "induc*") OR ("Free will" AND "reduc*") OR ("Free will" AND "threat*") OR ("Free will" AND "undermin*") OR ("Free will" AND "weak*") OR ("Determinis*" AND "belie*") OR ("Determinis*" AND "manipulat*") OR ("Determinis*" AND "experiment*") OR ("Determinis*" AND "group") OR ("Determinis*" AND "induc*") OR ("Determinis*" AND "encourag*") OR

Table 1. Overview of All Published Experiments Included in the *P*-Curve Analysis and/or the Meta-Analyses Assessing the Influence of Free Will Belief Manipulations on Downstream Consequences.

Experiment	DV	Included in antisocial meta-analysis	Included in *p*-curve analysis
Vohs and Schooler (2008); Exp. 1	Cheating	Yes	Yes
Vohs and Schooler (2008); Exp. 2	Cheating	Yes	Yes
Baumeister et al. (2009); Exp. 1	Helping	Yes	Yes
Baumeister et al. (2009); Exp. 3	Aggression	Yes	Yes
Stillman and Baumeister (2010); Exp. 1	Learning	No	Yes
Stillman and Baumeister (2010); Exp. 2	Learning	No	Yes
Rangel & Keller (2011); Exp. 6	In-group preference	No	Yes
Rigoni et al. (2011); Exp. 1	Preconscious motor preparation	No	Yes
Alquist et al. (2013); Exp. 2	Conformity	No	Yes
Alquist et al. (2013); Exp. 3	Conformity	No	Yes
Evans (2013); Exp. 1	Willingness to help	Yes	Yes
Rigoni et al. (2013); Exp. 1	Action monitoring	No	Yes
MacKenzie et al. (2014); Exp. 2	Gratitude	No	Yes
MacKenzie et al. (2014); Exp. 3	Gratitude	No	Yes
MacKenzie et al. (2014); Exp. 4	Gratitude	No	Yes
Shariff et al. (2014); Exp. 2	Punishment	No	Yes
Shariff et al. (2014); Exp. 3	Punishment	No	Yes
Zhao et al. (2014); Exp. 2	Prejudice	Yes	Yes
Zhao et al. (2014); Exp. 3	Prejudice	Yes	Yes
Open Science Collaboration (2015); Exp. 1	Cheating	Yes	No
Alquist et al. (2015); Exp. 1	Counterfactual thinking	No	Yes
Alquist et al. (2015); Exp. 2	Counterfactual thinking and intention perception	No	Yes
Plaks and Robinson (2015); Exp. 4	Moral judgments	No	Yes
Rigoni et al. (2015); Exp. 1	Error detection	No	Yes
Crescioni et al. (2016); Exp. 3	Meaningfulness of life	No	Yes
Crescioni et al. (2016); Exp. 4	Meaningfulness of life	No	Yes
Goodyear et al. (2016); Exp. 1	Assigned moral responsibility	No	Yes
Protzko et al. (2016); Exp. 1	Cooperation	Yes	Yes
Schrag et al. (2016); Exp. 1	Risk behavior	No	Yes
Seto and Hicks (2016); Exp. 1	Self-alienation	No	Yes
Seto and Hicks (2016); Exp. 2	Reported authenticity	No	Yes
Caspar & Vuillaume (2017); Exp. 1	Immoral behavior	Yes	Yes
Clark et al. (2017); Exp. 4	Punitive distress	No	Yes
Genschow et al. (2017a); Exp. 3a	Correspondence bias	No	Yes
Genschow et al. (2017a); Exp. 3b	Correspondence bias	No	Yes
Hannikainen et al. (2017); Exp. 4	Preference for autonomy vs. Welfare	No	Yes
Hannikainen et al. (2017); Exp. 5	Approval of paternalistic policies	No	Yes
Monroe et al. (2017); Exp. 1	Punishment, cheating, moral judgments	Yes	No
Monroe et al. (2017); Exp. 3	Punishment and moral judgments	No	No
Vonasch et al. (2017); Exp. 6	Addiction and self-control	No	Yes
Ching & Xu (2018); Exp. 1	Prejudice	Yes	Yes
Moynihan et al. (2019); Exp. 1	Meaninglessness	No	Yes
Nadelhoffer et al. (2020); Exp. 3	Cheating	Yes	No
Nadelhoffer et al. (2020); Exp. 4	Cheating	Yes	No
Genschow & Vehlow (2019); Exp. 1	Victim blaming	Yes	Yes

Note. Nadelhoffer er al. (2019) and Genschow and Vehlow (2019) were included in the meta-analysis as unpublished studies but were published later as Nadelhoffer et al. (2020) and Genschow and Vehlow (2021) after data analysis had already been completed. Experiments were only included in the *p*-curve analysis if the effect of the manipulation on the DV was significant. All experiments that were not included in the *p*-curve analysis (included: "no") are experiments without a significant effect on the DV. Articles were included in the meta-analyses on antisocial behavior, conformity, punishment, and cheating irrespective of whether the effect was significant or not. DV= dependent variable.

("Determinis*" AND "increas*") OR ("Determinis*" AND "enhanc*").

In addition to this Boolean search, we also looked for studies that cited the Vohs and Schooler (2008) paper. Furthermore, we included studies that were cited in the recent review paper on belief in free will by Ewusi-Boisvert and Racine (2018). Third, unpublished data were collected by sending requests to authors who had previously published articles using free will belief manipulations. Fourth, we sent around requests for unpublished data via different national and international mailing lists, including the mailing list of the German Psychology Association (DGPs), the mailing list of the Belgian Association for Psychological Science (BAPS), and the mailing list of the European Society for Cognitive and Affective Neuroscience (ESCAN). Finally, we asked for unpublished data via Twitter and different open fora of the Society of Personality and Social Psychology (SPSP). The search for unpublished data was terminated on August 15, 2019.

Screening Process

We screened titles and abstracts from 3,739 records obtained from the literature search. In addition, 110 unpublished experiments were included. Studies that were clearly not eligible based on the criteria described below were excluded (i.e., n = 3570). The remaining studies were then evaluated in more detail by screening the full-text articles. All together, 279 full-text articles were assessed for eligibility. This procedure resulted in a database of 84 eligible studies containing 145 experiments (50 published and 95 unpublished). The number of excluded and included studies can be seen in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart in Figure 1. The inclusion criteria were:

1. The studied population had to involve healthy adults.
2. The study had to include an experimental manipulation with the aim of reducing or increasing belief in free will or belief in determinism.
3. For the research question investigating the influence of free will belief manipulations on free will beliefs, but not for the research question investigating downstream consequences, the study had to contain a manipulation check administered after the manipulation—that is, a measure of belief in free will, determinism, or both. We focused on belief in free will and belief in determinism because these are the two most commonly measured beliefs. Importantly, we only included measures that considered free will and determinism as separate constructs and not measures that considered them as two opposite endpoints of the same scale because the latter approach does not allow us to distinguish between both beliefs.[1]
4. Sufficient statistical information had to be available to calculate the necessary effect sizes. The required information had to be either reported in the paper or obtained from the authors by email.

Coding and Reliability

Research Question 1: Can belief in free will be manipulated? The eligible studies were first coded by the first author. To evaluate the reliability of the coding, the third author was trained to use the coding manual and subsequently recoded 17 randomly selected published studies containing 27 experiments in total. To further identify errors, the initial coding was also checked by the third author. Disagreements and inconsistencies that arose during coding were resolved by the first three authors via discussion. Coding reliability was assessed with intraclass correlation coefficients (ICC) for continuous variables and with kappa coefficients for categorical variables. The reliability for continuous variables was high (all ICC $\geq$ .99). Similarly, the average reliability for categorical variables was high (κ = .96) and varied from .74 to 1.00. The coding manual is openly accessible at the Open Science Framework (OSF; https://bit.ly/2L69prl). The following variables were coded:

Publication status. We coded whether an experiment was published (m = 51) or unpublished (m = 95). An experiment was considered published when it appeared in an academic journal or book. Dissertations, preprints (if not published elsewhere), conference posters, and raw data were all considered unpublished. Experiments that were initially unpublished but then later published before the meta-analysis was completed remained coded as unpublished experiments to account for potential differences between the original and published results (e.g., additional data collection, different exclusion criteria, and different analytical choices).

Beliefs. We coded whether belief in free will or belief in determinism was measured. Effect sizes obtained with instruments measuring belief in free will and belief in determinism as opposite ends of a scale (e.g., the FWD) were not included. Effect sizes were coded as measuring belief in free will when they were obtained with the free will subscales of the FWI, the FAD-Plus, or the FAD. In addition, we also included self-made scales and individual items designed to measure belief or disbelief in free will. The disbelief scales were reverse coded.

Effect sizes were coded as measuring belief in determinism when they were obtained with the following scales: belief in determinism of the FWI (Nadelhoffer et al., 2014), biological determinism of the WiF (Melcher, 2019), genetic determinism of the BGD (Keller, 2005), social determinism of the BSD (Rangel & Keller, 2011), and scientific determinism of

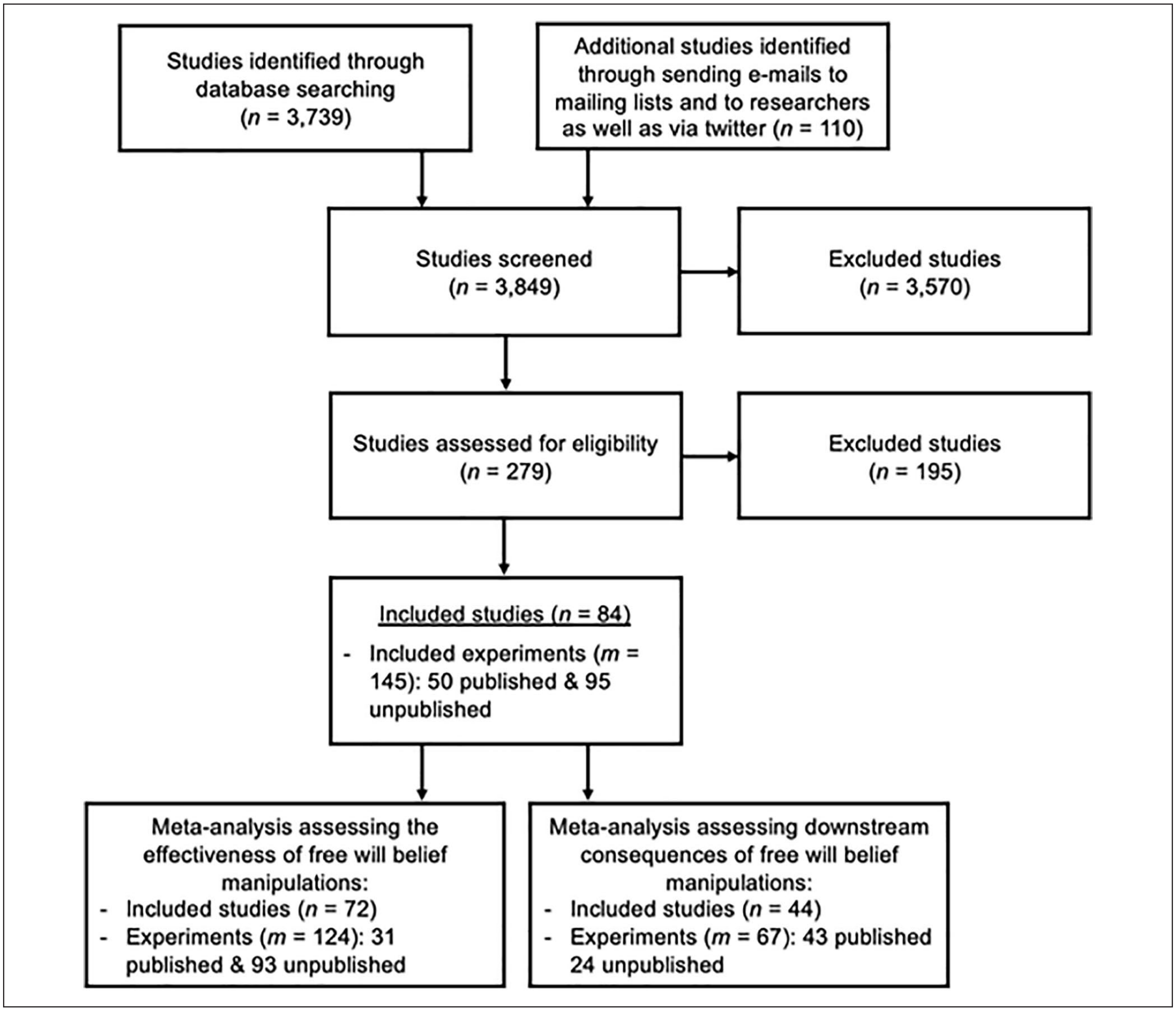

Figure 1. Flowchart of the screening process.
Note. The number of included or excluded studies (*n*) and experiments (*m*) are denoted. See text for additional detail.

the FAD-Plus or the FAD scale (Paulhus & Carey, 2011). In addition, as for free will beliefs, we also included self-made scales and individual items designed to measure either belief or disbelief in determinism. The disbelief items were reverse coded. We did not include the FAD/FAD-Plus' fatalistic determinism subscale because it does not capture determinism as construed by the other scales (i.e., the belief that future events are completely determined by prior events) but rather measures belief in "destiny" or "fate."

Scales. To compare the sensitivity of different scales, we coded which of the following scales was used to measure belief in free will and/or determinism: the FAD/FAD-Plus (Paulhus & Carey, 2011), the WiF (Melcher, 2019), the BGD (Keller, 2005), the BSD (Rangel & Keller, 2011), or self-made scales. The coding indicated that not enough experiments (i.e., $m < 3$) had used the WiF, BGD, and BSD to draw reliable conclusions. Furthermore, the self-made scales did not form an internally coherent cluster. Therefore, we restricted the scales analysis to the FWI and FAD/FAD-Plus.

For the FWI and FAD/FAD-Plus, we also coded not only the free will and determinism subscales but also all other subscales to investigate whether the influence of free will belief manipulations is specific to belief in free will and belief in determinism or also extends to other, related beliefs. More specifically, we coded the dualism subscale of the FWI and the fatalistic determinism and unpredictability subscales of the FAD/FAD-Plus.

Type of manipulation. To compare different types of manipulations, we coded whether a text, statements, a combination of text and statements, or a video was used to manipulate belief in free will.

Participant involvement. To test whether the effect of the manipulation depends on how deeply the participant had to process the stimulus material, we coded participant involvement. For some studies, participants were asked to attend the provided information carefully so that they could summarize it at the end of the experiment. In other studies, participants had to process the provided information directly after the manipulation by writing an essay or rewriting the provided messages. Yet in other studies, participants merely attended the provided information and did not have to reproduce any information. For participant involvement, we coded whether participants (a) had to summarize or rewrite the provided information directly after the manipulation, (b) had to summarize or rewrite the information at the end of the experiment, or (c) simply had to attend the provided information, such as by merely reading a text or watching a video.

Baseline condition. To test whether beliefs in free will and determinism can also be increased, in addition to decreased, we coded whether the baseline condition consisted of neutral messages or pro–free will messages.

Measurement moment. To investigate how long-lasting the free will belief manipulations are, we coded whether belief in free will or determinism was assessed directly after the manipulation or after the primary dependent variable was assessed.

Secondary moderators. We also coded several secondary moderators, namely, age, sex, continent, test location, and sample type. Age was coded as the mean age of participants in the sample. Sex was coded as the proportion of female participants. Continent refers to whether the sample was collected in the United States or on U.S.-based platforms such as MTurk versus in Europe or on Europe-based platforms such as Prolific (not enough experiments were conducted on other continents to draw reliable conclusions). Test location refers to whether the experiment was conducted in the laboratory or online. Finally, sample type refers to whether the participants were students or panel participants (e.g., MTurk or Prolific).

Research Question 2: Do free will belief manipulations have downstream consequences? In a first step, the first author coded all studies measuring downstream consequences of manipulating belief in free will. The following variables were coded: the dependent variable, whether the study included a manipulation check, whether the manipulation check was significant, and whether the experiment was published. A manipulation check was considered significant when at least one of the measured beliefs (i.e., free will or determinism) reached a p-value of $p < .05$. In a second step, the second author checked the first author's coding for errors. Disagreements and inconsistencies that arose during coding were resolved through discussion between both authors. The pooling of dependent variables into the clusters described below was done through mutual discussion.

