A Meta-Analytic Review of Moral Licensing

Irene Blanken[1], Niels van de Ven[1], and Marcel Zeelenberg[1]

Abstract
Moral licensing refers to the effect that when people initially behave in a moral way, they are later more likely to display behaviors that are immoral, unethical, or otherwise problematic. We provide a state-of-the-art overview of moral licensing by conducting a meta-analysis of 91 studies (7,397 participants) that compare a licensing condition with a control condition. Based on this analysis, the magnitude of the moral licensing effect is estimated to be a Cohen's *d* of 0.31. We tested potential moderators and found that published studies tend to have larger moral licensing effects than unpublished studies. We found no empirical evidence for other moderators that were theorized to be of importance. The effect size estimate implies that studies require many more participants to draw solid conclusions about moral licensing and its possible moderators.

Moral licensing theory posits that people who initially behave in a moral way can later display behaviors that are immoral, unethical, or otherwise problematic (e.g., Merritt, Effron, & Monin, 2010). For example, someone who has just spent some time volunteering for the local community center might later find it more acceptable to "forget" to report some additional income when filling out the tax return. The idea of moral licensing theory is that the prior good deed provides a "license" that allows one to perform morally questionable behavior later on. To date, various studies on moral licensing have been performed, a substantial subset of which has been published. However, the magnitude of the effect and the specific conditions under which moral licensing is likely to occur remain unclear. Therefore, in this article, we provide a comprehensive overview by performing a meta-analysis across all the available studies on moral licensing.

The mere existence of moral licensing, in which one allows oneself to engage in less moral behavior after a prior good deed, seems to be at odds with a number of well-established psychological findings and theories that stress consistency in behavior. Self-perception theory, for example, asserts that people infer their attitudes from observations of their own behavior, which ultimately affects their subsequent behavior (Bem, 1972). People who perform a good deed would thus see themselves as being a good and moral person which would therefore lead to future moral behavior as well. In addition, balance theory (Heider, 1946), cognitive dissonance theory (Festinger, 1957), the foot-in-the-door effect (Freedman & Fraser, 1966), and the sunk cost effect (Arkes

& Blumer, 1985), all highlight consistency as an essential motivator of human behavior (for reviews, see Abelson et al., 1968; Gawronski & Strack, 2012). Important social psychological theories thus predict that people like to be (and like to appear) good and moral in their actions, and especially so when past moral behavior has just been highlighted, as that makes it important to appear consistent. However, the moral licensing effect has been reported in many domains, including job hiring (Cascio & Plant, 2015; Monin & Miller, 2001), ambiguous racist attitudes (Choi & Crandall, 2008; Effron, Cameron, & Monin, 2009; Effron, Monin, & Miller, 2012; Mann, & Kawakami, 2012), donations to charity (Conway & Peetz, 2012; Sachdeva, Iliev, & Medin, 2009), consumer behavior (Khan & Dhar, 2006), and dishonest behavior (Jordan, Mullen, & Murnighan, 2011; Mazar & Zhong, 2010). As moral licensing has these adverse consequences for such a wide range of behaviors, research on this topic can give important insights into people's motivation and behavior.

It has been suggested that moral licensing can be interpreted as part of a larger moral self-regulation framework. The idea is that internal balancing of moral self-worth and the costs associated with pro-social behavior determine

[1]Tilburg University, The Netherlands

Corresponding Author:
Irene Blanken, Tilburg University, Warandelaan 2, Tilburg, 5037 AB, The Netherlands.
Email: i.blanken@tilburguniversity.edu

whether one will display (im)moral behavior (e.g., Sachdeva et al., 2009). When the moral image of oneself is established, an immoral action is allowed without the fear of losing that moral image (leading to licensing). Conversely, when one appears immoral to others, subsequent positive actions are needed to restore the moral image (leading to compensation or cleansing). However, further research is needed before it can be concluded that a general "balancing" mechanism is responsible for both the licensing and the cleansing effect (cf. Blanken, Van de Ven, Zeelenberg, & Meijers, 2014). The focus of the current meta-analysis will be entirely on the moral licensing effect. A meta-analysis on moral licensing will help in painting a clearer picture on what licensing is and when it occurs, and therefore also forms a solid basis for exploring how the processes of moral licensing and moral cleansing relate.

Definitions of Moral Licensing

On a theoretical level, the process of moral licensing is defined as

> When people are under the threat that their next action might be (or appear to be) morally dubious, individuals can derive confidence from their past moral behavior, such that an impeccable track record increases their propensity to engage in otherwise suspect actions. (Merritt et al., 2010, p. 344)

For our meta-analysis, we build on the more operational definition of moral licensing that Merritt et al. (2010, p. 344) provide: "Past good deeds can liberate individuals to engage in behaviors that are immoral, unethical, or otherwise problematic, behaviors that they would otherwise avoid for fear of feeling or appearing immoral." Moral licensing can be regarded as an example of the broader category of psychological licensing which is "the perception that one's behavioral history, social context, or category membership permit one to legitimately do or say something that otherwise would discredit the self" (Miller & Effron, 2010, p. 116). For example, being a member of a minority group can license one to criticize that group. Thus, moral licensing makes an appeal to someone's past good deeds, whereas the broader category of psychological licensing does not necessarily involve performing good behavior or displaying good intentions. Studies including this particular type of psychological licensing are not included in the meta-analysis, as we are interested in the behavioral consequences of acting morally.

Typical Examples of Moral Licensing

Moral licensing occurs for both individual and social behaviors. It touches on relevant everyday behaviors related to welfare, job hiring, ambiguous racial attitudes, charity donations, consumer purchases, and green consumption. For instance, participants who established their non-prejudiced attitudes by endorsing President Obama (Effron et al., 2009) or through selecting a Black person for a consulting firm job (Monin & Miller, 2001), were subsequently more likely to make pro-White judgments. In a similar vein, participants who previously received feedback that they were close to their goal of being regarded egalitarian toward Black persons seated themselves farther away from a Black confederate than participants who received feedback that they were not progressing toward their goal (Mann & Kawakami, 2012). Furthermore, participants who recalled their own moral actions subsequently displayed fewer pro-social intentions (Conway & Peetz, 2012; Jordan et al., 2011) and cheated more to get a higher payoff (Clot, Grolleau, & Ibanez, 2014; Jordan et al., 2011) than participants who did not recall their moral actions. Moral licensing also plays a role in the area of interpersonal decision making: Previous ethical behavior licensed participants to offer less money to another participant (Cornelissen, Bashshur, Rode, & Le Menestrel, 2013). In addition, participants who disclosed their conflict of interest to the other party, subsequently gave more corrupt advice (Cain, Loewenstein, & Moore, 2005, 2010).

Moral licensing does not seem to be a within-domain phenomenon; the licensed behavior can also take place in a different area. For instance, participants who imagined that they would volunteer to spend time doing community service subsequently preferred hedonic over utilitarian products (Khan & Dhar, 2006). Furthermore, in a virtual shopping paradigm, participants who bought ecofriendly products subsequently offered less money to another person in an ultimatum game and stole more money compared with participants who bought regular products (Mazar & Zhong, 2010).

The Current Meta-Analysis

The first moral licensing study was published in 2001 (Monin & Miller, 2001). Since then, more than 50 separate studies have been published in more than 20 articles. One reason for conducting the current meta-analysis was our observation that the moral licensing effect seems to conflict with one of the most established psychological findings that people want to be and appear consistent in their behavior. A second reason is based on our own research experiences and informal contacts with colleagues in the field, which indicated that it is not always easy to replicate the moral licensing effect (e.g., Blanken et al., 2014), which suggests that the effect may not be robust or subject to moderating factors. Thus, a meta-analysis seems important to (a) attain a good indication of the effect size so we know whether real-life interventions are useful, (b) be able to run power analyses to create well-powered studies (and lower the chance of type II errors), and (c) advance the existing theoretical framework through testing for possible boundary conditions of the effect. The purpose of this article is therefore to provide a state-of-the-art overview of moral licensing. We report a quantitative meta-analytical review through examining published and unpublished

experimental comparisons between a moral licensing and a control condition to determine the magnitude of the moral licensing effect and to identify the potential moderators of the effect.

Moderators of the Moral Licensing Effect

We will explore the conditions under which the moral licensing effect is likely to occur through analyzing theoretically meaningful and methodological moderators. We will examine the following theoretically relevant moderators: The type of moral licensing induction, the behavior measured in the dependent variable, and the domain in which the behaviors take place. We will also examine methodological factors that do not touch upon theoretical distinctions in moral licensing, but are related to the specific aspects of the research design and the current status of the research: article status and control condition.