P-curve. Rather than estimating a population effect size, p-curve investigates whether there is at least one effect in a set of significant effects that is not zero, by testing whether the p-values corresponding to these effects follow a flat or skewed distribution (Simonsohn et al., 2014). The p-curve analysis assessing the overall evidence for downstream consequences of free will manipulations included 35 studies with a total of 49 experiments (see OSF; https://bit.ly/2L69prl). In all, 39 of these experiments were published in academic journals or books. The remaining experiments were unpublished datasets. The p-values used in the p-curve analysis were coded according to the guidelines laid out in Simonsohn et al. (2014). If a study reported multiple relevant tests, we coded only the first test. We also report robustness analyses checking (a) whether the results changed when we used the second instead of the first test and (b) whether significant effects still remained after removing the most significant test (Simonsohn et al., 2015). Note that for this latter robustness analysis, we determined the p-value using a simulated null distribution to account for the fact that removing the lowest p-value affects the distribution of the test statistic.

Antisocial behavior meta-analysis. Upon reviewing the literature on downstream consequences, one clear set of variables arose—namely, antisocial versus prosocial behavior. Hence, we pooled together the 15 studies ($m = 21$) in this set and subjected them to a meta-analysis testing whether manipulating belief in free will influences social behavior. The prosocial dependent variables were helping and positive attitudes toward minorities. The antisocial dependent variables were immoral behavior, cheating, prejudice, malicious envy, and victim-blaming. Studies measuring prosocial behavior or prosocial attitudes were reverse coded.

Conformity, punishment, and cheating meta-analyses. Three dependent variables were measured in at least five experiments: cheating ($n = 5$; $m = 8$), conformity ($n = 4$, $m = 6$), and punishment ($n = 3$; $m = 9$).

Meta-Analytic Procedures

Robust variance estimation. Analyses were performed in R (v3.5.1) using the metafor (Viechtbauer, 2010) and robumeta (Fisher & Tipton, 2015) packages. The data were analyzed using random-effects models because we considered heterogeneity to be likely and because random-effects models converge on fixed-effects models in the absence of heterogeneity

(Field & Gillett, 2010). A critical assumption of such random-effects models is that the included effect sizes are statistically independent (Lipsey & Wilson, 2001), and violating this assumption is known to inflate the false-positive rate (Tanner-Smith et al., 2016). In the current meta-analysis, however, many studies measured multiple dependent variables or included multiple baseline conditions and therefore yielded more than one relevant effect size. As a result, the included effect sizes were not independent. To control for this dependency, we decided to empirically estimate standard errors that do not assume independent effect sizes using robust variance estimation (RVE; Fisher & Tipton, 2015; Hedges et al., 2010; Tanner-Smith et al., 2016; Tanner-Smith & Tipton, 2014).

In meta-analyses, effect sizes are typically weighted by their standard error. Within the RVE framework, two ways to calculate such weights have been proposed: hierarchical effects weights and correlated effects weights. The former is most appropriate when dependency originates mainly from effect sizes being nested in larger units (e.g., research groups), whereas the latter is more appropriate when dependency originates mostly from a single study providing multiple effect sizes. While RVE provides asymptotically unbiased standard errors regardless of how the weights are calculated, the choice of weights does influence statistical efficiency (Hedges et al., 2010). As we expected dependency to arise mainly from studies providing multiple effect sizes, we used correlated effects weights (Fisher & Tipton, 2015). These weights are inverse variance weights, where the denominator of the weight assigned to each effect size in an experiment is determined by the average variance in the experiment multiplied by the number of effect sizes provided by that experiment. Importantly, this ensures that the total weight of each experiment does not depend on the number of effect sizes it provides (e.g., Cracco et al., 2018). Of note, to estimate variability, a parameter ρ representing the correlation among the effect sizes has to be specified (Fisher & Tipton, 2015). This parameter is assumed to be the same for all experiments and typically has a negligible influence on the results. In the current meta-analysis, we used the default value of the robumeta package (i.e., $\rho = 0.80$) but also report sensitivity analyses where we vary this value.

Importantly, one problem with RVE is that it has an inflated false-positive rate when the number of studies is moderate to small or when skewed or unbalanced moderators are included (Tipton, 2015; Tipton & Pustejovsky, 2015). As a solution, small-sample (i.e., a limited number of studies in the meta-analysis) corrections have been proposed for both t tests (Tipton, 2015) and F tests (Tipton & Pustejovsky, 2015). As it is difficult to know when these corrections should be implemented and when they should not, it has been recommended to implement them for all RVE analyses, regardless of the meta-analytical sample size (Tipton, 2015). In the current meta-analysis, we follow this recommendation. Importantly, corrections to the t test are only valid when df is ≥ 4 (Tipton, 2015). As a result, we only report t tests with more than 4 df. In addition, for continuous moderators (e.g., age), we removed outlier values exceeding the weighted mean by more than 3 *SDs* because such outliers strongly reduce the available degrees of freedom and hence statistical power (Tanner-Smith et al., 2016). Weighted means were calculated by dividing a weight of 1 equally among the different values provided by the same experiment.

Effect size estimation. All included studies used a between-subjects design to test the effect of the manipulation. Hedge's g was used as a measure of effect size. Effects were coded so that positive effect sizes corresponded to stronger beliefs in free will/determinism (research question 1) and larger values on the outcome measures (research question 2) in the anti–free will condition than in the reference condition. Hedge's g was calculated by first calculating Cohen's d and then correcting these values using the *escalc* function of the metafor package in R (Viechtbauer, 2010). When means and standard deviations were reported, Cohen's d was calculated as follows:

$$d_s = \frac{M_{anti} - M_{ref}}{\sqrt{\frac{(n_{anti}-1)SD_{anti}^2 + (n_{ref}-1)SD_{ref}^2}{n_{anti}+n_{ref}-2}}}. \quad (1)$$

where *anti* refers to the anti–free will condition and *ref* to the reference (i.e., baseline) condition. When standard errors were provided instead of standard deviations, these were transformed to standard deviations and Cohen's d was calculated using Equation 1. When insufficient information was reported to use Equation 1, we instead calculated Cohen's d from the t- or F-value as follows:

$$d_s = t\sqrt{\frac{1}{n_{anti}} + \frac{1}{n_{ref}}} \text{ or } \sqrt{F}\sqrt{\frac{1}{n_{anti}} + \frac{1}{n_{ref}}}. \quad (2)$$

When we could not retrieve sufficient information to calculate an effect size, we contacted the corresponding author for the necessary information. Nevertheless, despite our efforts, we could not calculate the effect size for all coded studies. Specifically, for the first research question, we could not calculate any of the effect sizes for four experiments from three studies and only part of the effect sizes for three experiments from three studies. For the second research question, we were not able to code five experiments from five studies. Effects for which we could not calculate an effect size were not included in the meta-analysis.

Outliers. Outliers were defined as effect sizes exceeding the weighted mean effect size, calculated as before, by more than 3 *SD*s. For the first research question, this procedure identified one outlier for free will beliefs ($g = -2.89$) and one outlier for determinism beliefs ($g = 4.00$). These outliers were replaced by the effect size 3 *SD*s above the weighted mean effect size for free will beliefs ($g = -1.63$) and determinism beliefs ($g = 2.09$). There were no outliers for the second research question. In addition to replacing outliers, we also tested the influence of each individual effect size on the average effect size using a leave-one-out cross-validation procedure. This indicated that leaving out individual effect sizes did not strongly influence the average effect size and mostly influenced it similarly in both directions for both the first (free will beliefs: $\Delta g_{min} = -0.01$, $\Delta g_{max} = 0.01$; determinism beliefs: $\Delta g_{min} = -0.02$, $\Delta g_{max} = 0.01$) and second research questions (antisocial behavior: $\Delta g_{min} = -0.03$, $\Delta g_{max} = 0.04$; cheating: $\Delta g_{min} = -0.11$, $\Delta g_{max} = 0.12$; conformity: $\Delta g_{min} = -0.04$, $\Delta g_{max} = 0.11$; punishment: $\Delta g_{min} = -0.05$, $\Delta g_{max} = 0.05$).

Moderator correlations. To control for confounded moderators (Field & Gillett, 2010; Lipsey & Wilson, 2001), we computed the weighted associations, calculated as before, between the different moderators included in the analysis of the first research question with r for continuous–continuous and continuous–dichotomous pairs, with multiple R for continuous–polytomous pairs, and with Cramér's *V* for dichotomous–dichotomous, dichotomous–polytomous, and polytomous–polytomous pairs (Cracco et al., 2018). For all these measures, 0 means *no relationship* and 1 means *a perfect relationship*. When two moderators correlated $>.50$, we tested whether the moderator effects (if any) remained after controlling for the confounded moderator. When the contingency table of two confounded moderators did not contain empty cells, we controlled for their confounding influences by including both moderators in the same meta-regression model. Using sum coding, this allowed us to test the average effect of each moderator across the levels of the other moderator (for categorical moderators). When the contingency table did contain empty cells, we controlled for confounding by restricting the analysis of Moderator A to the level of Moderator B where the levels of Moderator A were most balanced to maximize statistical power (Tipton, 2015; Tipton & Pustejovsky, 2015).

Publication bias. For both research questions, we tested for publication bias and other small-study effects using a hierarchical Egger's regression test (Sterne & Egger, 2005). More specifically, we ran an RVE meta-regression predicting effect sizes from their standard error (Rodgers & Pustejovsky, 2020). Publication bias leads to a positive relationship between effect sizes and standard errors because only large effect sizes are statistically significant in studies with a large standard error (small N). To correct for potential publication bias, we used the PET-PEESE (precision-effect test and precision-effect estimate with standard errors) approach (Stanley & Doucouliagos, 2014) because unlike other popular approaches, such as three-parameter selection models (3PSM), it is easily implemented within the RVE framework (Rodgers & Pustejovsky, 2020). Moreover, simulation studies have shown that PET-PEESE performs reasonably well under most circumstances (Carter et al., 2019). In PET-PEESE, a bias-corrected effect is calculated by taking the intercept of two RVE meta-regressions that, respectively, regress the effect sizes onto their standard error (PET) and variance (PEESE). PET and PEESE have complementary strengths and weaknesses: Whereas PET tends to be overly conservative when there is a true effect, PEESE tends to be overly liberal when there is no true effect (Stanley & Doucouliagos, 2014). Therefore, we report both measures, as has recently been recommended (Carter et al., 2019). Importantly, in some cases, PET and PEESE overcorrect effect sizes, reversing their sign. Because these overcorrections are not meaningful (Carter et al., 2019), we report the corrected effect size as 0 when this occurs. Finally, in addition to using bias correction methods, we also compared published with unpublished studies and tested if the effects remained significant when only unpublished studies were considered.

Sensitivity analyses. To explore the robustness of our main results, we conducted four sensitivity analyses investigating how variations to our analysis procedure influenced (a) the effect size of the manipulation and (b) the estimated publication bias. First, we repeated the analyses while varying the ρ parameter of the RVE models from 0 to 1 in steps of 0.2. Second, we repeated the analyses after computing effect sizes preferentially from the reported statistics instead of from the means and standard deviations, as sometimes studies reported both. Although these two types of information should in principle lead to the same effect size, this is not always the case. Therefore, we decided to test how changing the type of data used to calculate the effect sizes influenced the results. Third, we repeated the analyses without replacing outlier values. Finally, we used a multilevel meta-analytical approach (Cheung, 2014) instead of RVE to test whether different methods of dealing with dependent effect sizes yielded similar effect sizes. A sensitivity analysis was considered to produce results identical to the main analysis as long as the difference between the corresponding estimates was ≤ 0.01.

Results

Research Question 1: Can the Belief in Free Will Be Manipulated?

Effect on belief in free will and belief in determinism. We first tested the overall effectiveness of the manipulation for both free will and determinism beliefs. This indicated that exposing participants to anti–free will manipulations decreased belief in free will, $g = -0.29$, $t(111) = -8.74$, $p < .001$, 95% confidence interval [CI] = [−0.35, −0.22], $m = 119$, $k = 148$, and increased belief in determinism, $g = 0.17$, $t(46.5) = 4.33$,

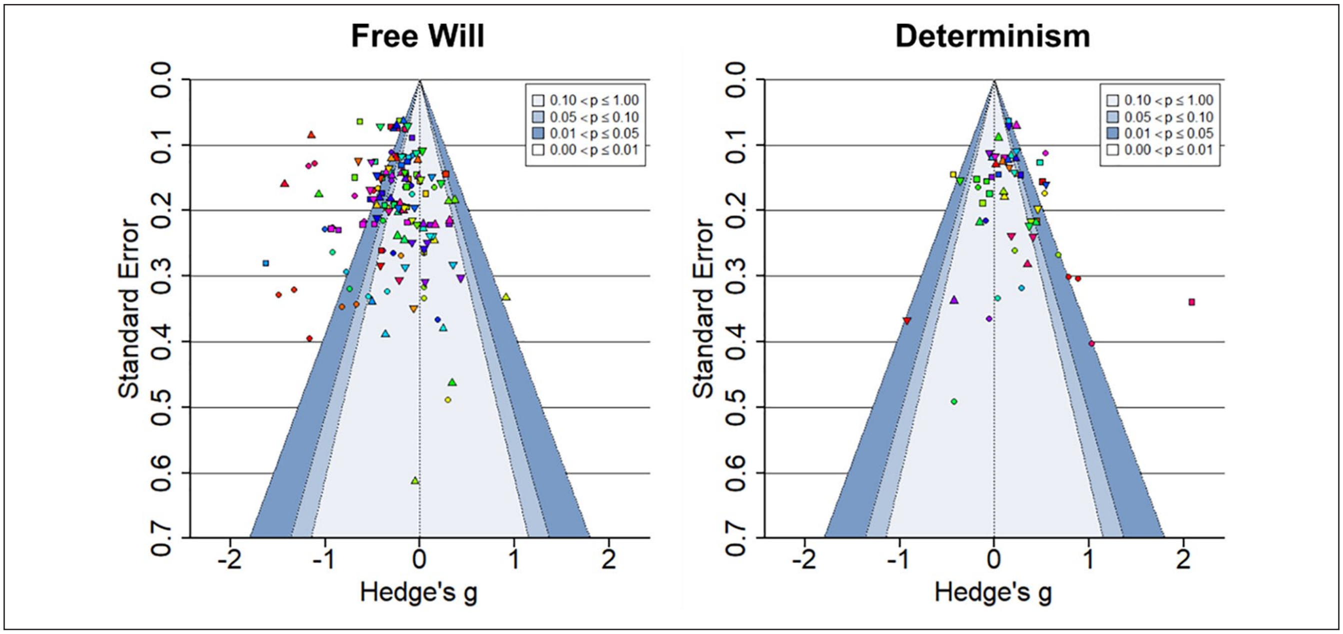

Figure 2. Funnel plots for belief in free will (left) and belief in determinism (right).
Note. Effect sizes from the same study are plotted in the same color with the same symbol.

$p < .001$, 95% CI = [0.09, 0.24], $m = 53$, $k = 61$. Next, we compared the influence of the manipulation on these two beliefs by reverse-coding the determinism belief effect sizes and comparing them with the free will belief effect sizes. This revealed a borderline nonsignificant difference, with the manipulation having a slightly stronger effect on free will beliefs than on determinism beliefs, $t(66) = -1.87$, $p = .066$, $m = 124$, $k = 209$. Finally, we looked at the I^2 coefficients, which indicated that a substantial portion of the variance for both belief in free will ($I^2 = 79\%$) and belief in determinism ($I^2 = 67\%$) was due to heterogeneity in the effect sizes.[2] In sum, these analyses indicate that anti–free will manipulations reduce the belief in free will and increase the belief in determinism.

Publication bias and small-study effects. A visual inspection of the funnel plots (see Figure 2) revealed a largely symmetrical effect size distribution for both free will and determinism beliefs. In line with this visual inspection, a hierarchical Egger test found no relationship between the standard error and the effect size for either free will, $t(41.3) = 0.58$, $p = .565$, or determinism beliefs, $t(19.4) = 0.76$, $p = .457$. Importantly, however, this does not necessarily mean that there is no publication bias in the literature. Indeed, our analysis included a large number of unpublished studies both for the free will ($m = 93$, $k = 116$) and for the determinism analysis ($m = 39$, $k = 44$), potentially masking the presence of publication bias. In support of this hypothesis, an analysis including publication status as a moderator showed that the effect of the manipulation was larger in published than in unpublished studies for both belief in free will, $t(36.8) = -3.43$, $p = .002$ (see Table 2), and belief in determinism, $t(18.9) = 2.14$, $p = .046$ (see Table 3). Importantly, however, the effect size of both free will beliefs, $t(86.4) = -6.58$, $p < .001$, and determinism beliefs, $t(34.4) = 2.75$, $p = .010$, remained significant even when only unpublished studies were included in the analysis. Similarly, correcting the free will belief effect size with PET ($g_z = -0.33$, $p < .001$) or PEESE ($g_z = -0.31$, $p < .001$) revealed that the effect size was still significant. Correcting the determinism belief effect size with PET made the effect disappear ($g_z = 0.06$, $p = .617$), but correcting it with PEESE did reveal a significant effect ($g_z = 0.13$, $p = .040$).