Moral Licensing Induction: Traits Versus Actions

We will compare the effects of moral licensing inductions related to prior moral traits versus prior moral actions. Conway and Peetz (2012) found that recalling moral actions led to licensing (subsequent less moral behavior), whereas recalling moral traits led to consistency (subsequent moral behavior). They argued that the recall of a moral action signals that progress toward the goal of being moral has been made, and for a subsequent choice between doing the morally right thing versus acting out of self-interest (e.g., whether or not to donate money to someone in need), the goal of being moral becomes less important (because one already made progress toward that goal) and self-interest is thus more likely to win. In contrast, recalling moral traits is thought to activate more abstract moral identity concerns. As the recall activates the general concept of morality, people will subsequently behave more according to moral norms (and thus no moral licensing is expected). This theory predicts that licensing effects should only occur when induced through good actions rather than good traits. Study 3 of Conway and Peetz (2012) provides initial support for this idea. In the current meta-analysis, we have the opportunity to further test this moderator.

Behavior in the Dependent Variable: Actual Versus Hypothetical

We will also investigate whether the moral licensing effect differs depending on whether the dependent variable is actual behavior rather than hypothetical behavior (for instance, donating actual money versus indicating the amount of money one would be willing to donate). Previous research revealed that people want to appear moral while avoiding the cost of being moral (Batson & Thompson, 2001; Dana,

Weber, & Kuang, 2007). That is, people will display good behavior, as long as the costs of good behavior do not override the benefits of self-interested behavior too much. As it is relatively effortless to display hypothetical good behavior (talk is cheap), in these cases there may be lower willingness to display undesirable behaviors that one needs to justify. Thus, one could expect that the moral licensing effect is larger when the dependent behavior consists of actual compared with hypothetical behavior.

Domain: Same Versus Different

We will further investigate whether the size of the moral licensing effect differs depending on whether the good and bad behaviors occur in the same or in a different domain. Miller and Effron (2010) pointed out that good behavior in one domain can not only license people to perform dubious behavior in the same domain, but also in unrelated domains (e.g., Mazar & Zhong, 2010). Theory and empirical work in the field of mental accounting reveals that people use different mental accounts to organize their finances (Thaler, 1985). For instance, if people receive a financial windfall in one domain (e.g., a refund for a delayed flight), they typically more easily spend it on something related (e.g., a luxurious dinner during that trip). If similar effects occur in the realm of moral licensing, one might expect that after doing something good in one domain, people more easily allow themselves to do something more questionable in the same domain later. Therefore, it could be expected that licensing effects are larger (smaller) when the moral and immoral behaviors are measured in the same (different) domain.

Article Status: Published Versus Unpublished Work

We will examine whether the magnitude of the moral licensing effect depends on whether the study is part of a published article or not. It could be expected that the moral licensing effect, like other empirical findings, is larger for studies in published articles as more positive findings than null- or negative-findings tend to be published (Ioannidis, 2005). We will later test for potential publication bias in various ways, but think it is essential to include study status as a control variable when we examine other possible moderators as well.

Control Condition: Neutral Versus Negative

We will further investigate whether the moral licensing effect differs depending on whether the moral licensing condition (in which people are asked to recall prior good behavior) is contrasted with a negative control condition (a recall of prior bad behavior) or with a neutral control condition (a recall of neutral behavior). This is important because the opposite pattern can exist as well: Positive behavior becomes more likely

after recalling previous negative behavior, which is called the moral cleansing effect (Conway & Peetz, 2012; Jordan et al., 2011; Sachdeva et al., 2009; Zhong & Liljenquist, 2006, but see also Blanken et al., 2014). Based on the theory of moral cleansing, one could expect that the moral licensing effect is larger when a negative control condition is used compared with a neutral one.

In the moral licensing literature, a number of other moderators has been proposed. Power was insufficient to include all these possible moderators, because (a) the current number of included studies ($n = 91$) does not allow for too many moderators and (b) many of these hypothesized moderators were only tested in very few studies, which would make comparisons unreliable due to the small sample. For completeness of our review of moral licensing effects we do mention the proposed moderators here, hoping that future research can further explore these possible moderators. Specific moderators that were identified (stronger licensing in conditions displayed in italics) are *free* versus forced choice good behavior (Bradley-Geist, King, Skorinko, Hebl, & McKenna, 2010), *high* versus low rationalizability of cheating (Brown et al., 2011), recalling *recent* versus distant good behavior (Conway & Peetz, 2012), having an *outcome-based* versus a rule-based mind-set (Cornelissen, Bashshur, et al., 2013), focusing on *goal progress* versus goal commitment (Mann & Kawakami, 2012), and having *no external incentive* versus having an external incentive for one's moral behavior (Clot, Grolleau, & Ibanez, 2013a; Khan & Dhar, 2006).

In addition, some studies tested the moderating effect of continuous personality variables that we could not include in the current meta-analysis. These were being *high* versus low in self-monitoring (Cornelissen, Karelaia, & Soyer, 2013), having a *high* versus low score on the modern racism scale (Effron et al., 2009; Effron et al., 2012), and having a *strong* versus weak pro-environmental identity when possibly engaging in licensing in the environmental domain (Meijers, Noordewier, Verlegh, & Smit, 2014).

Taken together, we provide a quantitative meta-analytical review of moral licensing through examining both published and unpublished research. We will investigate the estimated mean effect size of moral licensing and advance the existing theoretical framework on moral licensing through investigating several moderators.

Method

Data Collection

An extensive literature search was conducted to collect data on moral licensing, based on the definition of Merritt et al. (2010, p. 344) "Past good deeds [or good intentions] liberate individuals to engage in behaviors that are immoral, unethical, or otherwise problematic, behaviors that they would otherwise avoid for fear of feeling or appearing less moral." This included

searches in Web of Science and Google Scholar using the following keywords[1]: *(self-) licensing, moral licensing, psychological licensing, moral balancing, moral compensation, moral spillover, self-justification, (moral) credentialing, and (moral) credential(s).* Relevant journals (all RSS feeds from the *European Journal of Social Psychology, Journal of Consumer Psychology, Journal of Consumer Research, Journal of Experimental Social Psychology, Journal of Marketing Research, Journal of Personality and Social Psychology, Personality and Social Psychology Bulletin, Psychological Bulletin and Psychological Science*), conference proceedings, dissertations, and master theses were also checked. Furthermore, we called for relevant studies on moral licensing (both published and unpublished, both successful and unsuccessful) on the Society for Personality and Social Psychology (SPSP) internet forum (October 18th, 2012) as well as via various mailing lists (SPSP, the Society for Judgment and Decision Making [SJDM], and the Dutch Society of Social Psychologists [ASPO]). Finally, we presented a preliminary version of the meta-analysis at the annual SPSP Conference in New Orleans on January 17th, 2013. At this conference, we released another call for data on moral licensing. We considered all the studies we found and received until December 1st, 2014 for inclusion.

Inclusion Criteria

Studies were included if they met two criteria. First, the behavior that was measured had to meet our definition of moral licensing. This entails that the behavior has to take place in a moral domain. Licensing studies on self-regulation (e.g., Chiou, Yang, & Wan, 2011; De Witt-Huberts, Evers, & De Ridder, 2012; Fishbach & Dhar, 2005; Mukhopadhyay & Johar, 2009; Mukhopadhyay, Sengupta, & Ramanathan, 2008) do not fall under the definition of *moral* licensing and were thus not included in the current meta-analysis. In addition, this implies that the independent variable should consist of (intended) good *behavior* or the recall thereof. For instance, we included the studies by Mazar and Zhong (2010) where the independent variable consists of buying eco-friendly products in a virtual shopping paradigm, but we did not include the study by Eskine (2012) where the independent variable consists of participants merely being exposed to organic products, as the latter does not entail actual or hypothetical behavior. If there is no prior good deed (or recall thereof), there can also be no moral licensing according to our definition. Following this definition also implies that the dependent variable should measure actual or hypothetical behavior of the participants. Studies where the dependent variable consisted of evaluative judgments, for instance an evaluation of one's morality level (Jordan et al., 2011, Study 1) or the perception of general undesirable behavior (Effron, Monin, & Miller, 2013, Study 2), were therefore not included.

Second, reported statistics had to be adequate to calculate effect sizes. When important statistical information was lacking, authors were contacted for more information. When

authors did not respond to the initial request, two reminders were sent. Studies that did not meet our inclusion criteria can be found in Online Appendix I.

Dependent Variables

The dependent variables reported in the included studies comprise immoral behavior, such as cheating and stealing (Mazar & Zhong, 2010), or a decrease in moral behavior, such as donating less money to charity (Sachdeva et al., 2009). These behaviors are real, such as stealing money from the experimenter (Mazar & Zhong, 2010), or hypothetical, such as indicating that one would be willing to volunteer (Conway & Peetz, 2012). They were either measured on a continuous scale or as a dichotomous choice between a virtuous and a less virtuous option. All effect sizes were recoded so that positive effect sizes indicated associations between previous moral behavior and a subsequent decrease in moral behavior.

Moderators

In several studies that investigated possible moderators of the moral licensing effect, the authors predicted a moral licensing effect in one condition but no moral licensing in another condition. For instance, Conway and Peetz (2012) predicted that licensing would occur when participants recalled a recent moral act, but not when they recalled a temporally distant moral act. In these cases, our default was that we did not include the conditions where the authors did not expect a moral licensing effect. For the Conway and Peetz example on recalling recent versus distant moral acts, this implies that we only included the condition where participants recalled a recent moral act. Across all the studies that we included, we made two exceptions to this default. First, Bradley-Geist et al. (2010) predicted that participants who freely chose to write about a positive (vs. negative) experience with a member of a minority group would obtain a moral license, whereas participants who were forced to write about a positive (vs. negative) past experience with a member of a minority group would not obtain a moral license. Because in many other studies on moral licensing participants are specifically asked to write about moral behavior or moral traits in the past and thus do not have a choice to write about immoral versus moral behavior (Blanken, Van de Ven, & Zeelenberg, 2012; Blanken et al., 2014; Clot et al., 2013a; Conway & Peetz, 2012; Cornelissen, Bashshur, et al., 2013; Jordan et al., 2011; Sachdeva et al., 2009; Schüler, Lehnhardt, & Huber, 2012; Thomas & Showers, 2012; Young, Chakroff, & Tom, 2012), we decided to include the forced choice conditions from Bradley-Geist et al. (2010). Second, in the Study 3 of Conway and Peetz (2012), the authors predicted that participants who wrote about moral actions would obtain a moral license, whereas participants who wrote about moral traits would not obtain a moral license. We decided to include

both the moral action and the moral trait conditions because we tested for the difference between moral traits versus actions in our between-study moderator section.

For studies where authors predicted that one specific condition could have a larger licensing effect than the other condition, we included both effects. For instance, Monin and Miller (2001) predicted that participants who previously established non-racist credentials were more likely to display preference for a White over a Black person in a hypothetical job hiring task compared to participants who did not establish non-racist credentials. They also predicted that this effect could be larger for participants who completed the independent and dependent variables in front of the same (versus a different) experimenter. For this study, we thus included both the single experimenter audience and the different experimenter audience conditions, because the authors had predicted a licensing effect in both conditions.

Some studies tested the moderating effect of continuous personality variables (Cornelissen, Karelaia, & Soyer, 2013; Effron et al., 2009; Effron et al., 2012; Meijers et al., 2014). For instance, Effron et al. (2009) and Effron et al. (2012) measured whether scores on the modern racism scale moderated participants' preferences for White over Black persons. In these cases, we included the main effect size without distinguishing between participants who scored high versus low on the measured personality variable because (a) other studies on moral licensing do not differentiate between these specific variables and (b) according to general theorizing on moral licensing the general licensing induction should work for all individuals.

We analyzed the effects of between-study moderators, that is, moderators that we theoretically predicted to influence the magnitude of the moral licensing effect (as pointed out in the introduction section) by means of a regression model.

Overview of Analyzed Studies

The data set contains 91 different comparisons between a moral licensing and a control condition, with a total of 7,397 participants, reported in 22 published or forthcoming journal articles and 8 unpublished manuscripts. Table 1 provides an overview of the included studies.

Effect Size Measure

We calculated Cohen's d based on pooled standard deviations for all studies, with a positive d value indicating moral licensing. For calculation of the effect sizes of the continuous dependent variables, we used means and standard deviations. When information on means and standard deviations was lacking, t values were used for the calculation. For calculation of the effect sizes of dichotomous dependent variables, we used the reported χ^2 statistic or calculated the χ^2 statistic manually. For all effect sizes, we applied the small sample bias correction provided by Lipsey and Wilson (2001). The

Table 1. Details of All Studies Included in the Meta-Analysis (Dummy Coding for Moderators in Last Columns).

Author(s)	Year	Study	N	d	s^2	Licensing manipulation	Dependent variable	A	B	C	D	E
1 Blanken, van de Ven, Zeelenberg, and Meijers	2014	1	64	0.08	0.06	Pp wrote about themselves using positive trait words	Donation to charity	0	1	0	1	0
2 Blanken et al.	2014	2	91	0.27	0.04	Pp wrote about themselves using positive words	Cooperative behavior in a hypothetical commons dilemma	0	0	0	1	0
3 Blanken et al.	2014	3	567	−0.05	0.01	Pp wrote about themselves using positive trait words	Donation to charity and cooperative behavior in a hypothetical commons dilemma	0	x	0	1	0
4 Blanken, van de Ven, and Zeelenberg	2012	1	86	−0.08	0.05	Pp recalled having good traits	Participants indicated their willingness to fake illness at work to visit a concert in a hypothetical scenario	0	0	0	0	0
5 Blanken et al.	2012	1	86	0.07	0.05	Pp recalled performing good behavior	Willingness to fake illness at work to visit a concert in a hypothetical scenario	1	0	0	0	0
6 Blanken et al.	2012	2	38	0.05	0.11	Pp indicated whether they are planning to donate their organs after death	Everyday Cooperation Scale	1	0	0	0	0
7 Blanken et al.	2012	3	49	0.49	0.08	Pp recalled having good traits	Hypothetical donation to charity	0	0	0	0	0
8 Blanken et al.	2012	3	47	0.03	0.09	Pp recalled performing good behavior	Hypothetical donation to charity	1	0	0	0	0
9 Blanken et al.	2012	4	64	0.30	0.06	Pp recalled having good traits	Willingness to fake illness at work to visit a concert in a hypothetical scenario	0	0	0	0	0
10 Blanken et al.	2012	4	65	0.54	0.06	Pp recalled having good traits	Willingness to fake illness at work to visit a concert in a hypothetical scenario	0	0	0	0	0
11 Blanken et al.	2012	5	57	−0.11	0.07	Pp recalled helping a friend	Willingness to Volunteer Scale	1	0	0	0	0
12 Blanken et al.	2012	5	53	0.12	0.08	Pp recalled helping a stranger	Willingness to Volunteer Scale	1	0	0	0	0
13 Blanken et al.	2012	6	50	0.61	0.10	Pp recalled having good traits and helping a person who is walking on crutches	Giving "accidentally" overpaid (actual) money back to experimenter (y/n)	x	1	0	0	0
14 Blanken et al.	2012	6	49	0.48	0.10	Pp recalled performing good behaviors and helping a person who is walking on crutches	Giving "accidentally" overpaid (actual) money back to experimenter (y/n)	1	1	0	0	0
15 Blanken et al.	2012	7	48	0.63	0.09	Pp recalled having good traits	Willingness to pay for luxurious over purposeful goods	0	0	0	0	0
16 Blanken et al.	2012	7	51	0.22	0.08	Pp recalled performing good behavior	Willingness to pay for luxurious over purposeful goods	1	0	0	0	0
17 Blanken et al.	2012	8	94	0.86	0.07	Pp recalled having good traits	Everyday Cooperation Scale	0	0	0	0	0
18 Blanken et al.	2012	8	133	0.24	0.03	Pp recalled performing good behavior	Everyday Cooperation Scale	1	0	0	0	0
19 Blanken et al.	2012	9	65	−0.19	0.06	Pp recalled having good traits	Willingness to pay for luxurious over purposeful goods	0	0	0	0	0
20 Blanken et al.	2012	9	61	0.04	0.07	Pp recalled performing good behavior	Willingness to pay for luxurious over purposeful goods	1	0	0	0	0
21 Blanken et al.	2012	10	61	−0.34	0.07	Pp indicated whether they are planning to donate their organs after death + provided their address for a folder on organ donation	Everyday Cooperation Scale	1	0	0	0	0
22 Blanken et al.	2012	11	83	−0.25	0.05	Pp indicated preference for a Black doctor	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	0	0
23 Blanken et al.	2012	12	57	−0.67	0.08	Pp disagreed with racist statement	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	0	0
24 Bradley-Geist, King, Skorinko, Hebl, and McKenna	2010	1	38	−0.17	0.11	Pp were asked to write about a positive experience with a Black individual	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0
25 Bradley-Geist et al.	2010	1	35	0.66	0.13	Pp were given the choice to write about a positive or negative experience with a Black individual. Participants who chose to write about a positive experience were included	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0
26 Bradley-Geist et al.	2010	2	44	0.17	0.09	Pp were asked to write about a positive experience with a Hispanic individual	Willingness to hire White person for a job (over a Hispanic person) in a scenario[a] + tolerance for prejudice scale	1	0	1	1	0
27 Bradley-Geist et al.	2010	2	42	0.47	0.10	Pp were given the choice to write about a positive or negative experience with a Hispanic individual. Participants who chose to write about a positive experience were included	Willingness to hire White person for a job (over a Hispanic person) in a scenario + tolerance for prejudice scale	1	0	1	1	0
28 Bradley-Geist et al.	2010	4	56	−0.70	0.08	Pp were asked to write about a positive experience with a Black friend	Willingness to hire White person for a job (over a Hispanic person) in a scenario[a] + tolerance for prejudice scale	1	0	1	1	1

(continued)

545

Table 1. (continued)