Taken together, we find evidence for publication bias, but also find that publication bias is unlikely to explain the effect of the manipulation on free will beliefs. Although the effect on determinism beliefs disappeared after applying PET correction, it was still significant after applying PEESE correction, and an analysis including only unpublished studies likewise revealed a significant determinism effect. Given that PET is known to be overly conservative when there is a true effect (Stanley & Doucouliagos, 2014), the combined evidence suggests that not just belief in free will but also belief in determinism is affected by the manipulation, even after accounting for publication bias.

Primary moderators

Scales. We first investigated whether the effectiveness of the manipulation depends on whether free will and determinism were measured using the FWI scale or using the FAD scale. This revealed no significant difference between both scales for free-will beliefs, $t(74.9) = 1.53$, $p = .130$ (see Table 2), but a stronger effect when belief in determinism was

Table 2. Free Will Moderator Statistics.

Moderator	g/β	*SE*	95% CI	*m*	*k*	*F/t* (*df*)	*p*
Publication Status						3.43 (36.8)	.002
Unpublished	-0.23_a	0.04	[–0.30, –0.16]	93	116		
Published	-0.51_b	0.07	[–0.66, –0.36]	26	32		
Questionnaire						1.53 (74.9)	.130
FWI	-0.23_a	0.04	[–0.31, –0.14]	39	40		
FAD	-0.13_a	0.04	[–0.22, –0.05]	46	59		
Manipulation Type						6.70 (3, 26.3)	.002
Text	-0.18_a	0.03	[–0.24, –0.12]	58	63		
Statements	-0.29_a	0.06	[–0.41, –0.17]	35	54		
Text and Statements	-0.79_b	0.14	[–1.12, –0.45]	8	9		
Video	-0.41_a	0.11	[–0.64, –0.17]	12	12		
Text Type						0.71 (13.2)	.493
Crick	-0.17_a	0.03	[–0.23, –0.11]	48	51		
Other	-0.22_a	0.06	[–0.35, –0.08]	10	12		
Involvement						10.90 (2, 69.5)	<.001
None	-0.22_a	0.06	[–0.34, –0.10]	33	43		
Report Before	-0.46_b	0.06	[–0.59, –0.34]	44	53		
Report After	-0.14_a	0.03	[–0.20, –0.09]	32	39		
Measurement Moment						2.32 (92.3)	.022
Before	-0.38_a	0.07	[–0.52, –0.25]	51	68		
After	-0.21_b	0.03	[–0.27, –0.15]	49	58		
Baseline						4.11 (106.7)	<.001
Control	-0.18_a	0.03	[–0.23, –0.12]	72	80		
Pro–free will	-0.42_b	0.06	[–0.53, –0.31]	64	67		
Age	0.04	0.14	[–0.25, 0.32]	64	80	0.27 (28.2)	.793
Sex (% Female)	0.09	0.16	[–0.24, 0.42]	66	82	0.53 (31.5)	.598
Continent						2.50 (101.0)	.014
Europe	-0.19_a	0.05	[–0.29, –0.09]	54	62		
United States	-0.36_b	0.05	[–0.45, –0.27]	62	83		
Test Location						0.45 (77.2)	0.655
Lab	-0.29_a	0.05	[–0.40, –0.18]	42	49		
Online	-0.26_a	0.05	[–0.36, –0.15]	57	69		
Sample						0.81 (62.6)	0.422
Students	-0.30_a	0.05	[–0.39, –0.21]	50	63		
Panel	-0.37_a	0.08	[–0.53, –0.21]	29	32		

Note. Different subscripts indicate $p < .05$. All statistical tests are unsigned. g/β = Hedge's *g*/beta coefficient; 95% CI = 95% confidence interval; *m* = number of experiments; *k* = number of effect sizes; FWI = Free Will Inventory; FAD = Free Will and Determinism scale.

measured with the FAD than with the FWI, $t(21.8) = 2.29$, $p = .031$ (see Table 3). However, scale correlated highly with measurement moment for determinism beliefs (see Table 4). In particular, studies that used the FWI tended to measure beliefs at the end of the experiment rather than right after the manipulation, whereas studies using the FAD were more balanced with respect to measurement moment. To control for this confound, we fitted a model including both scale and measurement moment. This revealed that the difference between the FWI and FAD for determinism beliefs was no longer significant, $t(15.8) = 1.43$, $p = .172$, when measurement moment was controlled, although the numerical pattern went in the same direction as before.

Next, we investigated whether free will manipulations also influenced the other subscales of the FWI and FAD. The FWI analysis indicated that the manipulation influenced not only belief in free will, $t(25.9) = -3.73$, $p < .001$, and belief in determinism, $t(26.9) = 2.27$, $p = .031$, but also belief in dualism, $t(25.8) = -6.56$, $p < .001$. More specifically, it indicated that belief in free will and belief in dualism decreased, whereas belief in determinism increased following anti–free will manipulations. A direct comparison of the size of these three effects revealed a significant main effect of subscale, $F(2, 31.6) = 4.43$, $p = .020$, with significantly larger effect sizes for dualism than for determinism, $t(27.7) = 2.87$, $p = .008$, but no significant difference between free

Table 3. Determinism Moderator Statistics.

Moderator	g/β	*SE*	95% CI	*m*	*k*	*F/t* (df)	*p*
Publication Status						2.14 (18.9)	.046
Unpublished	0.12_a	0.05	[0.03, 0.21]	39	44		
Published	0.30_b	0.07	[0.15, 0.44]	14	17		
Questionnaire						2.29 (21.8)	.032
FWI	0.11_a	0.04	[0.02, 0.19]	30	31		
FAD	0.30_b	0.07	[0.14, 0.47]	16	18		
Manipulation Type						1.83 (2, 15.7)	.193
Text	0.09_a	0.04	[0.02, 0.17]	30	32		
Statements	0.26_a	0.08	[0.08, 0.43]	15	20		
Video	0.26_a	0.15	[–0.12, 0.64]	7	7		
Involvement						0.61 (2, 29.3)	.549
None	0.23_a	0.07	[0.08, 0.37]	16	19		
Report Before	0.18_a	0.11	[–0.05, 0.41]	14	17		
Report After	0.14_a	0.04	[0.06, 0.22]	21	22		
Measurement Moment						0.95 (33.2)	.349
Before	0.22_a	0.08	[0.05, 0.39]	18	22		
After	0.13_a	0.05	[0.03, 0.23]	28	29		
Baseline						1.05 (25.0)	.304
Control	0.13_a	0.03	[0.06, 0.20]	40	43		
Pro–free will	0.25_a	0.10	[0.02, 0.47]	18	18		
Age	0.03	0.13	[–0.25, 0.30]	42	49	0.21 (19.1)	.838
Sex (% Female)	–0.03	0.19	[–0.43, 0.38]	45	52	0.14 (17.3)	.893
Continent						0.99 (45.2)	.329
Europe	0.13_a	0.06	[0.01, 0.25]	30	35		
United States	0.21_a	0.05	[0.10, 0.31]	23	26		
Test Location						0.55 (42.6)	.586
Lab	0.19_a	0.06	[0.07, 0.32]	27	33		
Online	0.15_a	0.05	[0.04, 0.26]	22	23		
Sample						0.62 (28.7)	.543
Students	0.19_a	0.06	[0.07, 0.31]	31	37		
Panel	0.15_a	0.05	[0.03, 0.26]	14	15		

Note. Different subscripts indicate $p < .05$. All statistical tests are unsigned. g/β = Hedge's g/beta coefficient; 95% CI = 95% confidence interval; m = number of experiments; k = number of effect sizes; FWI = Free Will Inventory; FAD = Free Will and Determinism scale.

Table 4. Weighted Correlations Among Moderator Variables.

Moderator	Published	Scale	Type	Involvement	Moment	Baseline	Age	Sex	Continent	Location	Sample
Published	1.00	.43	.32	.25	.06	.00	.16	.16	.25	.08	.03
Scale	.28*	1.00	.40	.31	**.52***	.15	.27	.38	.29	.28	.26
Type	.20	.36***	1.00	**.59*****	**.55******	**.75*****	.23	.27	**.51*****	.10	.27
Involvement	.19	.37***	.42***	1.00	.39*	**.69*****	.14	.25	.50**	.21	.36*
Moment	.08	.42***	.42***	.43***	1.00	.01	.11	.04	.38**	.06	.05
Baseline	.13	.25	**.66*****	**.59*****	.18	1.00	.05	.20	.24	.07	.09
Age	.04	.30	.34*	.10	.17	.01	1.00	.42**	.25	**.84*****	**.75*****
Sex	.09	.18	.39**	.23	.10	.08	**.55*****	1.00	.36**	.43**	**.66*****
Continent	.26**	.21	.29*	.26*	.02	.11	.24*	.43***	1.00	.42**	**.69*****
Location	.02	.19	.18	.18	.08	.07	**.82*****	.44***	.42***	1.00	**.86*****
Sample	.15	.23	.33*	.45***	.05	.13	**.81*****	**.63*****	.50***	**.72*****	1.00

Note. The lower half of the matrix shows correlations for free will beliefs and the upper half of the matrix shows correlations for determinism beliefs. Relations between continuous–continuous variables and between continuous–dichotomous variables were assessed with Pearson's *r*. Relations between continuous–polytomous variables were assessed with multiple R. Relations between dichotomous–dichotomous, dichotomous–polytomous, and polytomous–polytomous variables were assessed with Cramér's *V*. Correlations > .50 are indicated in bold font. These correlations were controlled for whenever one of the two involved moderators produced a significant effect.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 5. FWI and FAD Statistics.

Moderator	g/β	*SE*	95% CI	*m*	*k*	*F/t* (*df*)	*p*
FWI						4.43 (2, 31.6)	.020
Free Will	$0.25_{a,b}$	0.07	[0.11, 0.38]	39	40		
Determinism	0.11_{a}	0.05	[0.01, 0.21]	30	31		
Dualism	0.29_{b}	0.05	[0.20, 0.39]	28	29		
FAD						6.52 (3, 9.8)	.011
Free Will	0.12_{a}	0.04	[0.04, 0.21]	46	59		
Scientific Determinism	0.40_{b}	0.08	[0.22, 0.57]	16	18		
Fatalistic Determinism	0.13_{a}	0.04	[0.02, 0.23]	13	14		
Unpredictability	-0.03_{c}	0.05	[−0.13, 0.07]	12	13		

Note. Different subscripts indicate $p < .05$. Effect sizes are coded such that larger effects in the expected direction lead to more positive effect sizes. All statistical tests are unsigned. g/β = Hedge's *g*/beta coefficient; 95% CI = 95% confidence interval; *m* = number of experiments; *k* = number of effect sizes; FWI = Free Will Inventory; FAD = Free Will and Determinism scale.

will and determinism, $t(34.4) = 1.86$, $p = .072$, or between free will and dualism, $t(33.9) = 0.57$, $p = .571$ (see Table 5).

The FAD analysis likewise showed that the manipulation changed not only belief in free will, $t(33) = -2.88$, $p = .007$, and belief in scientific determinism, $t(7.8) = 5.26$, $p < .001$, but also belief in fatalistic determinism, $t(5.3) = 3.10$, $p = .025$. More specifically, belief in free will decreased and belief in scientific and fatalistic determinism increased following anti–free will manipulations. No effect was found on belief in unpredictability, $t(8.3) = 0.67$, $p = .522$. Comparing the effectiveness of the manipulation across subscales again revealed a significant main effect, $F(3, 9.8) = 6.51$, $p = .011$. Post hoc tests showed that the effect on scientific determinism was stronger than on all other subscales (all *p*s ≤ .007) and that the effect on unpredictability was significantly weaker than the effect on the other subscales (all *p*s ≤ .045; see Table 5).

In sum, the analysis comparing the different scales indicated that anti–free will manipulations have similar effects irrespective of whether belief in free will and belief in determinism are measured with the FWI or with the FAD. Interestingly, the results also showed that the effects of the manipulation are not specific to belief in free will and belief in scientific determinism, but also extend to belief in dualism and belief in fatalistic determinism, albeit not to belief in unpredictability.

Type of manipulation. We first compared the effectiveness of the different manipulations (i.e., text, statements, text and statements combined, or video). This revealed a significant effect of manipulation type on free will beliefs, $F(3, 26.3) = 6.70$, $p = .002$ (see Table 2), but not on determinism beliefs, $F(2, 15.7) = 1.83$, $p = .193$ (see Table 3).

A further analysis of the free will belief effect showed that manipulations combining a text with statements were more effective than manipulations only presenting statements, $t(11.4) = -3.26$, $p = .007$, only presenting a text, $t(9.5) = -4.17$, $p = .002$, or only showing a video, $t(14.6) = -2.14$, $p = .050$. None of the other conditions differed significantly from each other (all *p*s ≥ .057; see Table 2). Next, we also compared the effectiveness of the Crick text, which has been used most often in the literature, with the effectiveness of other texts that have been used. This revealed no difference for free will beliefs, $t(13.2) = -0.71$, $p = .493$ (see Table 2). The determinism analysis did not have enough degrees of freedom (i.e., $df = 2.15$) to interpret.

Importantly, however, manipulation type correlated strongly with baseline condition for free will beliefs (see Table 2). In particular, a closer look at the coding revealed that studies combining a text with statements and studies using videos always had a pro–free will baseline. Therefore, to test if this could explain the effect of manipulation type, we did a control analysis restricting the analysis to those studies using a pro–free will baseline. While this changed the main effect of manipulation type from significant to marginally significant, $F(3, 21.9) = 2.86$, $p = .060$, it did not substantially change the pattern of results. In particular, manipulations combining a text with statements were still significantly more effective than manipulations only using statements ($p = .016$), and manipulations only using a text ($p = .016$) were marginally more effective than manipulations using a video ($p = .061$).

Taken together, attempts to manipulate beliefs in free will are most effective when combining texts arguing against free will with statements that help strengthen the message. They are more effective than only using statements, texts, or videos.

Participant involvement. For participant involvement, there was a significant main effect on free will beliefs, $F(2, 69.5) = 10.9$, $p < .001$ (Table 2), but not on determinism beliefs, $F(2, 29.3) = 0.61$, $p = .549$ (see Table 3). Further analysis of the free will effect showed that the manipulation was more effective when participants had to report on the content of the manipulation immediately after the manipulation compared with when they had to report on the content at the end of the experiment, $t(63.8) = -4.67$, $p < .001$, or not at all, $t(61.9) = -2.79$, $p = .007$. The latter two conditions, however, did not differ, $t(58.5) = 1.19$, $p = .240$. Importantly, like manipulation type, participant involvement also correlated highly with baseline condition for free will beliefs. Visual inspection showed that most studies requiring participants to report on the content of the manipulation immediately after the manipulation also used a pro–free will baseline, whereas other studies tended to use a neutral baseline. To control for this confound, we fitted a model including both participant involvement and baseline condition, which revealed that participant involvement remained significant, $F(2, 50) = 4.71$, $p = .013$.

In conclusion, anti–free will manipulations have the strongest effects when experimental procedures are rehearsed or verified directly after the manipulation, suggesting that participant involvement increases its effectiveness.

Measurement moment. For measurement moment, anti–free will manipulations had a larger influence on free will beliefs, $t(92.3) = -2.32$, $p = .022$ (see Table 2), when the belief was assessed before compared with after the primary dependent variable was measured (i.e., downstream consequences). The same statistical test was not significant for determinism beliefs, $t(33.2) = 0.95$, $p = .349$ (see Table 3).

In sum, anti–free will manipulations have the strongest effects when free will beliefs are measured directly after the experimental manipulation instead of at the end of the study. This suggests the effects of the manipulation fade over the course of the experimental session.