Author(s)	Year	Study	N	d	s²	Licensing manipulation	Dependent variable	A	B	C	D	E
29 Brown et al.	2011	1	97	0.41	0.04	Pp read four moral dilemmas and rated how likely they would be to behave in a pro-social fashion if they were to encounter such a situation	The extent to which participants cheat on a math task	1	1	0	1	0
30 Cain, Loewenstein, and Moore	2005	1	43	0.58	0.10	Pp disclosed their conflict of interest to other pp	Giving worse advice to increase own gains in a hypothetical conflict of interest	1	0	1	1	0
31 Cain, Loewenstein, and Moore	2010	1	348	−0.15	0.01	Pp disclosed their conflict of interest to other pp	Giving worse advice to increase own gains in a hypothetical conflict of interest	1	0	1	1	0
32 Cain et al.	2010	3	49	0.04	0.08	Pp disclosed their conflict of interest to other pp	Giving worse advice to increase own payoff	1	1	1	1	0
33 Cain et al.	2010	3	77	0.40	0.05	Pp disclosed their conflict of interest to other pp	Giving worse advice to increase own payoff	1	1	1	1	0
34 Cascio and Plant	2015	1	86	0.54	0.05	Pp were asked whether they were interested in taking part in a fundraiser (skip a meal for Red Cross)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	0	1	0
35 Cascio and Plant	2015	2	51	0.66	0.08	Pp were asked whether they were interested in taking part in a fundraiser (skip a meal for Red Cross)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	0	1	0
36 Cascio and Plant	2015	3	54	1.26	0.09	Pp were asked whether they would be willing to donate blood	Explicit racism–Attitudes Toward Black Scale (Example item: "I would rather not have Blacks live in the same apartment building I live in")	1	0	0	1	0
37 Cascio and Plant	2015	4	74	0.60	0.06	Pp were asked whether they were interested in taking part in a fundraiser (skip a meal for Red Cross)	Indicating stereotypes against Blacks	1	0	0	1	0
38 Choi, Crandall, and La	2014	2	116	0.44	0.04	Pp first evaluated a high-quality ad with a Black model and were next given a chance to express positive attitude toward the ad	Evaluation of a second (target) ad including a Black model on attractiveness	1	0	1	1	0
39 Choi et al.	2014	3	60	0.54	0.07	Pp first evaluated a high-quality ad with a Black model and were next given a chance to express positive attitude toward the ad	Evaluation of a second (target) ad including a Black model on attractiveness	1	0	1	1	0
40 Clot, Grolleau, and Ibanez	2014	1	100	0.35	0.04	Pp imagined that they previously performed a good deed	Cheating for higher payoff	1	1	0	1	0
41 Clot, Grolleau, and Ibanez	2013b	1	367	0.23	0.01	Pp imagined performing voluntary work	Pp indicated the extent to which they would take actual money out of this public funds	1	0	0	0	0
42 Clot, Grolleau, and Ibanez	2013a	1	192	0.30	0.02	Pp imagined engaging in a pro-environmental activity	Pp could allocate an amount (30€) between themselves and an environmental charity (either World Wildlife Fund or Greenpeace)	0	1	1	1	0
43 Conway and Peetz	2012	1	51	0.33	0.08	Pp recalled moral behavior they performed during the previous week (recently)	Willingness to Volunteer Scale	1	0	0	1	1
44 Conway and Peetz	2012	2	39	0.58	0.11	Pp described moral behavior in the recent past (1 week ago)	Willingness to Volunteer scale	1	0	0	1	1
45 Conway and Peetz	2012	3	65	0.00	0.06	Pp were asked to visualize having certain moral characteristics, and wrote about what having these traits would mean for their personality	Donation to charity; pp entered a draw and they could indicate how much of the price money would be donated to charity	0	1	0	1	1
46 Conway and Peetz	2012	3	65	0.79	0.07	Pp were asked to visualize performing moral behaviors, and wrote about how they would perform these behaviors	Donation to charity; pp entered a draw and they could indicate how much of the price money would be donated to charity	1	0	1	1	1
47 Cornelissen, Bashshur, Rode, and Le Menestrel	2013	1	48	0.59	0.09	Pp remembered an episode in the past where they did something ethical	Amount of money offered in actual dictator game	1	1	1	1	1
48 Cornelissen, Bashshur, et al.	2013	2	40	0.84	0.11	Pp remembered an episode in the past where they did something ethical	Amount of money offered in actual dictator game	1	0	1	1	1
49 Cornelissen, Bashshur, et al.	2013	3	50	0.53	0.08	Pp remembered an episode in the past where they did something ethical	Cheating for higher payoff	1	0	1	1	1
50 Cornelissen, Karelaia, and Soyer	2013	1	70	0.57	0.06	Pp could indicate whether they supported UNICEF	Providing help for organization by investing time to develop slogans	1	1	1	0	1
51 Cornelissen et al.	2013	2	92	0.36	0.04	Pp indicated whether they supported Oxfam	Contribute part of participation fee to Oxfam	1	1	1	0	1
52 Effron, Cameron, and Monin	2009	1	84	0.44	0.05	Pp indicated that they would vote for President Obama (displaying preference for Black over White president)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0

(continued)

Table 1. (continued)

Author(s)	Year	Study	N	d	s²	Licensing manipulation	Dependent variable	A	B	C	D	E
53 Effron et al.	2009	2	40	0.35	0.10	Pp indicated that they would vote for President Obama (displaying preference for Black over White president)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0
54 Effron et al.	2009	3	50	0.16	0.08	Pp indicated that they would vote for President Obama (displaying preference for Black over White president)	Willingness to allocate funds to a group of White (over Black) people in a scenario	1	0	1	1	0
55 Effron, Monin, and Miller	2012	3	157	0.25	0.03	Pp displayed non-racist behavior while having a racist alternative to their behavior (accused White over Black guy for crime)	Willingness to hire White person for a job (over a Black person) in a scenario and willingness to allocate funds to a group of White (over Black) people	1	0	1	1	0
56 Jordan, Mullen, and Murnighan	2011	2	68	0.42	0.06	Pp recalled a situation in which they helped other people	Indicate willingness to donate money to charity, donate blood, volunteer	1	0	0	1	0
57 Jordan et al.	2011	3	76	0.98	0.06	Pp recalled a situation in which they helped other people	The extent to which participants cheat on math task	1	1	0	1	0
58 Khan and Dhar	2006	1	108	0.62	0.04	Pp imagined that they had volunteered to perform community service	Preferring a vicious over a less vicious item	1	0	0	1	0
59 Khan and Dhar	2006	2a	94	0.60	0.05	Pp imagined donating money to charity	Preferring a vicious over a less vicious item	1	0	0	1	0
60 Khan and Dhar	2006	2b	80	0.56	0.06	Pp imagined donating money to charity	Preferring a vicious over a less vicious item	1	0	0	1	0
61 Khan and Dhar	2006	3	80	0.39	0.05	Pp imagined helping a foreign student	Pp received $2 and could indicate whether and how much money they would donate to charity	1	0	0	1	0
62 Khan and Dhar	2006	4	80	0.46	0.06	Pp imagined that they had volunteered to perform community service	Preferring a vicious over a less vicious item	1	0	0	1	0
63 Khan and Dhar	2006	5	66	0.56	0.06	Pp imagined that they had volunteered to perform community service	Preferring a vicious over a less vicious item	1	0	1	1	0
64 Leonard	2012	1	32	−0.29	0.13	Pp imagined refusing an offer to fake volunteering less hours than required	Indicate how much money they would keep from a wallet they found in a hypothetical scenario	1	0	0	0	1
65 Leonard	2012	1	92	−0.04	0.04	Pp imagined refusing an offer to fake volunteering less hours than required	Indicate how much money they would keep from a wallet they found in a hypothetical scenario	1	0	0	0	1
66 Leonard	2012	2	43	−0.02	0.09	Pp imagined refusing an offer to fake volunteering less hours than required	Indicate how much money they would keep from a wallet they found in a hypothetical scenario	1	0	0	0	1
67 Leonard	2012	3	64	0.01	0.06	Pp imagined that they found a wallet on the floor and decided not to take any of the money	Pp indicated to what extent they would refuse an offer to fake volunteering less hours than required	1	0	0	1	1
68 Mann and Kawakami	2012	1	44	0.73	0.10	Pp received feedback that they were drawing closer to their goal (try to have positive evaluations of Black people whenever they were presented with an image of Blacks)	Interpersonal closeness task—seating distance toward Black person	1	1	1	1	1
69 Mann and Kawakami	2012	2	94	0.55	0.04	Pp received feedback that they were drawing closer to their goal (try to have positive evaluations of Black people whenever they were presented with an image of Blacks)	Racial attitudes	1	0	1	1	1
70 Mann and Kawakami	2012	3	30	0.86	0.15	Pp received visual feedback during a task that indicated that they were becoming more positive toward Blacks based on physiological responses measured with a LifeShirt	Interpersonal closeness task—seating distance toward Black person	1	1	1	1	1
71 Mazar and Zhong	2010	2	76	0.29	0.05	Pp were assigned to an online store with more green (ecofriendly) than conventional products. Pp made purchases in this store	Amount of money offered in an actual dictator game	1	1	0	1	0
72 Mazar and Zhong	2010	3	81	0.53	0.05	Pp were assigned to an online store with more green (ecofriendly) than conventional products. Pp made purchases in this store	Lying about performance to gain extra money for the experiment and stealing money from envelope	1	1	0	1	0
73 Meijers, Noordewier, Verlegh, and Smit	2014	1	40	−0.04	0.10	Pp imagined buying environmentally friendly sneakers	Sustainable intentions measured through the Minton and Rose Behavioral Intentions Scale	1	0	1	0	0
74 Meijers et al.	2014	1	40	0.83	0.11	Pp imagined buying environmentally friendly sneakers.	Sustainable intentions measured through the Minton and Rose Behavioral Intentions Scale	1	0	1	0	0
75 Meijers et al.	2014	2	88	0.60	0.05	Pp entered a webshop with heavily environmental apparel and were asked to compose an outfit.	Environmental Concern and pro-environmental intentions	1	0	1	0	0
76 Meijers et al.	2014	2	88	−0.24	0.05	Pp entered a webshop with heavily environmental apparel and were asked to compose an outfit	Environmental Concern and pro-environmental intentions	1	0	1	0	0
77 Merritt et al.	2012	2	70	0.48	0.06	Pp had a chance to establish credentials by identifying ambiguous behaviors as racist	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0

(continued)

547

Table 1. (continued)

Author(s)	Year	Study	N	d	s²	Licensing manipulation	Dependent variable	A	B	C	D	E
78 Monin and Miller	2001	1	140	0.26	0.03	Pp indicated whether they considered each of five negative statements about women to be right or wrong, for instance: "Most women are not smart"	Willingness to hire a woman for a job (over a man) in a scenario	1	0	1	1	0
79 Monin and Miller	2001	2	110	0.71	0.04	Pp hired a women (showing that they are not sexist) or a Black guy (showing that they are not racist)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0
80 Monin and Miller	2001	3	21	0.91	0.21	Pp hired a Black guy (showing that they are not racist)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0
81 Monin and Miller	2001	3	19	0.89	0.23	Pp hired a Black guy (showing that they are not racist)	Willingness to hire White person for a job (over a Black person) in a scenario	1	0	1	1	0
82 Sachdeva, Iliev, and Medin	2009	1	29	0.60	0.14	Pp wrote about themselves using positive words	Donation to charity	0	1	0	1	0
83 Sachdeva et al.	2009	2	18	1.10	0.26	Pp wrote about themselves using positive words	Donation to charity	0	1	0	1	1
84 Sachdeva et al.	2009	3	31	0.57	0.13	Pp wrote about themselves using positive words	Cooperative behavior in a hypothetical commons dilemma	0	0	0	1	0
85 Schüler, Lehnhardt, and Huber	2012	1	62	−0.42	0.07	Pp recalled a situation in which they helped other people	Dictator game where pp could earn points for lottery	1	1	0	0	0
86 Spektor	2014	1	278	0.00	0.01	Pp remembered and wrote down a concrete situation within the last 7 days in which they felt righteous or honorable (moral condition)	Pp were told that they automatically participated in a €100 draw and could donate an amount of their choice to one or more charities in case they won	1	1	0	0	1
87 Susewind and Hoelzl	2014	1	84	0.38	0.05	Pp focused on progress toward sustainable goals	Pp indicated their pro-social intentions	1	1	0	1	0
88 Susewind and Hoelzl	2014	2	62	0.44	0.07	Pp brainstormed on a topic that clearly benefited others and made progress on their goals	Pp divided €6 between themselves and another participant of the study that was randomly assigned to them, knowing that the other participant has to accept their decision	1	1	0	1	1
89 Thomas and Showers	2012	2	76	−0.30	0.05	Pp recalled a time when they did something especially moral or ethical	The extent to which participants cheat on math task	1	1	0	0	1
90 Young, Chakroff, and Tom	2012	1	66	0.41	0.06	Pps were asked to describe 5 good deeds	Donation amount to charity - hypothetical amount, actual y/n and actual amount	1	x	0	1	0
91 Zhong, Ku, Lount, and Murnighan	2009	2	68	0.29	0.06	Pp were asked to imagine that they would display ethical behavior in a specific situation	Willingness to display ethical behavior in hypothetical situation	1	0	1	1	0

Note. N = number of participants included in study; d = Cohen's d effect size; s^2 = within-study variance; A = licensing induction: traits (0) versus actions (1); B = behavior measured in the dependent variable: actual (1) versus hypothetical (0); C = domain: same (1) versus different (0); D = article status: published (1) versus unpublished (0); E = control condition: neutral (0) versus negative (1).

[a]Data on this measure were not included as there were insufficient statistics to calculate a Cohen's d effect size.

548

equations that we used to calculate the effect sizes can be found in Online Appendix II.

When the moral licensing condition (e.g., recall positive behavior) was compared with both a neutral (e.g., recall neutral behavior) and a negative (e.g., recall bad behavior) control condition, we always report the comparison between the moral licensing condition and the neutral control condition. Differences between comparisons with negative conditions versus comparisons with control conditions are analyzed in the between-moderator section. Several studies reported the effect of one independent variable on multiple dependent variables. For instance, Jordan et al. (2011) measured the effect of recalling (un)ethical behavior on (a) allowing an answer to a math task to appear on the screen, (b) whether the participant used that provided answer, and (c) the number of answers completed before participants started to cheat. To avoid statistical dependencies, in these situations we combined the different effects into a single effect size by averaging the multiple effect sizes (Hedges & Olkin, 1985; Marascuillo, Busk, & Serlin, 1988; Rosenthal & Rubin, 1986). Some studies included two moral licensing conditions. For instance, Blanken et al. (2012) measured the effects of displaying prior good behavior and of displaying having been a good person on willingness to help and compared both conditions with a similar control condition. In these cases, we report a separate effect size for each moral licensing condition. This also enabled us to code these conditions independently for the moderator section.

Random-Effects Model

We examined the overall effect size of the moral licensing effect using a random-effects model, as there was no a priori reason to assume that the true effect size is exactly the same across all studies (Hedges & Vevea, 1998). The model treats the heterogeneity as purely random; $\theta_i = \mu + u_i$, where $u_i \sim N(0, \tau^2)$ and τ^2 is the residual heterogeneity estimated with the DerSimonian–Laird estimator (DerSimonian & Laird, 1986). The average true effect μ is calculated via weighted least squares with weights equal to $w_i = 1 / (v_i + \hat{\tau}^2)$, where $\hat{\tau}^2$ is the estimate of τ^2 and v_i is the sample variance (Viechtbauer, 2010).

Results

Mean Effect of Moral Licensing

The random-effects meta-analysis ($N = 91$; 57 published and 34 unpublished studies) produced a mean effect size of moral licensing of $d = 0.31$ (95% CI [0.23, 0.38]).[2] The null hypothesis H0: $\mu = 0$ was rejected ($Z = 8.24$, $p < .001$), showing that there is a significant moral licensing effect across the studies we analyzed. The between-study variance is $\tau^2 = .06$ (95% CI [.03, .11]) with $I^2 = 54.58\%$ of the total variation due to heterogeneity among true effects; in other words, 54.58% of the

variability may be attributable to systematic between-study differences (Higgins & Thompson, 2002). Also, the test for heterogeneity is significant $Q(df = 90) = 198.17$, $p < .001$, implying that other possible moderators are influencing the magnitude of the moral licensing effect. The forest plot of the meta-analysis is depicted in Figure 1. To conclude the main analysis, the moral licensing effect is small to medium in effect size. In addition, there is substantial variation of the effect size between studies. In the next section, we discuss possible publication bias and whether specific between-study moderating variables are responsible for this variation.

The Impact of Publication Bias

The existence of a positive publication bias in the licensing literature was assessed via two methods: by creating a funnel plot and by investigating article status (published vs. unpublished) as a between-study moderator. First, we created a funnel plot of the effects sizes of the published data ($n = 57$) against their corresponding standard errors. If there is no positive publication bias, the funnel plot should be roughly symmetrical around the true effect size estimate, because without a publication bias an equal amount of studies should find smaller effects than the true effect size as there are studies that find larger effects. If there is a positive publication bias, high powered studies should be close to the true effect size and be present on both sides, whereas studies with smaller samples and higher variability would be more likely to only appear as being larger than the average effect size. Smaller studies that find no significant effect are unlikely to be published, and therefore the bottom side of the funnel plot remains relatively empty in the presence of a positive publication bias.

Figure 2 shows the effect sizes in the random-effects funnel plot with the filled-in data. In this plot, the effect size estimates from the included published studies are represented as black dots. The white dots represent the estimated number of missing studies ($n = 21$). Interpretation of the symmetry of the black dots in the funnel plot is rather subjective (Thornton & Lee, 2000), but a regression analysis can be conducted with the standard error as predictor of the observed outcomes. When there is a publication bias, the effect sizes are positively related to the standard error (showing that studies with smaller sample sizes find larger effects; Egger, Davey Smith, Schneider, & Minder, 1997). The regression test for funnel plot asymmetry using the weighted regression with multiplicative dispersion model showed a significant effect, $t(55) = 5.72$, $p < .001$, indicating the presence of a positive publication bias.