Baseline condition. Anti–free will manipulations had a larger effect on free will beliefs when the baseline was a pro–free will condition than when it was a neutral condition, $t(106.7) = -4.11$, $p < .001$ (Table 2). A similar effect was also apparent for determinism beliefs, but did not reach statistical significance, $t(25) = 1.05$, $p = .304$ (see Table 3). As discussed earlier, however, baseline condition correlated highly with manipulation type and manipulation task. As there were no studies in some cells of the Baseline Condition × Manipulation Type cross-table, we decided to control for manipulation type by looking at the effect of baseline condition for the manipulation type where baseline condition was most balanced (i.e., statements) to optimize power (Tanner-Smith et al., 2016). This revealed that even when restricting the analysis to those studies using statements, anti–free will messages were more effective when the baseline condition was a pro–free will condition than when it was a neutral condition ($p = .032$). To control for participant involvement, we fitted a model including both baseline condition and participant involvement, which did not change the results ($p = .024$).

Taken together, comparing anti–free will messages with pro–free will messages results in larger effect sizes than comparing anti–free will messages with neutral messages. This suggests that belief in free will can not only be reduced but can also be increased by experimental manipulations.

Secondary moderators

Age. The mean age of the sample did not influence the effectiveness of the manipulation for either free will beliefs, $t(28.2) = 0.27$, $p = .793$ (see Table 2), or determinism beliefs, $t(19.1) = 0.21$, $p = .838$ (see Table 3).

Sex. The proportion of female participants in the sample did not influence the effect of the manipulation on either free will beliefs, $t(31.5) = 0.53$, $p = .598$ (see Table 2), or determinism beliefs, $t(17.3) = -0.14$, $p = .893$ (see Table 3).

Continent. The manipulation had a stronger effect on samples collected in the United States or on U.S.-based platforms (e.g., MTurk) than on samples collected in Europe or on Europe-based platforms (e.g., Prolific) for free will beliefs, $t(101) = -2.50$, $p = .014$ (see Table 2), but not for determinism beliefs, $t(45.2) = 0.99$, $p = .329$ (see Table 3).

Test location. Whether the experiment was conducted in the lab or online did not influence the effect of the manipulation on either free will beliefs, $t(77.2) = 0.45$, $p = .655$ (see Table 2), or determinism beliefs, $t(42.6) = -0.55$, $p = .586$ (see Table 3).

Sample type. The effectiveness of the manipulation did not differ between student samples and samples collected via online platforms for either free will beliefs, $t(62.6) = 0.81$, $p = .422$ (see Table 2), or determinism beliefs, $t(28.7) = 0.62$, $p = .543$ (see Table 3).

Sensitivity Analyses

To test the robustness of our main results, we ran four sensitivity analyses investigating how changing the analysis procedure influenced (a) the effect size of the manipulation and (b) the evidence for publication bias. First, we varied the ρ parameter between 0.1 and 1.0 in steps of 0.2. This resulted in identical effect sizes for both free will beliefs and determinism beliefs. Second, we calculated effect sizes preferentially from the test statistics instead of from the means and standard deviations. This again resulted in identical estimates for both free will beliefs and determinism beliefs. Third, we repeated the analysis without replacing outlier values. This led to slightly larger effect sizes for determinism beliefs ($g_z = 0.19$) but identical effect sizes for free will beliefs. Finally, we conducted a multilevel meta-analysis instead of using RVE. This resulted in identical effect sizes for both free will beliefs and determinism beliefs. Thus, none of the four changes to our analysis procedure had a notable influence on the manipulation's effect size. Similarly, the same four changes also did not have a notable influence on the publication bias results, except that the PEESE-corrected effect size

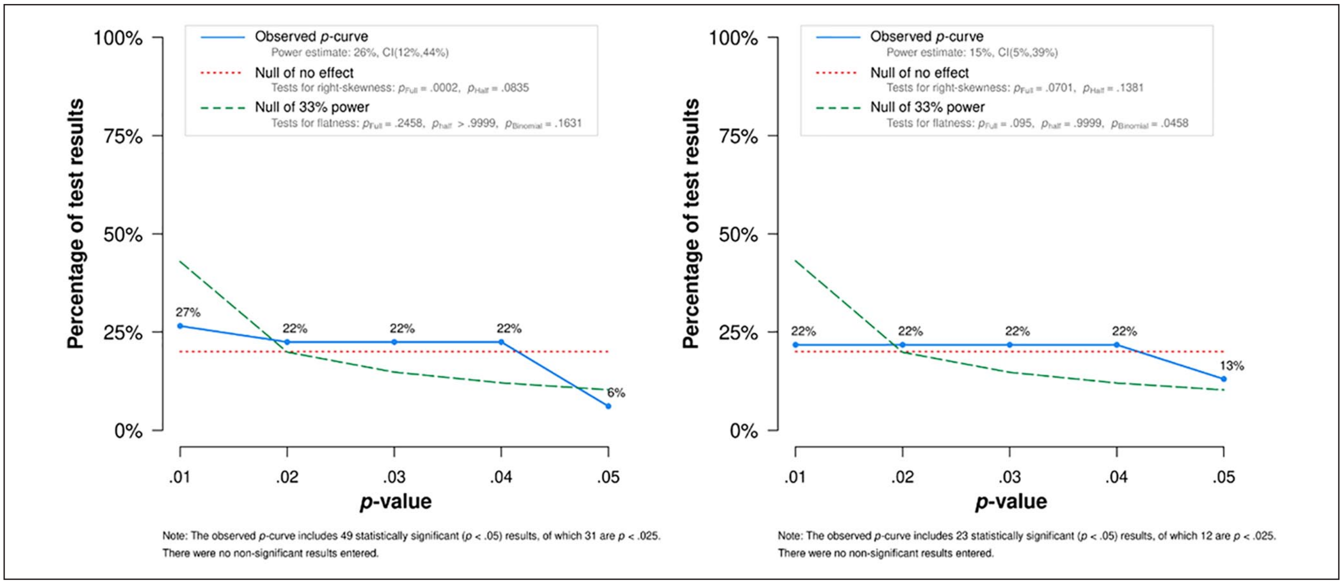

Figure 3. *P*-curve on the entire data set (left) and on the studies with a significant manipulation check (right).
Note. See the OSF folder (https://bit.ly/2L69prl) for the evidence tables.

for determinism beliefs was no longer significant when outliers were not replaced ($g_z = 0.12$, $p = .201$). For a full overview of the results of the different sensitivity analyses, see OSF: https://bit.ly/2L69prl.

Summary of the effectiveness of free will belief manipulations. In sum, our analyses indicate that anti–free will manipulations are able to successfully influence belief in free will and related beliefs. The strongest effects occur when anti–free will texts (e.g., the Crick text) are paired with anti–free will statements. Manipulation effects are strongest when measured directly after the manipulation, and group differences are largest when the anti–free will condition is compared with a pro–free will condition. However, we found no moderation by the demographics investigated here.

Research Question 2: Do Free Will Belief Manipulations Have Downstream Consequences?

P*-curve analysis.* A p-curve is considered to contain evidential value if the null-hypothesis that a zero true effect size underlies all effect sizes can be rejected. This null-hypothesis can be rejected if the half p-curve, considering only p-values $<.025$, is significantly right-skewed at $p < .05$ or if the full p-curve, considering all p-values $< .05$, and half p-curve are both right-skewed at $p < .10$. Similarly, a p-curve is considered to signal the absence of evidential value if the half p-curve is significantly flatter than the curve expected with 33% power at $p < .05$ or if both the full and the half p-curve are flatter than 33% power at $p < .10$ (Simonsohn et al., 2015).

A first p-curve analysis with 49 experiments testing the influence of anti–free will manipulations on behavior, attitudes, and cognition revealed that both the full ($z_{\text{full}} = -3.54$, $p_{\text{full}} < .001$) and half p-curve ($z_{\text{half}} = -1.38$, $p_{\text{half}} = .084$) had a p-value $< .10$, and hence that the set of studies contained evidential value. However, a visual inspection of the p-curve revealed that it was almost entirely flat (see Figure 3). Although the flatness test did not find evidence for the absence of an evidential value ($z_{\text{half}} = 5.86$, $p_{\text{half}} > .999$), a robustness check, as recommended by Simonsohn et al. (2015), indicated that removing the single most significant p-value across all included studies (i.e., Ching & Xu, 2018) was sufficient to render the test of evidential value nonsignificant ($z_{\text{half}} = -0.70$, $p_{\text{half}} = .128$). A second robustness check indicated that evidential value was stronger ($z_{\text{full}} = -4.57$, $p_{\text{full}} < .001$, $z_{\text{half}} = -2.49$, $p_{\text{half}} = .006$) when the second rather than the first result was used from studies reporting multiple relevant results, and this was still true even after removing the most significant test.

A potential explanation for why the main analysis did not find clear evidence for downstream effects could be that some of the included studies were not able to confirm that their manipulation changed participants' belief in free will. That is, a study cannot be expected to have downstream consequences if it did not successfully manipulate the belief in free will in the first place. To test this hypothesis, we ran a second p-curve analysis including only those studies with a statistically significant manipulation check (i.e., $p < .05$). Across 23 tests, this revealed no evidence for either the presence ($z_{\text{half}} = -1.09$, $p_{\text{half}} = .138$) or absence ($z_{\text{half}} = 3.66$, $p_{\text{half}} = .999$) of evidential value (Figure 3). A robustness check using the second instead of the first reported test in studies with multiple relevant tests indicated the presence of evidential value ($z_{\text{full}} = -1.87$, $p_{\text{full}} = .031$, $z_{\text{half}} = -1.48$, $p_{\text{half}} = .069$), but removing the single most significant test was again

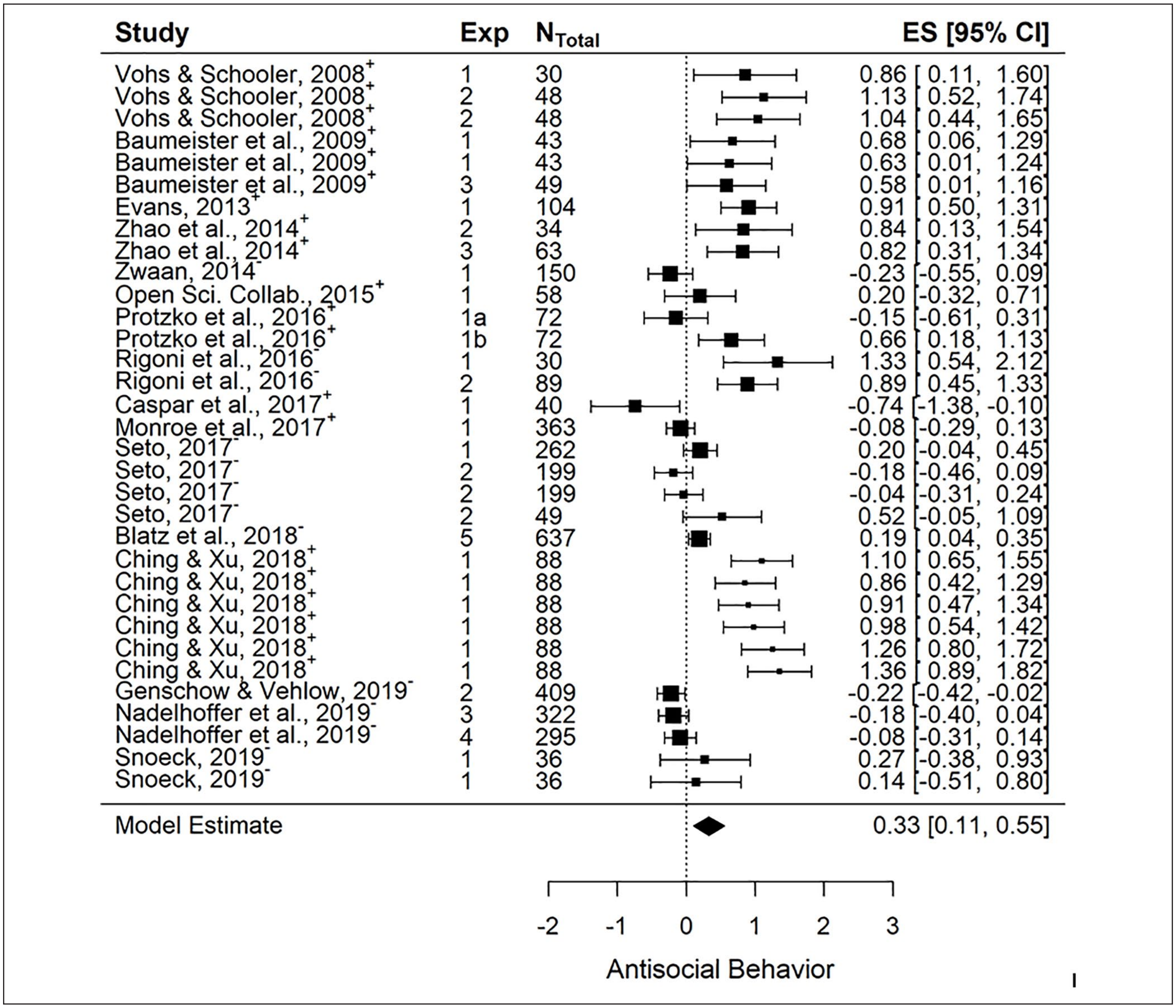

Figure 4. Forest plot of the effect of the manipulation on antisocial behavior.

Note. Prosocial effect sizes are reverse coded. Published studies are annotated with "+" and unpublished studies with "-." Nadelhoffer er al. (2019) and Genschow and Vehlow (2019) were included in the meta-analysis as unpublished studies but were published later as Nadelhoffer et al. (2020) and Genschow and Vehlow (2021) after data analysis had already been completed. CI = confidence interval.

sufficient to make this evidence disappear ($z_{\text{half}} = -0.71$, $p_{\text{half}} = .103$).

Taken together, the *p*-curve analysis finds only little evidence for the hypothesis that manipulating belief in free will has downstream consequences on behavior, attitudes, and cognition. That is, although some analyses provided weak evidence for downstream consequences, this evidence disappeared in all but one case when the single lowest *p*-value was excluded (Simonsohn et al., 2015). At the same time, there was also no conclusive evidence for the absence of an effect. Hence, the *p*-curve found no clear evidence for, but also not against, the hypothesis that free will belief manipulations have downstream consequences.