Second, we tested for study status (published vs. unpublished) as a between-study moderator. Article status (published vs. unpublished) turned out to significantly influence the estimated moral licensing effect size, with published studies having larger effects ($d = 0.43$, $SE = 0.04$) than

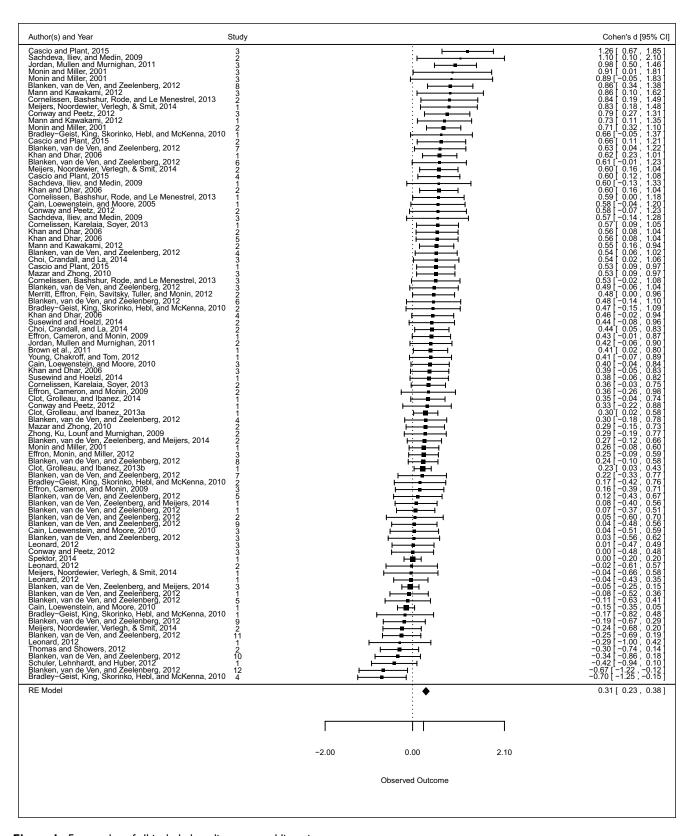| Author(s) and Year	Study		Cohen's d [95% CI]
Cascio and Plant, 2015	3		1.26 [0.67 , 1.85]
Sachdeva, Iliev, and Medin, 2009	2		1.10 [0.10 , 2.10]
Jordan, Mullen and Murnighan, 2011	3		0.98 [0.50 , 1.46]
Monin and Miller, 2001	3		0.91 [0.01 , 1.81]
Monin and Miller, 2001	3		0.89 [−0.05 , 1.83]
Blanken, van de Ven, and Zeelenberg, 2012	8		0.86 [0.34 , 1.38]
Mann and Kawakami, 2012	3		0.86 [0.10 , 1.62]
Cornelissen, Bashshur, Rode, and Le Menestrel, 2013	2		0.84 [0.19 , 1.49]
Meijers, Noordewier, Verlegh, & Smit, 2014	1		0.83 [0.18 , 1.48]
Conway and Peetz, 2012	3		0.79 [0.27 , 1.31]
Mann and Kawakami, 2012	1		0.73 [0.11 , 1.35]
Monin and Miller, 2001	2		0.71 [0.32 , 1.10]
Bradley−Geist, King, Skorinko, Hebl, and McKenna, 2010	1		0.66 [−0.05 , 1.37]
Cascio and Plant, 2015	2		0.66 [0.11 , 1.21]
Blanken, van de Ven, and Zeelenberg, 2012	7		0.63 [0.04 , 1.22]
Khan and Dhar, 2006	1		0.62 [0.23 , 1.01]
Blanken, van de Ven, and Zeelenberg, 2012	6		0.61 [−0.01 , 1.23]
Meijers, Noordewier, Verlegh, & Smit, 2014	2		0.60 [0.16 , 1.04]
Cascio and Plant, 2015	4		0.60 [0.12 , 1.08]
Sachdeva, Iliev, and Medin, 2009	1		0.60 [−0.13 , 1.33]
Khan and Dhar, 2006	2		0.60 [0.16 , 1.04]
Cornelissen, Bashshur, Rode, and Le Menestrel, 2013	1		0.59 [0.00 , 1.18]
Cain, Loewenstein, and Moore, 2005	1		0.58 [−0.04 , 1.20]
Conway and Peetz, 2012	2		0.58 [−0.07 , 1.23]
Sachdeva, Iliev, and Medin, 2009	3		0.57 [−0.14 , 1.28]
Cornelissen, Karelaia, Soyer, 2013	1		0.57 [0.09 , 1.05]
Khan and Dhar, 2006	2		0.56 [0.08 , 1.04]
Khan and Dhar, 2006	5		0.56 [0.08 , 1.04]
Mann and Kawakami, 2012	2		0.55 [0.16 , 0.94]
Blanken, van de Ven, and Zeelenberg, 2012	4		0.54 [0.06 , 1.02]
Choi, Crandall, and La, 2014	3		0.54 [0.02 , 1.06]
Cascio and Plant, 2015	1		0.53 [0.09 , 0.97]
Mazar and Zhong, 2010	3		0.53 [0.09 , 0.97]
Cornelissen, Bashshur, Rode, and Le Menestrel, 2013	3		0.53 [−0.02 , 1.08]
Blanken, van de Ven, and Zeelenberg, 2012	3		0.49 [−0.06 , 1.04]
Merritt, Effron, Fein, Savitsky, Tuller, and Monin, 2012	1		0.48 [0.00 , 0.96]
Blanken, van de Ven, and Zeelenberg, 2012	6		0.48 [−0.14 , 1.10]
Bradley−Geist, King, Skorinko, Hebl, and McKenna, 2010	2		0.47 [−0.15 , 1.09]
Khan and Dhar, 2006	4		0.46 [−0.02 , 0.94]
Susewind and Hoelzl, 2014	2		0.44 [−0.08 , 0.96]
Choi, Crandall, and La, 2014	2		0.44 [0.05 , 0.83]
Effron, Cameron, and Monin, 2009	1		0.43 [−0.01 , 0.87]
Jordan, Mullen and Murnighan, 2011	2		0.42 [−0.06 , 0.90]
Brown et al., 2011	1		0.41 [−0.06 , 0.89]
Young, Chakroff, and Tom, 2012	1		0.41 [−0.07 , 0.89]
Cain, Loewenstein, and Moore, 2010	3		0.40 [−0.04 , 0.84]
Khan and Dhar, 2006	3		0.39 [−0.05 , 0.83]
Susewind and Hoelzl, 2014	1		0.38 [−0.06 , 0.82]
Cornelissen, Karelaia, Soyer, 2013	2		0.36 [−0.03 , 0.75]
Effron, Cameron, and Monin, 2009	2		0.36 [−0.26 , 0.98]
Clot, Grolleau, and Ibanez, 2014	1		0.35 [−0.04 , 0.74]
Conway and Peetz, 2012	1		0.33 [−0.22 , 0.88]
Clot, Grolleau, and Ibanez, 2013a	1		0.30 [0.02 , 0.58]
Blanken, van de Ven, and Zeelenberg, 2012	4		0.30 [−0.18 , 0.78]
Mazar and Zhong, 2010	2		0.29 [−0.15 , 0.73]
Zhong, Ku, Lount and Murnighan, 2009	2		0.29 [−0.19 , 0.77]
Blanken, van de Ven, Zeelenberg, and Meijers, 2014	2		0.27 [−0.12 , 0.66]
Monin and Miller, 2001	1		0.26 [−0.08 , 0.60]
Effron, Monin, and Miller, 2012	3		0.25 [−0.09 , 0.59]
Blanken, van de Ven, and Zeelenberg, 2012	8		0.24 [−0.10 , 0.58]
Clot, Grolleau, and Ibanez, 2013b	1		0.23 [0.03 , 0.43]
Blanken, van de Ven, and Zeelenberg, 2012	7		0.22 [−0.33 , 0.77]
Bradley−Geist, King, Skorinko, Hebl, and McKenna, 2010	2		0.17 [−0.42 , 0.76]
Effron, Cameron, and Monin, 2009	3		0.16 [−0.39 , 0.71]
Blanken, van de Ven, and Zeelenberg, 2012	5		0.12 [−0.43 , 0.67]
Blanken, van de Ven, Zeelenberg, and Meijers, 2014	1		0.08 [−0.40 , 0.56]
Blanken, van de Ven, and Zeelenberg, 2012	1		0.07 [−0.37 , 0.51]
Blanken, van de Ven, and Zeelenberg, 2012	2		0.05 [−0.60 , 0.70]
Blanken, van de Ven, and Zeelenberg, 2012	9		0.04 [−0.48 , 0.56]
Cain, Loewenstein, and Moore, 2010	3		0.04 [−0.51 , 0.59]
Blanken, van de Ven, and Zeelenberg, 2012	3		0.03 [−0.56 , 0.62]
Leonard, 2012	3		0.01 [−0.47 , 0.49]
Conway and Peetz, 2012	3		0.00 [−0.48 , 0.48]
Spektor, 2014	1		0.00 [−0.20 , 0.20]
Leonard, 2012	2		−0.02 [−0.61 , 0.57]
Meijers, Noordewier, Verlegh, & Smit, 2014	1		−0.04 [−0.66 , 0.58]
Leonard, 2012	1		−0.04 [−0.43 , 0.35]
Blanken, van de Ven, Zeelenberg, and Meijers, 2014	3		−0.05 [−0.25 , 0.15]
Blanken, van de Ven, and Zeelenberg, 2012	1		−0.08 [−0.52 , 0.36]
Blanken, van de Ven, and Zeelenberg, 2012	5		−0.11 [−0.63 , 0.41]
Cain, Loewenstein, and Moore, 2010	1		−0.15 [−0.35 , 0.05]
Bradley−Geist, King, Skorinko, Hebl, and McKenna, 2010	1		−0.17 [−0.82 , 0.48]
Blanken, van de Ven, and Zeelenberg, 2012	9		−0.19 [−0.67 , 0.29]
Meijers, Noordewier, Verlegh, & Smit, 2014	2		−0.24 [−0.68 , 0.20]
Blanken, van de Ven, and Zeelenberg, 2012	11		−0.25 [−0.69 , 0.19]
Leonard, 2012	1		−0.29 [−1.00 , 0.42]
Thomas and Showers, 2012	2		−0.30 [−0.74 , 0.14]
Blanken, van de Ven, and Zeelenberg, 2012	10		−0.34 [−0.86 , 0.18]
Schuler, Lehnhardt, and Huber, 2012	1		−0.42 [−0.94 , 0.10]
Blanken, van de Ven, and Zeelenberg, 2012	12		−0.67 [−1.22 , −0.12]
Bradley−Geist, King, Skorinko, Hebl, and McKenna, 2010	4		−0.70 [−1.25 , −0.15]
RE Model		◆	0.31 [0.23 , 0.38]

−2.00 0.00 2.10

Observed Outcome

Figure 1. Forest plot of all included studies on moral licensing.

Note. This figure was created using the metafor forest (res) function in the R metafor package. RE model = random-effects model; observed outcome = the obtained average effect size.
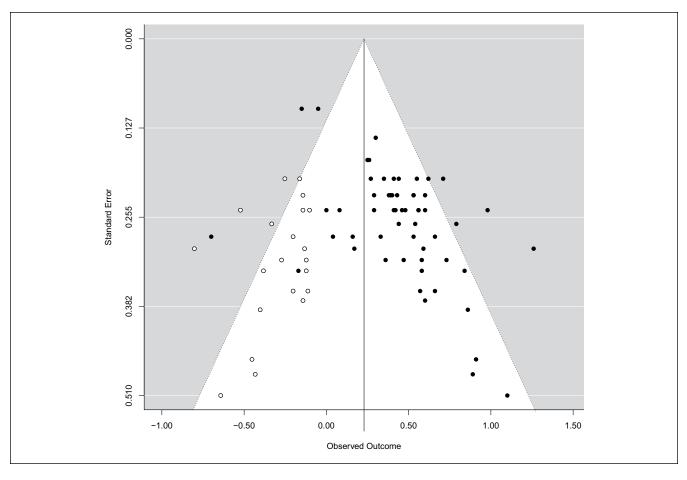
Figure 2. Trimmed and filled funnel plot of all published studies (*n* = 57) on moral licensing.
Note. The black dots represent the published studies on moral licensing. The white dots represent the estimated missing studies (*n* = 21). This figure was created using the metafor funnel (rtf) function in the R metafor package.

unpublished studies (*d* = 0.11, *SE* = 0.06), $Q_M(1) = 20.66$, *p* < .001.[3] Both methods thus indicate the existence of a publication bias. We decided to keep the article status moderator in the overall meta-regression model to control for its effects because other between-study moderators might be mainly tested in published studies.

Between-Study Moderators

To estimate the amount of variance in effect sizes of the included studies that could be explained by the theoretically predicted between-study moderators, we fitted a meta-regression model including all these moderators in the R script (Viechtbauer, 2010). Two independent coders coded the potential between-study moderators (see also the last column of Table 1). When there was disagreement, the two coders talked to the primary researcher and came to a joint conclusion. For three studies (Blanken et al., 2012; Blanken et al., 2014; Young et al., 2012), it was not possible to code one of the moderators. For instance, the effect size of Young et al. (2012) consisted of merged dependent variables. As one variable measured hypothetical donation behavior and another variable measured actual donation behavior, it was

not possible to indicate whether the behavior was actual or hypothetical. We excluded these three studies, leaving *n* = 88 studies for the moderator analyses. The initial inter-rater reliability for each moderator was as follows: Independent variable: Trait versus Action Cohen's κ = .97; Dependent variable: Actual versus Hypothetical behavior Cohen's κ = .92; Domain: Same versus Different Cohen's κ = .64; Control condition: Neutral versus Negative Cohen's κ = 1.00. Initially, the moderator "Dependent variable: A decrease in good behavior versus an increase in bad behavior" was also coded. As coding this moderator seemed problematic for both independent coders (for instance, is a preference for a White over a Black job applicant an increase in bad behavior vs. a decrease in good behavior), we decided not to include this moderator in the main regression model.

First, we conducted separate analyses for the different moderators. The type of independent variable (trait vs. action) did not influence the average effect size, $Q_M(1) = 0.05$, *p* = .819. The type of behavior measured in the dependent variable (actual vs. hypothetical) did not influence the average effect size, $Q_M(1) = 0.86$, *p* = .354. Domain (same vs. different) did not influence the average effect size, $Q_M(1) = 0.18$, *p* = .670. As mentioned in the publication bias section, article status

Table 2. Between-Study Moderator Estimates in the Meta-Regression.

Moderator	β estimate	SE	Z	p
Intercept	.48	.09	5.30	<.001
Induction: Actions (d = 0.31) versus traits (d = 0.33)	−.04	.10	−0.39	.696
Dependent variable: Actual behavior (d = 0.36) versus hypothetical behavior (d = 0.28)	.02	.08	0.27	.791
Domain: Same (d = 0.29) versus different (d = 0.32)	−.10	.08	−1.32	.186
Status: Published (d = 0.43) versus unpublished (d = 0.11)	.34	.08	4.57	<.001
Control condition: Negative (d = 0.30) versus neutral (d = 0.31)	−.01	.09	−0.12	.905

(published versus unpublished) did significantly influence the estimated moral licensing effect size. The type of control condition (neutral vs. negative) did not influence the average effect size, $Q_M(1) = 0.03$, $p = .860$. Next, we analyzed the moderators together in a multiple regression model to control for collinearity. Table 2 provides an overview of the estimated β coefficients for each between-study moderator and the z-scores, standard errors, and *p* values. Also, in this multiple regression model the results showed that studies in published articles tend to have larger effect sizes than studies from unpublished work, β = .34, *p* < .001. Apart from this moderator, no other moderators were found to have a significant effect in the model. Thus, type of independent variable (trait or action), type of dependent variable (actual or hypothetical behavior), domain of dependent variable (same versus different), and type of control condition (neutral versus negative), did not moderate the moral licensing effect size.

Discussion

In the current meta-analysis, we aimed to give a state-of-the-art overview of moral licensing by examining the magnitude of the moral licensing effect and testing for potential moderators. We found an overall average effect size of *d* = 0.31, 95% CI [0.23, 0.38], that is statistically different from zero, suggesting that there is a small-to-medium moral licensing effect (Cohen, 1992). To contextualize what is meant by "small-to-medium," we compared this effect size with other relevant effect sizes in the field. Social psychological effects typically yield a value of *r* equal to .21 (approximately 4% variance explained; Richard, Bond, & Stokes-Zoota, 2003). This translates to a Cohen's *d* value of 0.43. The moral licensing effect size is thus slightly smaller than the average effect size in social psychology. Of course, this does not imply that the moral licensing effect has little theoretical or practical relevance.

An important consequence of this small-to-medium effect size is that properly powered studies on moral licensing need far more participants than are typically used. We used G*power (Faul, Erdfelder, Lang, & Buchner, 2007) to determine that one would need 165 participants per condition to have 80% statistical power to find an effect of *d* = 0.31.[4] A post hoc power analysis on all studies using our current effect size estimate finds that on average, the studies only

have 28% power. Increasing sample size in moral licensing studies will help the researcher, as it increases the chance of finding an effect. Furthermore, it will allow for more solid conclusions, for instance about the various moderators that have been tested in different studies.

Potential Moderators and Underlying Mechanisms

In the following section, we discuss the moderators that we tested for in our meta-analysis and relate the findings to previously proposed mechanisms of moral licensing. Note that we also tested for (and found) a publication bias in our moderator tests, and we will return to that topic in a later section.