Antisocial behavior. A random-effects meta-analysis with robust variance estimation indicated an effect size of $g = 0.33$ for the influence of anti–free will manipulations on anti–social behavior, $t(21.1) = 3.10$, $p = .005$, 95% CI = [0.11, 0.55], $m = 23$, $k = 33$.[3] However, a visual inspection of the forest plot (see Figure 4) suggested that this significant effect was mostly driven by studies with smaller sample sizes, whereas studies with larger sample sizes tended to report null results. This was confirmed by a hierarchical Egger test, indicating a significant positive relationship between the effect size and the standard error, $\beta = 3.01$, $t(12.8) = 3.52$, $p = .004$. In other words, while the meta-analysis suggested an effect of the manipulation on antisocial

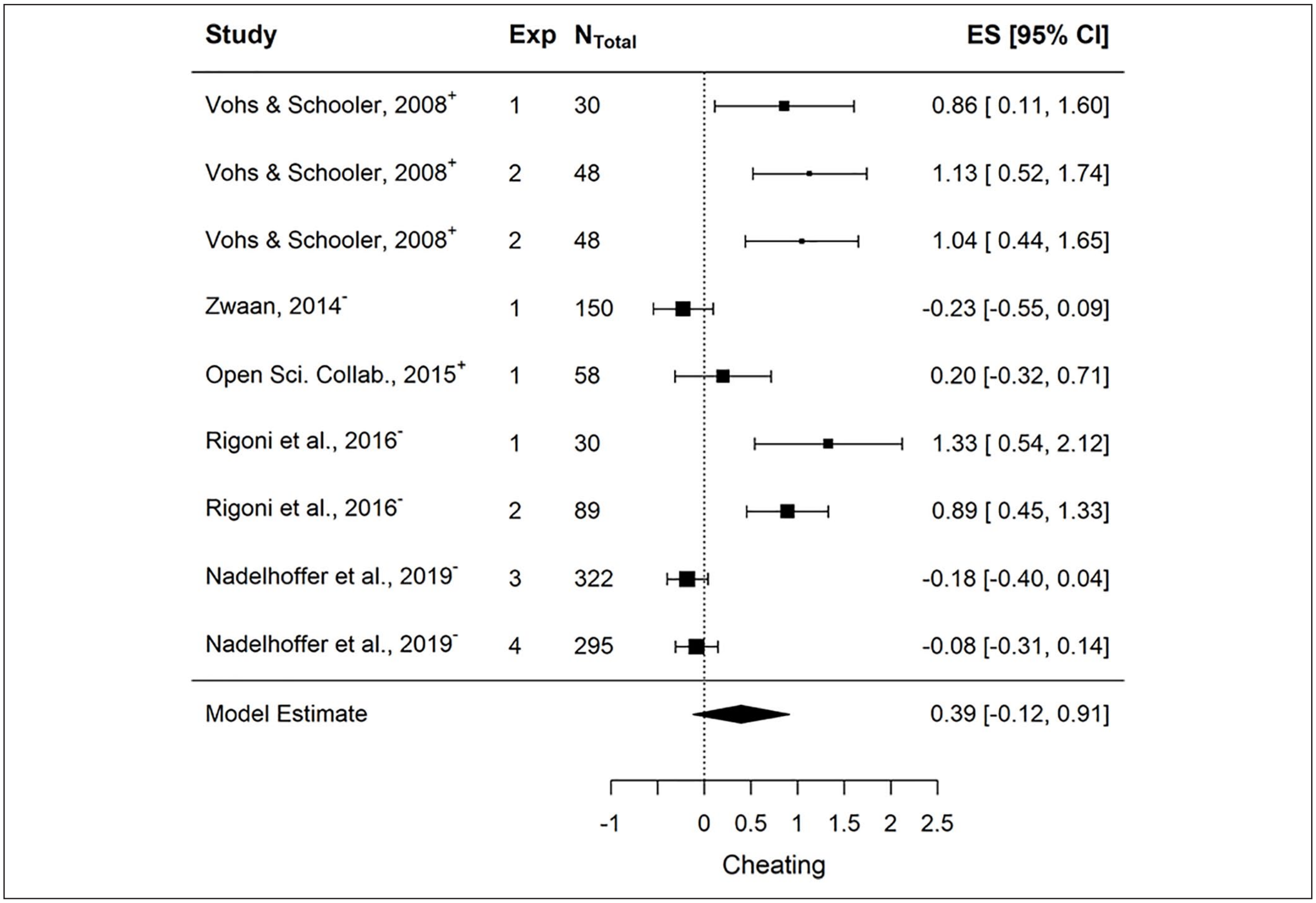

Figure 5. Forest plot of the effect of the manipulation on cheating.
Note. Published studies are annotated with "+" and unpublished studies with "-." Nadelhoffer er al. (2019) and Genschow & Vehlow (2019) were included in the meta-analysis as unpublished studies but were published later as Nadelhoffer et al. (2020) and Genschow and Vehlow (2021) after data analysis had already been completed. CI = confidence interval.

behavior, additional analyses indicated that this effect may have been driven by small-study bias. Correcting the effect size for this bias using PET ($g = 0.00$) or PEESE ($g = 0.00$) resulted in substantially reduced and nonsignificant effect sizes. Similarly, even though there was no significant difference between published ($g = 0.50$) and unpublished ($g = 0.15$) studies, $t(19.6) = 1.76$, $p = .095$, the effect was descriptively larger for published studies and was no longer significant if only unpublished studies were included in the analysis, $t(8.6) = 1.16$, $p = .279$. Taken together, this indicates that there is insufficient evidence that anti–free will manipulations influence antisocial behavior.

However, including all studies in a meta-analysis of this sort may be problematic because studies that did not change the belief in free will (e.g., failed at the manipulation check) could not hope to have downstream consequences on antisocial behavior. That said, a moderator analysis comparing effect sizes in studies with a significant manipulation check ($g = 0.30$, $m = 10$, $k = 13$) to effect sizes in studies with no manipulation check or a nonsignificant manipulation check ($g = 0.37$, $m = 13$, $k = 20$) revealed no significant difference, $t(19.4) = -0.34$, $p = .736$. If anything, the effect size was even numerically smaller in the set of studies reporting a successful manipulation check, and this effect size was not significant, $t(8.7) = 2.11$, $p = .065$, 95% CI = [−0.02, 0.61].

In sum, the analysis showed that the effect of anti–free will manipulations on antisocial behavior was no longer significant after controlling for publication and small sample biases. This was true even when we only included studies that found a significant effect of the manipulation on belief in free will and indicates that there is insufficient evidence for the idea that manipulating belief in free will influences antisocial behavior.

Cheating, conformity, and punishment. No effects of anti–free will manipulations were found on cheating, $g = 0.39$, $t(6.8) = 1.81$, $p = .114$, 95% CI = [−0.12, 0.91], $m = 8$, $k = 9$, conformity, $g = 0.26$, $t(4.7) = 1.63$, $p = .168$, 95% CI = [−0.16, 0.68], $m = 6$, $k = 9$, or punishment, $g = -0.15$, $t(7.5) = -1.60$, $p = .151$, 95% CI = [−0.36, 0.07], $m = 9$, $k = 9$.

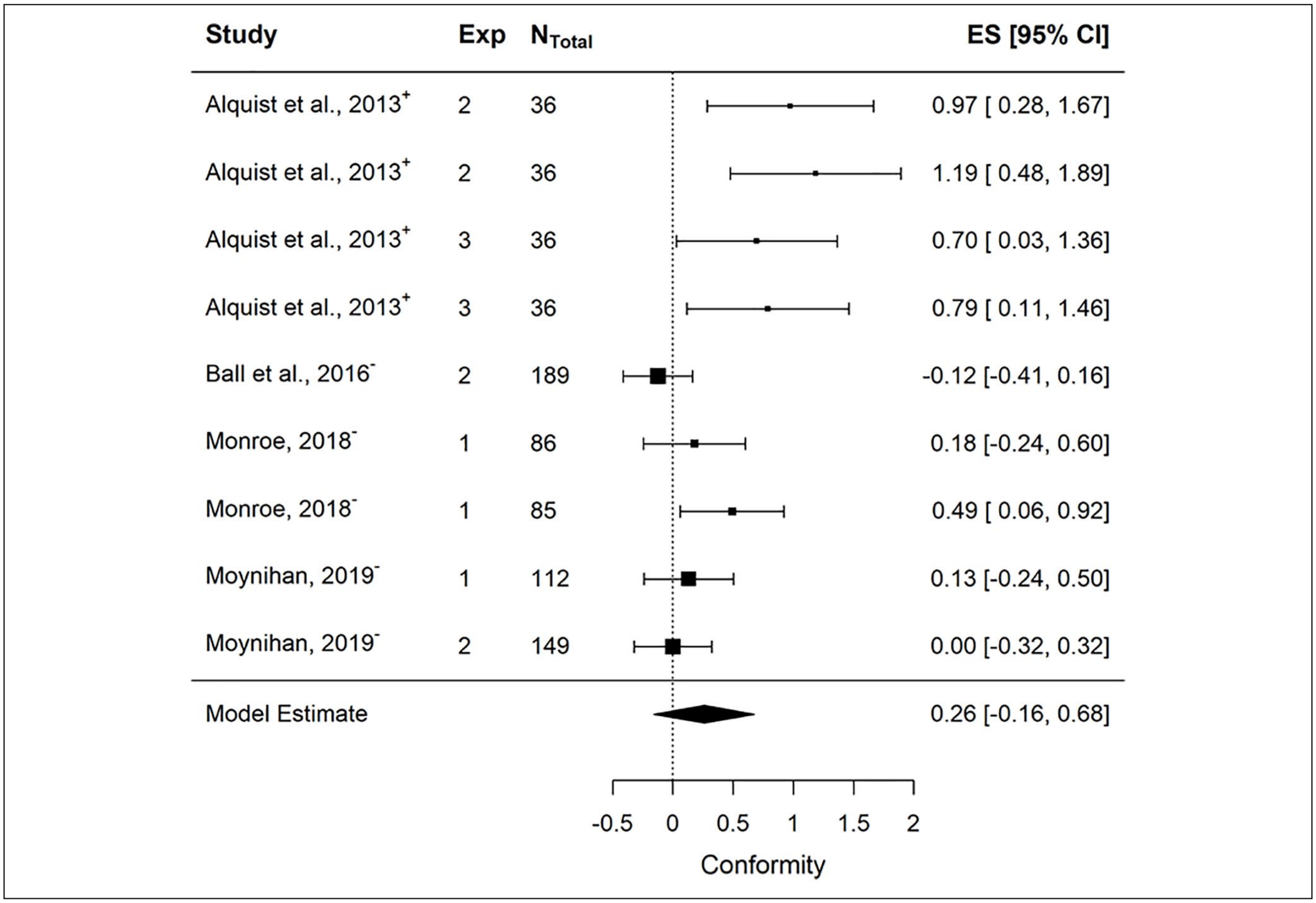

Figure 6. Forest plot of the effect of the manipulation on conformity.
Note. Published studies are annotated with "+" and unpublished studies with "-."CI = confidence interval.

In line with the antisocial behavior meta-analysis, visual inspection of the forest plots (see Figures 5–7) indicated that studies with smaller sample sizes tended to report significant effects, whereas studies with larger sample sizes tended to report null effects. This was confirmed by a hierarchical Egger test, showing significant relationships between the effect size and the standard error for all three dependent variables (all $p \leq .014$). Correcting the effect sizes for these small-study biases using PET (all $g = 0.00$) or PEESE (all $g = 0.00$) eliminated all possible evidence. However, the Egger and PET/PEESE results should be interpreted with care, as the limited *df*s (all ≤ 4.1) likely led to an inflated false-positive rate (Tipton, 2015). Similarly, for cheating and conformity, there were not enough studies to reliably test if the effect remained significant when including only the unpublished studies (both $df \leq 3.9$), but in both cases, effect sizes were numerically smaller for unpublished ($g_{\text{cheating}} = 0.23$, $g_{\text{conformity}} = 0.04$) than for published studies ($g_{\text{cheating}} = 0.69$, $g_{\text{conformity}} = 0.90$). For punishment, effect sizes were also numerically smaller for unpublished studies ($g = -0.10$) than for published studies ($g = -0.23$), but this difference did not reach significance, $t(6.1) = -0.61$, $p = .563$.

Finally, we investigated if the effect was modulated by whether a significant manipulation check was reported and if it was significant when including only those studies reporting a significant manipulation check. However, for cheating and conformity, there were insufficient studies to test either of these two hypotheses (all $df \leq 3.77$). Although the cheating effect size was descriptively larger for studies reporting a significant manipulation check ($g = 0.94$, $m = 3$, $k = 4$) than for studies not reporting a significant manipulation check ($g = 0.02$, $m = 5$, $k = 5$), this difference should not be taken at face value, given the limited number of studies and given that three of the four included effect sizes with a significant manipulation check came from the same paper (i.e., Vohs & Schooler, 2008). The conformity effect size was descriptively smaller for studies reporting a significant manipulation check ($g = 0.10$, $m = 3$, $k = 4$) than for studies not reporting a significant manipulation check ($g = 0.49$, $m = 3$, $k = 5$). For punishment, the difference between studies that did or did not report a significant manipulation check was not significant, $t(6.0) = 0.24$, $p = .821$. There were not enough studies to test whether the effect remained significant when only including studies reporting a significant

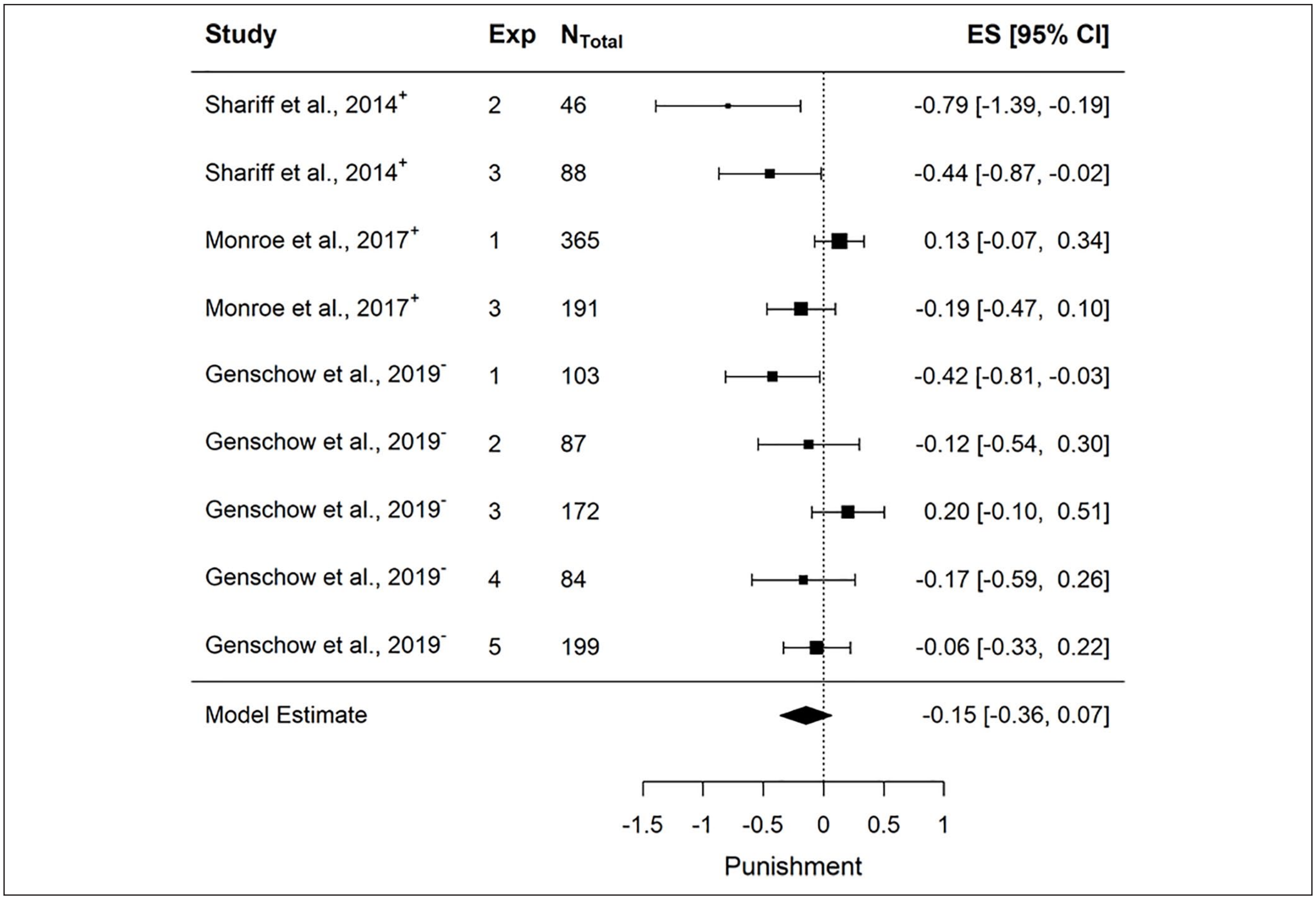

Figure 7. Forest plot of the effect of the manipulation on punishment.
Note. Published studies are annotated with "+" and unpublished studies with "-."CI = confidence interval.

manipulation check (df = 3.9). However, if anything, the effect was numerically smaller in studies that reported a significant manipulation check ($g = -0.13$, $m = 5$, $k = 5$) than in studies that did not report a significant manipulation check ($g = -0.18$, $m = 4$, $k = 4$).

In conclusion, after controlling for small-study biases, the effect sizes of free will belief manipulations on cheating, conformity, and punishment became statistically indistinguishable from zero.

Sensitivity analyses. To test the robustness of our main results, we used the same approach as for the first research question. That is, for all four dependent measures, we investigated how changing the analysis procedure influenced (a) the effect size of the manipulation and (b) the evidence for publication bias. In contrast to the first research question, however, we now ran only three instead of four sensitivity analyses because there were no outlier values in the data of the second research question. First, we varied the ρ parameter from 0.1 to 1.0 in steps of 0.2. This led to identical effect sizes for all four dependent measures. Second, we calculated effect sizes preferentially from the test statistics instead of from the means and standard deviations. This led to a slightly larger effect size for conformity (g_z = 0.28, p = .138), but did not change the effect sizes of the other dependent variables. Finally, we conducted a multilevel meta-analysis instead of using RVE. This resulted in slightly higher effect sizes for antisocial behavior (g_z = 0.35, p = .004), cheating (g_z = 0.43, p = .093), and conformity (g_z = 0.33, p = .128) but did not change the effect size for punishment. Thus, in line with the first research question, varying the analysis strategy did not have a notable influence on the results. The same was also true for the publication bias results, both in the sense that all sensitivity analyses indicated evidence for such a bias and in the sense that correcting the bias made the effects disappear regardless of the analysis approach. For a full overview of the sensitivity analysis results, see OSF: https://bit.ly/2L69prl.

Discussion

Past research has shown that reducing individuals' belief in free will affects societally relevant behaviors such as cheating (Vohs & Schooler, 2008) and other antisocial behaviors

(Baumeister et al., 2009; Zhao et al., 2014). These and similar results have been used by some scholars as an argument that anti–free will and deterministic viewpoints should be kept away from society, because they may change the way people interact with each other. However, a number of studies reported difficulties in replicating some of the key results in the field (Genschow et al., 2020; Giner-Sorolla et al., 2016; Monroe et al., 2017; Open Science Collaboration, 2015; Schooler et al., 2014; Shariff & Vohs, 2014). While these failed replications call into question the societal relevance of belief in free will, it is not yet clear what caused them. In this article, we tested three possible explanations that may account for the failed replications. First, it could be that the failed replications are false negatives. That is, they were not able to detect an effect that is actually real. Second, it could be that the manipulations commonly used in the literature do not alter individuals' belief in free will and thus have no impact on other behaviors. Third, it could be that free will belief manipulations reduce belief in free will, but that this does not have any downstream consequences.

Our meta-analysis favors the third explanation. That is, we found that beliefs related to free will can be effectively manipulated by commonly used experimental manipulations, although the effects were rather small, with $g = -0.29$ for free will beliefs and $g = 0.17$ for determinism beliefs. However, we did not find evidence for downstream consequences on attitudes, behavior, or cognition. In the remainder of this article, we discuss the implications of our results for the processes underlying free will belief manipulations, the potential reasons for why these manipulations did not have downstream consequences, the societal implications of the meta-analysis, possible steps for future research, and the limitations of the present meta-analysis.

Processes Underlying Free Will Belief Manipulations

The analysis of our first research question on the effectiveness of free will belief manipulations indicated that beliefs related to free will can be influenced by experimental manipulations. Although we found evidence for publication bias, we also found that this cannot explain the effect. To investigate the conditions under which these manipulations were effective, we conducted several moderator analyses. The results of these analyses further the understanding of free will belief manipulations manifold.

First, previous research indicated that most people believe that they have free will (Baumeister et al., 2009; Nahmias et al., 2005). Based on this finding, it has been assumed that free will beliefs can only be decreased (for an overview, see Ewusi-Boisvert & Racine, 2018). While this assumption has never been systematically investigated, the present meta-analysis allowed us to test whether experimental manipulations can also increase the belief in free will. The results demonstrate that belief in free will can be reduced as well as increased by experimental means.

Second, the changes in the beliefs were relatively weak. An open question is whether the conceptualization of free will in the manipulation influences the magnitude of the effect. Indeed, philosophical definitions of free will are rather diverse and complex (Carey & Paulhus, 2013) and may thus be impalpable for laypeople. If true, free will belief manipulations relying on such definitions are unlikely to have a strong impact on laypeople's beliefs. To mitigate this problem, future research could investigate whether manipulations that focus more specifically on laypeople's concepts of free will produce stronger effects.

Third, from a methodological point of view, it is important to know which manipulations most reliably alter belief in free will. By comparing all the different manipulations used in previous research, our analysis suggests that an approach in which participants have to both read a text and reproduce statements produces the largest effects.

Fourth, an interesting and open question was how deeply participants have to process the stimulus information for the manipulation to succeed. To answer this question, we investigated whether the effectiveness of the manipulation depended on the degree to which participants had to engage with the task (e.g., by summarizing or rewriting presented messages). Manipulations worked best when participants had to summarize the stimulus information directly after the manipulation. In other words, actively processing the provided information (e.g., by summarizing it) increased the effect of the manipulation. This indicates that participants' involvement in the task strengthens the effect of free will belief manipulations—a finding that is in line with previous research on the self-generation effect (Slamecka & Graf, 1978) where self-generated information increases memory performance of that information.

Fifth, past research left unanswered the question of how long the effects of free will belief manipulations last. Measuring beliefs at the end of the experiment assumes the manipulation lasts through the duration of the experiment; this is not necessarily the case, however. For instance, it could also be that free will belief manipulations do not fully change people's beliefs but merely activate an anti–free will mindset for a brief moment. To test how long-lasting the effects of free will belief manipulations are within the span of an experiment, we investigated if the effect of the manipulation differed depending on whether beliefs were measured immediately after the manipulation or at the end of the experiment (i.e., after completing another task). At least for belief in free will, the effect reduced over time. While the same was not true for belief in determinism, it is important to note that deterministic beliefs were generally less influenced by moderators. This could mean that determinism beliefs are more robust than beliefs in free will, but there are also several alternative explanations. For example, the meta-analysis included fewer studies measuring belief in determinism than belief in free will and the determinism moderator analyses may therefore have had less statistical power to detect such effects (Hempel et al., 2013). Alternatively, it could be that because belief in determinism was less influenced by the manipulation as such, there was a floor effect, making it difficult to further reduce the effect of the manipulation.

Sixth, the meta-analysis yielded a borderline nonsignificant trend indicating that anti-free will manipulations influenced beliefs in free will slightly more strongly than beliefs in determinism. Given that this effect was not very strong and not significant based on conventional levels of significance, it should be interpreted with care. Nevertheless, if confirmed by future research, a potential explanation could be that existing manipulations target beliefs in free will more strongly than beliefs in determinism. As beliefs in free will are not typically correlated with belief in determinism (Nadelhoffer et al., 2014), it is not surprising that such manipulations would have different effects on both beliefs. The stronger effect of anti–free will manipulations on free will beliefs could thus indicate that these manipulations specifically influence the belief in free will and only influence other beliefs to a lesser degree. However, in contrast to this view, the results of our meta-analysis suggest that the influence of anti–free will manipulations tends to be rather unspecific. For example, we found that the effect of anti–free will manipulations is at least equally strong, if not stronger, on belief in dualism than on belief in free will. Similarly, we found that belief in free will manipulations influence not only belief in free will and belief in scientific determinism but also belief in fatalistic determinism. There are different possible explanations for these unspecific effects. Some of the beliefs are correlated with each other and are thus quite unspecific themselves. For example, belief in free will correlates positively with belief in dualism (Nadelhoffer et al., 2014; Wisniewski et al., 2019). In addition, it is likely that the manipulations affect not only beliefs related to free will, but also other psychological and cognitive factors. For, instance, it is conceivable that challenging the fundamental belief that free will exists leads to a general feeling of confusion and uncertainty about the world.

Insufficient Evidence for Downstream Consequences and Its Potential Reasons

While we found clear evidence that beliefs related to free will can be experimentally manipulated, there was little evidence that these free will manipulations have downstream consequences. That is, both *p*-curve analyses across all dependent variables and meta-analyses on internally coherent sets of dependent variables found either no evidence for downstream consequences or weak evidence that, in all but one instance, disappeared when correcting for extreme values or small-sample effects. This remained true even if only studies with significant manipulation checks (i.e., significant effects on belief in free will and/or determinism) were used in the analysis. This indicates that there is currently insufficient evidence for downstream consequences.

How can this be reconciled with the fact that some primary studies did find downstream consequences? The results of our meta-analysis suggest that one likely candidate is publication bias. Indeed, the results of the Egger test indicated that studies with larger effect sizes tended to be studies with higher standard errors. This is a typical sign of publication bias because studies with small samples or noisy measurements only reach statistical significance if the effect size is large (Sterne & Egger, 2005). Hence, a potential reason why we could not confirm previous studies finding downstream consequences is that these studies are false positives.

Nevertheless, it is important to note that absence of evidence does not equal evidence of absence. Indeed, the *p*-curve analysis lumped together a variety of different variables, and the meta-analysis focused only on a limited subsample of outcomes for which there were enough data points. Thus, we cannot rule out that there are other specific variables that are influenced by free will belief manipulations. For example, we did not have enough data to investigate the influence of free will belief manipulations on neurocognitive processes (Rigoni et al., 2011, 2012, 2015), feelings of alienation (Seto & Hicks, 2016), attributions of other people's actions (Genschow et al., 2017a), perceived meaningfulness of life (Crescioni et al., 2016; Moynihan et al., 2019), perceived gratitude (MacKenzie et al., 2014), counterfactual thinking (Alquist et al., 2015), or risk-taking behavior (Schrag et al., 2016). For these and potential other variables that have not been studied yet, it remains open to what degree they are influenced by free will belief manipulations, although one of the effects on neurocognitive processes recently failed to replicate (Eben et al., 2020).

It is also possible that the failure to find robust evidence for previously reported downstream effects of free will belief manipulations has to do with subtle differences in the manner in which the experimental protocols were implemented by the research teams that were versus were not successful in finding such effects. Differences in experimental outcomes between labs might reflect variations in the implementation of effective experimental procedures (e.g., ensuring that participants were sufficiently motivated, believed the experimental ruse, and followed the protocol), could be due to the inclusion of artifacts that produced false-positive effects (e.g., introducing experimenter demands characteristics), or could be driven by the belief of the research teams to find an effect (Doyen et al., 2012). Future studies, perhaps by research teams that have previously been effective in finding downstream effects, might profitably examine whether procedural differences can be identified that reliably differentiate between when such effects are versus are not observed.

Even if currently employed manipulations do turn out, as the present findings suggest, to not produce robust downstream effects, this does not necessarily mean that such effects cannot happen. It might be that existing manipulations are not suited for this purpose. In other words, it might well be that experimentally reducing beliefs in free will has downstream consequences but that established free will belief manipulations are not able to produce them. There are several reasons why this might be the case. As mentioned earlier, the manipulations are rather unspecific and do not only affect belief in free will and determinism, but also other beliefs. This is problematic for at least two reasons. On the one hand, an unspecific manipulation may be detrimental in finding downstream consequences because the different factors influenced by the manipulation may counteract the effect of free will beliefs on the dependent variable. On the other hand, an unspecific manipulation opens the question to which degree downstream consequences (if any) are actually driven by free will beliefs or rather by other beliefs and psychological variables.

In addition, existing manipulations of free will beliefs have only weak effects on free will ($g = -0.29$) and determinism ($g = 0.17$) beliefs. Given that any effect of anti–free will manipulations on attitudes and behavior is likely smaller than their effect on the beliefs they purport to change, this makes finding evidence for downstream consequences particularly challenging. Indeed, research on the relationship between behavioral intentions and actual behavioral change shows that the influence of behavioral interventions on behavior is roughly twice as small as their influence on behavioral intentions (Webb & Sheeran, 2006). The same problem is likely true for belief in free will and its downstream consequences. In this sense, it may be helpful to relate our findings to the idea of "attitude strength" (Petty & Krosnick, 2014). Attitude strength is defined "as the extent to which attitudes manifest the qualities of durability and impactfulness" (Petty & Krosnick, 2014; p. 3). Durability refers to the degree to which an attitude can be changed over time and how strongly it resists attacks (e.g., persuasion attempts). Impactfulness relates to how strongly an attitude influences information processing, judgments, and behaviors. Applying the idea of attitude strength to beliefs in free will may help to explain why we did not find evidence for downstream consequences. That is, free will beliefs are most likely to influence attitudes, behavior, and cognition when they are strong. However, people with strong beliefs are least likely to be influenced by manipulation. Moreover, even if people were influenced by the manipulation, our meta-analysis indicates that these new beliefs are likely to be weak. Given that weak beliefs are less likely to influence behavior, this might explain why we did not find evidence for downstream consequences in the current meta-analysis.

Related to this idea, another potential reason for the difficulty to find downstream effects could be that the effect of the manipulation on free will beliefs fades over time, as our results indicate. As a result, it is possible that in some experiments, the effect of the manipulation had already disappeared when the dependent variable was measured. From this perspective, downstream consequences may be visible only when they are measured directly after the manipulation. Future research should test this hypothesis more directly.

Finally, it is possible that some participants are not committed to a particular belief about free will but nevertheless have strong moral convictions about behaviors like cheating (see, for example, Dubljević, 2013), for instance. In this case, manipulating belief in free will may influence beliefs in free will without having downstream consequences, because moral beliefs are held separate from their other beliefs. Future research could try to test this hypothesis in more detail.

Taken together, many possible explanations exist for why we could not find evidence for downstream consequences in this meta-analysis. An important task for future research will be to investigate these explanations and to determine whether and under which circumstances free will belief manipulations influence attitudes, behavior, and cognition.

Societal Implications

Whether free will exists is part of a long-standing philosophical debate (e.g., Dennett, 2015; Van Inwagen, 1983). However, ever since cognitive neuroscientists and psychologists started claiming that humans' perception of free will is nothing more than an illusion (e.g., Crick, 1994; Harris, 2012; Wegner, 2002), anti–free will viewpoints have become in vogue not only in academia (e.g., Greene & Cohen, 2004) but also in popular media (e.g., Chivers, 2010; Griffin, 2016; Racine et al., 2017; Wolfe, 1997). When psychological research found that presenting individuals with such anti–free will viewpoints influences fundamental behavior, cognition, and attitudes, the question arose whether the public press should publish such anti–free will viewpoints. While some philosophers argue that undermining people's belief in free will would have catastrophic consequences, as free will forms the basis for moral behavior (e.g., Smilansky, 2000, 2002), other philosophers argue that disbelieving in free will might also have positive effects because it could lead to abandoning retribution-based morality and illusory beliefs in a just world (Caruso, 2014; Greene & Cohen, 2004; Nadelhoffer, 2011). The present research adds to this debate by suggesting that confronting individuals with anti–free will viewpoints might not have as strong consequences as has been previously assumed.

Although these manipulations affect people's beliefs in free will, we did not find evidence for effects on behavior, cognition, or attitudes. This is in line with recent findings indicating that although professional judges' beliefs in free will are influenced by reading anti–free will texts, reading these viewpoints does not influence their judgments (Genschow et al., 2020). However, it is important to keep in mind that all these findings only speak to the effect of presenting individuals with a single anti–free will viewpoint. While such a short exposure may not have downstream consequences, it remains an open question whether more concentrated and repeated presentations of anti–free will messages, as may happen in real-life, could nevertheless have important consequences. Indeed, preliminary support for this possibility comes from the results of the present meta-analysis, demonstrating that a combined manipulation of presenting participants with a text and statements has the strongest impact on individuals' belief in free will. It would be interesting to test to what extent repeated exposure to anti–free will messages may have stronger effects on individuals' belief in free will and thereby also lead to downstream consequences.

In the same vein, although we did not find sufficient support for the idea that conventional belief in free will manipulations have downstream consequences, it is important to note that we do not necessarily argue that belief in free will itself has no impact. That is, on an interindividual level, belief in free will may well contribute to societally relevant behaviors such as retributive punishment and anti- or prosocial behavior—to name just a few examples. Indeed, several studies could reliably replicate the link between belief in free will and retributive punishment on a correlational level (Genschow et al., 2017a; Martin et al., 2017). Similar strong correlations were also found between belief in free will and job satisfaction (Feldman et al., 2018), intentional attributions (Genschow et al., 2017a, 2019a; Genschow & Lange, in press), as well as between free will beliefs and just world beliefs, religious worldviews, and a conservative worldview (Carey & Paulhus, 2013; Genschow & Vehlow, 2021). This suggests that on a correlational level, belief in free will may well be connected to societally relevant behaviors.

Taken together, there is a debate about whether anti–free will viewpoints should be discussed in the public media. Our findings suggest that the influence on society may be weaker than previously assumed. In this respect, we would like to argue that discussions about the implications of believing in free will should distinguish between scientific facts and philosophical speculations (Schooler, 2010) as well as acknowledge methodological limitations of the cited research (Racine et al., 2017).

Limitations of the Present Meta-Analysis

There are a few limitations to this meta-analysis that call for a careful discussion. First, all publication bias correction methods have downsides (Carter et al., 2019; Stanley, 2017) and should hence be interpreted with care. We used PET and PEESE because they are easily incorporated within the RVE framework (Rodgers & Pustejovsky, 2020) and have been shown to retain reasonable false-positive rates across a wide range of scenarios (Carter et al., 2019). However, PET-PEESE can suffer from low power, especially when sample sizes are small, heterogeneity is high, or when there is either very little or very heavy use of questionable research practices (Carter et al., 2019; Rodgers & Pustejovsky, 2020). Therefore, in addition to using bias correction methods, we also compared published with unpublished studies and tested if the effects remained significant if only unpublished studies were considered. While no single method is perfect by itself, we believe that by combining these different methods, we were able to get a clearer overview of the underlying true effects and the degree to which they were inflated by publication bias.

Second, the current meta-analysis included a large number of unpublished studies. This approach is consistent with recent calls to include unpublished research in meta-analyses as a means to counter publication bias (Polanin et al., 2016). However, a potential downside is that the unpublished studies may have been of inferior quality to the published research. One way to rule this out is to code study quality. Existing tools focus strongly on clinical intervention studies, however, and are therefore not easily applied to the research synthesized here (Armijo-Olivo et al., 2012). Instead, the main indicators of study quality in research on free will belief manipulations (and other experimental social psychological research) are blinded and random assignment, sample size, and the use of validated manipulations and scales. With respect to blinded and random assignment, we had initially planned to include a moderator coding whether the experimenter was blind to the manipulation. However, after completing the coding, it was clear that this variable was almost completely confounded with study location, such that online studies were blinded, whereas papers on studies conducted in the lab did not provide information about this variable. Indeed, experimenter blindness could only be coded for seven studies conducted in the lab, of which three were our own unpublished studies. A similar problem exists for random assignment. While the random assignment is standard in the field, whether or not group assignment was indeed random is often not explicitly reported in the paper, making it difficult to code.

The three other indicators of study quality (i.e., sample size, manipulation, and scale) were explicitly addressed in the meta-analysis. With respect to sample size, the average number of participants per group was slightly larger in unpublished ($N = 96$) than in published studies ($N = 85$). With respect to the manipulation, the best validated method to manipulate belief in free will is the Crick text. Of the 55% published and 45% unpublished studies using a text to manipulate beliefs in free will, 59% of the published and 91% of the unpublished studies used the Crick text. Finally, with respect to the scale, 53% of the published and 74% of the unpublished studies used a validated scale. Hence, a deeper analysis of the study quality indicates that, if anything, unpublished studies were of higher quality than published studies. More generally, methodological differences between published and unpublished studies in the current meta-analysis were small. This was also confirmed by our analysis of moderator correlations (see Table 4), which found no evidence for substantial correlations between publication status and any of the coded variables.

Third, although no correlations were observed with publication status, large correlations were observed for other moderators. Such confounding of moderators is inevitable in meta-analytic research (Lipsey & Wilson, 2001) and was addressed by performing control analyses analyzing confounded moderators together. Whereas most effects remained even after controlling for these confounds, truly controlling for moderator overlap is difficult and these findings should hence be tested more directly in future empirical work.

Finally, meta-analytical moderator analyses are known to often be underpowered (Hempel et al., 2013). It is therefore possible that some of the nonsignificant moderator effects found here could be explained by a lack of power. However, it is worth noting that at least for the meta-analysis on beliefs in free will and determinism, most analyses included a relatively large number of studies (see Tables 2 and 3). While we are not aware of research that has systematically assessed the statistical power of RVE meta-analyses under different conditions, previous simulation work using regular meta-analytical approaches suggests that given the parameters of the current meta-analysis, many of the moderator analyses were, in fact, well powered to detect even fairly small effect sizes (Hempel et al., 2013).

How to Move Forward? Possible Steps and Recommendations for Future Research

Our results offer several promising routes for future research. First, future research should continue investigating the exact underlying mechanisms of free will belief manipulations to further increase the understanding of these manipulations. For example, recent research suggests that it is important to consider not just beliefs but also attitudes toward free will (Cracco et al., 2020). From this perspective, a potential avenue for future research could be to investigate whether attitudes toward free will alter the effect of anti–free will manipulations.

Second, to better test the downstream consequences of free will belief manipulations, researchers should aim to develop manipulations that (a) more specifically manipulate belief in free will and determinism, but not other factors and (b) lead to larger effect sizes, for example, by using dual-approach manipulations where participants read and repeat the presented messages. Implementing these changes may allow testing whether belief in free will manipulations has a meaningful societal relevance.

Third, future research could also investigate the somewhat surprising results obtained in this meta-analysis. For example, an interesting finding is that anti–free will manipulations appeared to have stronger effects in the United States than in Europe. A potential explanation for this effect might be that individual responsibility and personal agency is more central to U.S. culture than to European culture. Indeed, although previous investigations suggest that irrespective of culture most people believe in free will (Sarkissian et al., 2010; Wisniewski et al., 2019), there is extensive literature showing cross-cultural differences in how individuals construe agency and choice (e.g., Furnham et al., 1994; Markus & Kitayama, 2003). In line with a recent call for cross-cultural replications in social psychology (Genschow et al., 2021), an interesting avenue for future research could therefore be to explore the influence of the manipulation in different continents and cultures. Such research could test not only whether there is a different influence of free will belief manipulations on the beliefs themselves but also whether cultural differences moderate the degree to which these manipulations have downstream consequences.

Fourth, besides cultural differences and the moderators investigated in our meta-analysis, there might be other factors that moderate the influence of free will belief manipulations on beliefs, attitudes, behavior, and cognition. Indeed, there is a discussion regarding the degree to which hidden moderators and high context sensitivity influence psychological effects and account for the success of replications (e.g., Stroebe & Strack, 2014; Van Bavel et al., 2016). Future research should investigate hidden moderators and test how context-dependent free will belief manipulations are.

Fifth, an interesting question that we did not investigate in this meta-analysis is the extent to which the dependent measures investigated here can be influenced by other variables. For example, it could be that psychological variables related to free will such as locus of control (Rotter, 1966) or self-efficacy (Bandura, 1977), for example, have stronger downstream consequences. Directly comparing the effects of these different variables would be interesting, because there are subtle differences between them that could tell us exactly what determines the studied behavior. For example, whereas belief in free will can be defined as an attitude, locus of control is a personality dimension (Waldman et al., 1983). Hence, if the locus of control is associated more strongly with cheating than belief in free will, this could indicate that internal attribution traits are more important for cheating than internal attribution attitudes. While such an analysis goes beyond the scope of the present meta-analysis, future research could investigate whether free will belief manipulations relate to different behaviors, attitudes, and thoughts than manipulations focusing on other psychological variables.

Finally, although we did not find sufficient support for the hypothesis that free will belief manipulations have downstream consequences, it might still be that on a correlational level, free will beliefs relate to individuals' behavior, cognition, and attitudes. Future research could, thus, (a) test to which degree previous findings reported in the literature can be replicated on a correlational level and (b) investigate whether other relevant societal factors are modulated by individuals' belief in free will.

Summary

The present meta-analysis finds support for the idea that beliefs related to free will can be experimentally manipulated. These effects are stronger when participants are presented with a combination of texts and statements, when they have to report on the content of the manipulation, and when belief in free will is measured directly after the manipulation. Moreover, beliefs related to free will can be experimentally increased as well as decreased. However, the used manipulations produce rather weak effects and are rather unspecific in the sense that beliefs and concepts other than belief in free will are also affected. Although we find support for the idea that individuals' belief in free will can be experimentally manipulated, the meta-analysis did not find evidence for the idea that these manipulations have meaningful downstream consequences. These findings call into question prior claims of a causal relationship between belief in free will and attitudes, behavior, and cognition (a number of which were made by authors of this article). Further research is warranted to determine whether free will beliefs are related to attitudes, behavior, and cognition on an interindividual (correlational) level and might have downstream consequences when stronger and more specific manipulations are used.

Authors' Note

Coding manual, data, and analysis scripts are open accessible at the Open Science Framework (OSF; https://bit.ly/2L69prl)

We did not apply for ethical approval because we did not conduct a study on humans, but rather reviewed existing literature in a meta-analysis.

Author Contributions

Oliver Genschow developed the research idea; all co-authors gave feedback on the basic research idea; Oliver Genschow conducted the literature research; Oliver Genschow, Emiel Cracco, and John Protzko collected unpublished data; Oliver Genschow, Emiel Cracco, and Jana Schneider screened and coded the data; Emiel Cracco conducted the analyses; Oliver Genschow & Emiel Cracco drafted the paper, all authors gave feedback on the drafted manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Oliver Genschow https://orcid.org/0000-0001-6322-4392
Emiel Cracco https://orcid.org/0000-0003-4043-5992
Jana Schneider https://orcid.org/0000-0003-3531-1696
John Protzko https://orcid.org/0000-0001-5710-8635

Notes

1. Based on this approach we had to exclude only two published and two unpublished experiments.
2. Note that robust variance estimation (RVE) uses a simplistic method to estimate I^2 and that this should therefore be seen as a rough indicator of heterogeneity rather than as a precise estimate (Tanner-Smith et al., 2016).
3. Note that after data collection had already been completed, Many Labs 5 was published, containing five replications of Vohs and Schooler's (2008) Experiment 1 (Buttrick et al., 2020). While these five experiments are not included in the meta-analysis, we did check whether adding them changed the results. For antisocial behavior, the estimated effect size with these five experiments included was $g = 0.29$ ($m = 28$, $k = 38$, $p = .004$). For cheating, it was $g = 0.27$ ($m = 13$, $k = 14$, $p = .082$). In both cases, a hierarchical Egger test found evidence for small-study bias (both $p = .004$) and both the precision-effect test (PET) and the precision-effect estimate with standard errors (PEESE) indicated that the corrected effect size was $g = 0$. In other words, the antisocial behavior and cheating meta-analyses yielded very similar results with and without these five experiments.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Akin, M., Braun, I., Daenhardt, K., Fischer, S., Kollmann, O., Müller, F., Pistillo, E., Schmidt, K., Stritzel, C., Ueffing, D., Weber, A., Werkle, A., & Wilke, J. (2016). *Macht der Glaube an Willensfreiheit dankbarer?* [Does the belief in free will increase gratefulness?] [Unpublished thesis].

*Alquist, J. L., Ainsworth, S. E., & Baumeister, R. F. (2013). Determined to conform: Disbelief in free will increases conformity. *Journal of Experimental Social Psychology*, *49*, 80–86.

*Alquist, J. L., Ainsworth, S. E., Baumeister, R. F., Daly, M., & Stillman, T. F. (2015). The making of might-have-beens: Effects of free will belief on counterfactual thinking. *Personality and Social Psychology Bulletin*, *41*, 268–283.

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2012). Assessment of study quality for systematic reviews: A comparison of the Cochrane collaboration risk of bias tool and the effective public health practice project quality assessment tool: Methodological research. *Journal of Evaluation in Clinical Practice*, *18*, 12–18.

*Ball, E.-C., Degen, D., Fohrer, A., Grunwald, M., Hintze, S., Höhn, V., Kästel, M., Koziol, P., Matz, M., Papadakis, A., Porth, T., Renner, M., Seyder, M., Willems, S., & Ziske, A. (2016). *Der Glaube an den freien Willen und die Neigung zur Konformität* [The belief in free will and the tendency to conformity.] [Unpublished thesis].

*Balodis, K. (2016). *Influence of belief in free will on intertemporal decisions* [Unpublished bachelor thesis].

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

*Bastart, J., Redersdorff, S., & Martinot, D. (2015). Le libre arbitre au service du jugement émis envers des victimes de sexism [Free will in the service of judging victims of sexism]. *Psychologie Française*, *60*, 223–236.

*Baumeister, R. F., Masicampo, E., & DeWall, C. N. (2009). Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin*, *35*, 260–268.

*Becker, L., Kandt, P., Leupelt, A., & Matthies, I. (2018). *Das Ebenbild Gottes oder ein Sklave seiner Triebe?* [Unpublished thesis].

*Blatz, L., Genschow, O., & Crusius, J. (2017). *Untitled data* [Data set].

*Bomke, N., Goldapp, M., & Schlack, A. (2017). *Der Glaube an den freien Willen, Dualismus und Ambivalenz* [Unpublished thesis].

*Braem, S., & Vermeylen, L. (2016). *Untitled data* [Data set].

Brass, M., Furstenberg, A., & Mele, A. R. (2019). Why neuroscience does not disprove free will. *Neuroscience & Biobehavioral Reviews*, *102*, 251–263.

Buttrick, N. R., Aczel, B., Aeschbach, L. F., Bakos, B. E., Brühlmann, F., Claypool, H. M., Hüffmeier, J., Kovacs, M., Schuepfer, K., Szecsi, P., Szuts, A., Szöke, O., Thomae, M., Torka, A., Walker, R. J., & Wood, M. J. (2020). Many labs 5: Registered replication of Vohs and Schooler (2008), experiment 1. *Advances in Methods and Practices in Psychological Science*, *3*(3), 429–438.

Carey, J. M., & Paulhus, D. L. (2013). Worldview implications of believing in free will and/or determinism: Politics, morality, and punitiveness. *Journal of Personality*, *81*, 130–141.

Carpenter, C. J. (2020). Meta-analyzing apples and oranges: How to make applesauce instead of fruit salad. *Human Communication Research*, *46*, 322–333.

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*, 115–144.

Caruso, G. D. (2014). (Un) just deserts: The dark side of moral responsibility. *Southwest Philosophy Review*, *30*, 27–38.

*Caspar, E. A., & Vuillaume, L. (2017). The influence of (dis) belief in free will on immoral behavior. *Frontiers in Psychology, 8*, Article 20.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychological Methods*, *19*, 211–229.

*Ching, B. H.-H., & Xu, J. T. (2018). The effects of gender neuroessentialism on transprejudice: An experimental study. *Sex Roles*, *78*, 228–241.

Chivers, T. (2010, October 12). Neuroscience, free will and determinism: "I'm just a machine." *The Telegraph*. http://www.telegraph.co.uk/news/science/8058541/Neuroscience-free-will-and-determinism-Im-just-a-machine.html

*Clark, C. J. (2019). *Untitled data* [Data set].

*Clark, C. J., Baumeister, R. F., & Ditto, P. H. (2017). Making punishment palatable: Belief in free will alleviates punitive distress. *Consciousness and Cognition*, *51*, 193–211.

Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., Radkova, I., Deschrijver, E., & Brass, M.. (2018). Automatic imitation: A meta-analysis. *Psychological Bulletin*, *5*, 453–500.

Cracco, E., González-García, C., Hussey, I., Braem, S., & Wisniewski, D. (2020). Cultural pressure and biased responding in free will attitudes. *Royal Society Open Science*, *7*, 191824.

*Crescioni, A. W., Baumeister, R. F., Ainsworth, S. E., Ent, M., & Lambert, N. M. (2016). Subjective correlates and consequences of belief in free will. *Philosophical Psychology*, *29*, 41–63.

Crick, F. (1994). *The astonishing hypothesis: The science search for the soul*. Touchstone.

*Crone, D. L., & Levy, N. L. (2019). Are free will believers nicer people?(Four studies suggest not). *Social Psychological and Personality Science*, *10*, 612–619.

*Darrall, A. (2017). *Freiheit! Für unsere Handlungen, aber nicht für unseren Willen. Konsequenzen einer philosophischen Theorie zum freien Willen* [Freedom! For our actions, but not for our will. Consequences of a philosophical theory of free will.] [Unpublished master's thesis].

Dennett, D. C. (2015). *Elbow room: The varieties of free will worth wanting*:. MIT Press.

*De Souter, L. (2017). *The influence of belief in free will on pseudo-random behavior* [Unpublished master thesis].

*Diel, K., & Genschow, O. (2018). *Untitled data* [Data set].

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLOS ONE, 7*, Article e29081.

Dubljević, V. (2013). Autonomy in neuroethics: Political and not metaphysical. *AJOB Neuroscience*, *4*, 44–51.

Eben, C., Zhang, C., Cracco, E., Brass, M., Billieux, J., & Verbruggen, F. (2020). Are post-error adjustments influenced by beliefs in free will? A failure to directly replicate Rigoni, Wilquin, Brass and Burle, 2013. *Royal Society Open Science*, *7*, 200664.

*Evans, J. (2013). The moral psychology of determinism. *Philosophical Psychology*, *26*, 639–661.

Ewusi-Boisvert, E., & Racine, E. (2018). A critical review of methodologies and results in recent research on belief in free will. *Neuroethics*, *11*, 1–14.

Feldman, G., Farh, J.-L., & Wong, K. F. E. (2018). Agency beliefs over time and across cultures: Free will beliefs predict higher job satisfaction. *Personality and Social Psychology Bulletin*, *44*, 304–317.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, *63*, 665–694.

Fisher, Z., & Tipton, E. (2015). *Robumeta: An R-package for robust variance estimation in meta-analysis*. arXiv:*1503*.02220.

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B*, *367*, 2213–2223.

Furnham, A., Kirkcaldy, B. D., & Lynn, R. (1994). National attitudes to competitiveness, money, and work among young people: First, second, and third world differences. *Human Relations*, *47*, 119–132.

*Genschow, O. (2019). *Untitled data* [Data set].

Genschow, O., Hawickhorst, H., Rigoni, D., Aschermann, E., & Brass, M. (2020). Professional judges' disbelief in free will and punishment. *Social Psychological and Personality Science*, *12*, 357–362.

Genschow, O., & Lange, J. (in press). Belief in free will is related to internal attribution in self-perception. *Social Psychological and Personality Science*. https://doi.org/10.1177/19485506211057711

*Genschow, O., Rigoni, D., & Brass, M. (2014). *Untitled data* [Data set].

*Genschow, O., Rigoni, D., & Brass, M. (2016a). *Untitled data* [Data set].
*Genschow, O., Rigoni, D., & Brass, M. (2016b). *Untitled data* [Data set].
*Genschow, O., Rigoni, D., & Brass, M. (2017a). Belief in free will affects causal attributions when judging others' behavior. *Proceedings of the National Academy of Sciences*, *114*, 10071–10076.
*Genschow, O., Rigoni, D., & Brass, M. (2017b). *Untitled data* [Data set].
Genschow, O., Rigoni, D., & Brass, M. (2019a). The hand of god or the hand of Maradona? Believing in free will increases perceived intentionality of others' behavior. *Consciousness and Cognition*, *70*, 80–87.
*Genschow, O., Rigoni, D., & Brass, M. (2019b). *Untitled data* [Data set].
Genschow, O., & Vehlow, B. (2021). Free to blame? The role of free will belief in victim blaming. *Consciousness and Cognition*, *88*, 103074.
Genschow, O., Westfal, M., Crusius, J., Bartosch, L., Feikes, K. I., Pallasch, N., & Wozniak, M. (2021). Does social psychology persist over half a century? A direct replication of Cialdini et al.'s (1975) classic door-in-the-face technique. *Journal of Personality and Social Psychology*, *120*, e1–e7.
Giner-Sorolla, R., Embley, J., & Johnson, L. (2016). *Replication of Vohs & Schooler (2008, PS, Study 1)*. https://osf.io/i29mh/
*Goodyear, K., Lee, M. R., O'Hara, M., Chernyak, S., Walter, H., Parasuraman, R., & Krueger, F. (2016). Oxytocin influences intuitions about the relationship between belief in free will and moral responsibility. *Social Neuroscience*, *11*, 88–96.
Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*, 1775–1785.
Griffin, A. (2016, April 30). Free will could all be an illusion, scientists suggest after study shows choice may just be brain tricking itself. *Independent*. http://www.independent.co.uk/news/science/free-will-could-all-be-an-illusion-scientists-suggest-after-study-that-shows-choice-could-just-be-a7008181.html
Hallett, M. (2007). Volitional control of movement: The physiology of free will. *Clinical Neurophysiology*, *118*, 1179–1192.
*Hanisch, A. (2015). *Frei handeln oder frei wollen? Experimentelle Manipulation des Glaubens an Willens- und Handlungsfreiheit* [Acting freely or wanting to act freely? Experimental manipulation of the belief in free will and action.] [Unpublished bachelor thesis].
*Hannikainen, I., Cabral, G., Machery, E., & Struchiner, N. (2017). A deterministic worldview promotes approval of state paternalism. *Journal of Experimental Social Psychology*, *70*, 251–259.
*Harms, J., Liket, K., Protzko, J., & Schölmerich, V. (2017). Free to help? An experiment on free will belief and altruism. *PLOS ONE*, *12*, Article e0173193.
*Harpster, R. (2019). *Untitled data* [Data set].
Harris, S. (2012). *Free will*. Simon and Schuster.
Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65.
Hempel, S., Miles, J. N., Booth, M. J., Wang, Z., Morton, S. C., & Shekelle, P. G. (2013). Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews*, *2*, 107.
*Highhouse, S., & Rada, T. B. (2015). Different worldviews explain perceived effectiveness of different employment tests. *International Journal of Selection and Assessment*, *23*, 109–119.
*Jettinghoff, W. (2017). *Untitled data* [Data set].
Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, *88*, 686–702.
*Lantian, A., & Badea, C. (2019). *Untitled data* [Data set].
Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. 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.
Libet, B., Wright, E. W., & Gleason, C. A. (1993). Readiness-potentials preceding unrestricted' spontaneous' vs. pre-planned voluntary acts. In B. Libet (Ed.), *Neurophysiology of consciousness* (pp. 229–242). Birkhäuser.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE.
*Lynn, M., Van Dessel, P., & Brass, M. (2013). The influence of high-level beliefs on self-regulatory engagement: Evidence from thermal pain stimulation. *Frontiers in Psychology, 4*, Article 614.
*MacKenzie, M. J. (2015). *How does free will belief influence guilt?* [Unpublished manuscript].
*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*, 1423–1434.
Markus, H. R., & Kitayama, S. (2003). Models of agency: Sociocultural diversity in the construction of action. In V. Murphy-Berman & J. J. Berman (Eds.), *Cross-cultural differences in perspectives on the self* (pp. 18–74). University of Nebraska Press.
Martin, N. D., Rigoni, D., & Vohs, K. D. (2017). Free will beliefs predict attitudes toward unethical behavior and criminal punishment: A global analysis. *Proceedings of the National Academy of Sciences*, *114*, 7325–7330.
*Melcher, W. (2016). *Untitled data* [Data set].
*Melcher, W. (2019). *Free will in psychological research: Considerations on methodic procedure and reproducibility of results*. Leuphana University Lüneburg.
*Mercier, B. (2017). *Untitled data* [Data set].
*Mercier, B., Wiwad, D., Piff, P. K., Aknin, L. B., Robinson, A. R., Shariff, A., & Vazire, S. (2020). Does Belief in Free Will Increase Support for Economic Inequality? *Collabra: Psychology*, *6*, 25.
*Monroe, A. E. (2018). *Untitled data* [Data set].
*Monroe, A. E., Brady, G., & Malle, B. F. (2017). This isn't the free will worth looking for general free will beliefs do not influence moral judgments, agent-specific choice ascriptions do. *Social Psychological and Personality Science*, *8*, 191–199.
*Monroe, A. E., & Malle, B. F. (2010a). From uncaused will to conscious choice: The need to study, not speculate about people's folk concept of free will. *Review of Philosophy and Psychology*, *1*, 211–224.
*Monroe, A. E., & Malle, B. F. (2010b). *Untitled data* [Data set].
*Moynihan, A. B., & Igou, E. R. (2019). *Untitled data* [Data set].
*Moynihan, A. B., Igou, E. R., & van Tilburg, W. A. (2019). Lost in the crowd: Conformity as escape following disbelief in free will. *European Journal of Social Psychology*, *49*, 503–520.

Murray, D., & Nahmias, E. (2014). Explaining away incompatibilist intuitions. *Philosophy and Phenomenological Research*, *88*, 434–467.

Nadelhoffer, T. (2011). The threat of shrinking agency and free will disillusionism. *Conscious Will and Responsibility: A Tribute to Benjamin Libet*, 173–188.

*Nadelhoffer, T. (2012). *Untitled data* [Data set].

*Nadelhoffer, T., Shepard, J., Crone, D. L., Everett, J. A., Earp, B. D., & Levy, N. (2020). Does encouraging a belief in determinism increase cheating? Reconsidering the value of believing in free will. *Cognition*, *203*, 104342.

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.

Nahmias, E., Morris, S. G., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, *18*, 561–584.

Nahmias, E., Morris, S. G., Nadelhoffer, T., & Turner, J. (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research*, *73*, 28–53.

Nichols, S. (2004). The folk psychology of free will: Fits and starts. *Mind & Language*, *19*, 473–502.

Nichols, S. (2006). Folk intuitions on free will. *Journal of Cognition and Culture*, *6*, 57–86.

Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous*, *41*, 663–685.

*Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716.

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.

Pereboom, D. (2006). *Living without free will*. Cambridge University Press.

Petty, R. E., & Krosnick, J. A. (2014). *Attitude strength: Antecedents and consequences*. Psychology Press.

*Plaks, J. E., & Robinson, J. S. (2015). Construal level and free will beliefs shape perceptions of actors' proximal and distal intent. *Frontiers in Psychology*, *6*, Article 777.

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, *86*, 207–236.

*Protzko, J., & Genschow, O. (2018). *Untitled data* [Data set].

Protzko, J., Ouimette, B., & Schooler, J. (2015). *Believing there is no free will corrupts intuitive cooperation*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2490855

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

Racine, E., Nguyen, V., Saigle, V., & Dubljevic, V. (2017). Media portrayal of a landmark neuroscience experiment on free will. *Science and Engineering Ethics*, *23*, 989–1007.

Rakos, R. F., Laurene, K. R., Skala, S., & Slane, S. (2008). Belief in free will: Measurement and conceptualization innovations. *Behavior and Social Issues*, *17*, 20–40.

*Rangel, U., & Keller, J. (2011). Essentialism goes social: Belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology*, *100*, 1056–1078.

*Rigoni, D. (n.d.). *Untitled data* [Data set].

Rigoni, D., & Brass, M. (2014). From intentions to neurons: Social and neural consequences of disbelieving in free will. *Topoi*, *33*, 5–12.

*Rigoni, D., Kühn, S., Gaudino, G., Sartori, G., & Brass, M. (2012). Reducing self-control by weakening belief in free will. *Consciousness and Cognition*, *21*, 1482–1490.

*Rigoni, D., Kühn, S., Sartori, G., & Brass, M. (2011). Inducing disbelief in free will alters brain correlates of preconscious motor preparation the brain minds whether we believe in free will or not. *Psychological Science*, *22*, 613–618.

*Rigoni, D., Pourtois, G., & Brass, M. (2015). “Why should I care?” Challenging free will attenuates neural reaction to errors. *Social Cognitive and Affective Neuroscience*, *10*, 262–268.

*Rigoni, D., Trager, B. M., Rys, W., Brass, M., Vallacher, R., & Vohs, K. D. (2016). *The importance of being earnest: Valuing honesty attenuates anti-social consequences of disbelieving in free will* [Unpublished manuscript].

*Rigoni, D., Wilquin, H., Brass, M., & Burle, B. (2013). When errors do not matter: Weakening belief in intentional control impairs cognitive reaction to errors. *Cognition*, *127*, 264–269.

Rodgers, M. A., & Pustejovsky, J. E. (2020). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*, *26*, 141–160.

Rose, D., & Nichols, S. (2013). The lesson of bypassing. *Review of Philosophy and Psychology*, *4*, 599–619.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*, 1–28.

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*, 29–41.

Sarkissian, H., Chatterjee, A., De Brigard, F., Knobe, J., Nichols, S., & Sirker, S. (2010). Is belief in free will a cultural universal? *Mind & Language*, *25*, 346–358.

Schooler, J. W. (2010). What science tells us about free will. In R. F. Baumeister, A. R. Mele, & K. D. Vohs (Eds.), *Free will and consciousness: How might they work?* (pp. 191–218). Oxford University Press.

*Schooler, J. W. (2014). *Untitled data* [Data set].

Schooler, J. W., Nadelhoffer, T., Nahmias, E., & Vohs, K. D. (2014). Measuring and manipulating beliefs and behaviors associated with free will. In A. R. Mele (Ed.), *Surrounding free will: Philosophy, psychology, neuroscience* (pp. 72–94). Oxford University Press.

*Schrag, Y. (n.d.). *Untitled data* [Data set].

*Schrag, Y., Tremea, A., Lagger, C., Ohana, N., & Mohr, C. (2016). Pro free will priming enhances “risk-taking” behavior in the Iowa gambling task, but not in the balloon analogue risk task: Two independent priming studies. *PLOS ONE*, *11*, Article e0152297.

*Seminck, N. (2019). *The influence of belief in free will on imitation of inappropriate behavior* [Unpublished master thesis].

*Seto, E. (n.d.). *Untitled data* [Data set].

*Seto, E. (2017). *The interplay of true self-knowledge on belief in free will and moral outcomes* [Unpublished dissertation].

*Seto, E., & Hicks, J. A. (2016). Disassociating the agent from the self: Undermining belief in free will diminishes true self-knowledge. *Social Psychological and Personality Science*, *7*, 726–734.

*Seto, E., Kim, J., & Hicks, J. A. (2020). The illusion of time: Testing the bidirectional relationship between belief in free will and temporal horizons. *Social Psychological and Personality Science*, *12*, 176–185.

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

Shariff, A. F., & Vohs, K. D. (2014). The world without free will. *Scientific American*, *310*, 76–79.

Shepard, J., & Reuter, S. (2012). Neuroscience, choice, and the free will debate. *AJOB Neuroscience*, *3*, 7–11.

Shepherd, J. (2012). Free will and consciousness: Experimental studies. *Consciousness and Cognition*, *21*, 915–927.

*Siller, L. (2019). *Glaube an den freien Willen: Halten sich die Effekte einer experimentellen Manipulation über die Zeit?* [Belief in free will: Do the effects of experimental manipulation hold over time?] [Unpublished master thesis].

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2015). *Official user-guide to the P-curve*. https://www.p-curve.com/guide.pdf

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604.

Smilansky, S. (2000). *Free will and illusion*. Oxford University Press.

Smilansky, S. (2002). Free will, fundamental dualism, and the centrality of illusion. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 489–505). Oxford University Press.

*Snoeck, A. (2019). *The effect of disbelief in free will on prejudice toward Arabic people* [Unpublished master thesis].

Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543–545.

Stanley, T. D. (2017). Limitations of PET-PEESE and other meta-analysis methods. *Social Psychological and Personality Science*, *8*, 581–591.

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*, 60–78.

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 99–110). Wiley.

*Stillman, T. F., & Baumeister, R. F. (2010). Guilty, free, and wise: Determinism and psychopathy diminish learning from negative emotions. *Journal of Experimental Social Psychology*, *46*, 951–960.

Stillman, T. F., Baumeister, R. F., Vohs, K. D., Lambert, N. M., Fincham, F. D., & Brewer, L. E. (2010). Personal philosophy and personnel achievement: Belief in free will predicts better job performance. *Social Psychological and Personality Science*, *1*, 43–50.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, *9*, 59–71.

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*, 13–30.

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology*, *2*, 85–112.

*Thelen, A. (2018). *Der Einfluss des Glaubens an einen freien Willen auf selbstwertdienliche Verzerrungen* [The influence of belief in free will on self-serving biases] [Unpublished master thesis].

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, *20*, 375–393.

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*, 604–634.

*Trager, B. (2016). *Untitled data* [Data set].

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, *113*, 6454–6459.

*Van Dessel, P. (2012). *Untitled data* [Data set].

Van Inwagen, P. (1983). *An essay on free will*. MIT Press.

*Vassena, E. (2019). *Untitled data* [Data set].

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, *6*, 473–482.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the Metafor package. *Journal of Statistical Software*, *36*, 1–48.

Viney, W., Waldman, D. A., & Barchilon, J. (1982). Attitudes toward punishment in relation to beliefs in free will and determinism. *Human Relations*, *35*, 939–949.

*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., Clark, C. J., Lau, S., Vohs, K. D., & Baumeister, R. F. (2017). Ordinary people associate addiction with loss of free will. *Addictive Behaviors Reports*, *5*, 56–66.

Waldman, D. A., Viney, W., Bell, P. A., Bennett, J. B., & Hess, S. (1983). Internal and external locus of control in relation to beliefs in free will and determinism. *Psychological Reports*, *53*, 631–634.

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *132*, 249–268.

Wegner, D. (2002). *The illusion of free will*. MIT Press.

Wisniewski, D., Deutschländer, R., & Haynes, J.-D. (2019). Free will beliefs are better predicted by dualism than determinism beliefs across different cultures. *PLOS ONE*, *14*, Article e0221617.

Wolfe, T. (1997). Sorry, but your soul just died. *Independent*. https://www.independent.co.uk/arts-entertainment/sorry-but-your-soul-just-died-1276509.html

*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*, Article e91572.

*Zwaan, R. A. (n.d.). *Untitled data* [Data set].

*Zwaan, R. A. (2014). The value of believing in free will: A replication Attempt. *Zeistgeist: Psychological Experimentation, Cognition, Language, and Academia*. https://rolfzwaan.blogspot.com/2013/03/the-value-of-believing-in-free-will.html