Moral licensing induction: Traits versus actions. Conway and Peetz (2012) hypothesized that recalling prior good actions would lead to moral licensing, whereas recalling prior good traits would lead to consistency. They found initial support for this hypothesis. Our meta-analysis could not confirm this finding. That is, we found no difference in size of the moral licensing effect, based on whether the prior good deed was coded as an action or a trait. For our moderator section, we included both recalled actions and performed actions (such as buying ecofriendly products; Mazar & Zhong, 2010). Conway and Peetz found that recalled moral actions elicited self-licensed behaviors, whereas recalled moral traits provoked consistent good behaviors (Study 3). It may be the case that the effect obtained by Conway and Peetz is specific to recall paradigms, although we have no theoretical idea why that would be the case. Future research could further test whether and how this moderator is important for moral licensing to occur.

Behavior in the dependent variable: Actual versus hypothetical. It is cheap and easy for people to display hypothetical behavior (i.e., to state good intentions) that is in accordance with their previous laudable behavior, which would make a consistency effect more likely. Therefore, we expected that perhaps moral licensing would be stronger for actual good behavior compared with hypothetical good behavior. However, the meta-analysis showed no such differences between actual versus hypothetical behavior. There was only one study in the data

set that directly investigated similar actual versus hypothetical behavior as a dependent measure in the licensing paradigm. Young et al. (2012) investigated the effects of recalling good (vs. bad or neutral) deeds on hypothetical *and* actual donations on charity. They found a consistency effect, such that people who recalled prior good deeds donated nearly twice as much to charity compared with people who recalled prior bad deeds or who recalled a neutral conversation. They did not find a licensing effect. They also did not find any differences between reported hypothetical and actual donations, similar to the results of our meta-analysis. However, note that the actual donation request in that study was directly placed after the hypothetical donation item, which might facilitated this effect.

Maybe, in the case of a hypothetical dependent variable, people are less tempted to display immoral behavior. When confronted with an actual choice with real consequences, decisions may be different. In these so-called "affectively rich" states (Rottenstreich & Hsee, 2001), people are more likely to focus predominantly on the favorability of the outcomes. Therefore, more research is needed that not only focuses on whether the behavior is hypothetical or not, but also how tempting it is.

Moreover, we think that more careful consideration of the dependent variables being used in moral licensing studies is essential. For example, quite some research on moral licensing includes scales like the Everyday Cooperation Scale (De Hooge, Zeelenberg, & Breugelmans, 2007) and the Willingness to Volunteer Scale (DeVoe & Pfeffer, 2007) as the dependent measures. Such scales measure people's stated intention to want to help out (and perform good behavior). However, it could be the case that people overstate their willingness to help others, perhaps partly due to social desirability concerns. If a moral licensing condition causes people to indicate a lower willingness to help others, this could indeed reflect licensing. However, it could also imply that people who recall prior good behavior actually become more honest. When they are asked whether they want to volunteer, they might be more thoroughly thinking about it and answering it truthfully, which is usually that one actually does not have the time to volunteer. Both such a consistency process and moral licensing would predict lower scores on for example the Willingness to Volunteer Scale, but the underlying process is completely different. Researchers should be aware of these possible confounds when studying licensing with hypothetical behavioral measures.

Domain: Same versus different domain. If effects similar to mental accounting (Thaler, 1985) occur in the realm of moral licensing, licensing effects would be larger when the moral and immoral behaviors were measured in the same domain compared with that in a different domain. However, the meta-analysis did not find differences between immoral behaviors in the same domain versus in a different domain. Effron and Monin (2012) reasoned that the relative effectiveness of same-domain versus different-domain moral licensing depends on whether the immoral behavior being licensed is blatantly bad or only ambiguously bad. They tested this idea in studies of observers' inclination to license the behaviors of other persons and found that different-domain licensing worked better than same-domain licensing when the bad deeds were blatant. In contrast, same-domain licensing worked better than different-domain licensing when the bad deeds were ambiguous (i.e., could be construed as non-problematic). Our meta-analysis could not control for the ambiguity of the licensed behavior. We could not make inferences about how blatantly bad the behavior in the dependent variables actually was for participants, because it was not clear to what extent they actually justified the licensed behaviors. Effron and Monin tested their predictions for observer-licensing, but not for self-licensing. Future research on self-licensing could therefore clarify whether the ambiguity of the licensed behavior plays a role in the distinction between same- versus different-domain licensing.

Control condition: Neutral or negative control condition. We were surprised by the absence of a difference in whether the licensing condition (e.g., recall a previous moral action) was compared with a *neutral* control condition (e.g., recall a previous trip to the shopping mall) or a *negative* control condition (e.g., recall a previous immoral action). This finding is surprising because researchers have documented the moral cleansing effect, the finding that recalling previous immoral behavior leads to more subsequent moral behavior (the exact opposite of the moral licensing effect; see Conway & Peetz, 2012; Jordan et al., 2011; Sachdeva et al., 2009; Zhong & Liljenquist, 2006). With the existence of a moral cleansing effect, one would expect that the moral licensing effect should be larger when the positive condition is compared with the negative condition than a neutral condition. After all, if recalling immoral behavior leads to more moral behavior compared with a neutral control condition, the contrast with the recalling moral behavior condition should be even stronger.

There is some discussion about whether the moral cleansing effect is indeed as strong as initial research suggested (see Blanken et al., 2014). However, another possibility is that perhaps recalling or performing bad or immoral behavior can induce people to feel good about themselves. Research on the ease of retrieval explanation of the availability heuristic shows that reminders of behavior only elicit feelings that are congruent with these behaviors, if it is easy for participants to retrieve these behaviors (Schwarz et al., 1991). If participants find it difficult to recall immoral behavior (or can only come up with an instance that was a very long time ago), they might infer from this that they are quite moral persons themselves. Furthermore, exposure to extreme stimuli from a category can also sometimes remind people of the other extreme of this category (Herr, Sherman, & Fazio, 1983). By activating the concept of immoral behavior, the

manipulations might also have activated the concept of morality. For these reasons, some of the manipulations that induced immoral behavior might have inadvertently also activated moral behavior somewhat, thereby leading to a moral licensing effect that is as large as moral licensing contrasted to neutral behaviors.

Taken together, none of the examined moderators that were based on existing theorizing on moral licensing significantly moderated the moral licensing effect in the current meta-analysis. There are various possible explanations for this, in addition to the ones described above. Moderators itself might again have moderators for when the effect occurs or not. The studies that were included in the meta-analysis used very different designs and included different populations. Variations in experimental design may have resulted in very different outcomes with respect to the moderating variables. In addition, small sample sizes of several included studies may have made it more difficult to perform a proper assessment of the predicted moderators as that made effect size estimates less precise. We therefore recommend researchers to critically revise the research paradigms and increase sample sizes that they use to study the moral licensing effect and, if necessary, revise the theory.

Other possible mechanisms. The literature on moral licensing does not lack explanatory mechanisms. Conway and Peetz (2012), for instance, proposed that recalling *distant* moral behavior makes one's moral values salient and induces people to act consistent with these moral values, whereas recalling *recent* moral behavior leads to moral licensing. The reason is that recent moral behavior reminds people that they made goal-progress toward their goal of being a good person. Having satisfied that goal, people can then focus attention on other goals, which might include for example selfish behavior. Conway and Peetz indeed found that recalling temporally distant moral behavior led to subsequent good behavior (consistency), whereas recalling recent moral behavior decreased people's tendency to engage in good behavior (licensing). Consistent with that idea, Fishbach and Dhar (2005) found that inducing a sense of achieved goal progress can initiate behaviors associated with licensing in the self-regulation domain such as making unhealthy food choices and spending less time on course work. However, it should be noted that in other studies that find a moral licensing effect, people recall past moral behavior in which it is not specified that it has to be recent or distant moral behavior (e.g., Jordan et al., 2011; Sachdeva et al., 2009). Therefore, we did not have sufficient information to be able to include the recency of the prior positive behavior as a possible moderator for the moral licensing effect in our meta-analysis.

Another possible mechanism for moral licensing was proposed by Miller and Effron (2010), who distinguished two different pathways of moral licensing, namely, (a) moral licensing via credits and (b) moral licensing via credentials. The first pathway, moral licensing via *credits*, asserts that

engaging in good behaviors endows people with credits that function to balance out subsequent questionable behaviors, like some sort of moral currency. A person can perceive his recent voluntary work as a credit license to decline a donation request. Thus, a person who obtained a credit might think "I have done something good so I can now do something bad." Importantly, the perceived meaning of undesirable acts does not change, but the expending of one's previously earned credit "nullifies" the current bad deed, which makes it acceptable. The deed is still seen as negative, but acceptable. The undesirable behavior is thus accepted, because it is offset by prior good behavior. The second pathway, moral licensing via *credentials*, entails that one's behavioral history provides a license by changing the way subsequent behavior is construed. These credentials function like a character witness on whom one can repeatedly call to testify that subsequent dubious behavior is not wrong or immoral. For instance, previous unprejudiced behavior establishes oneself as an unbiased person. A subsequent discriminatory attitude is thought of as less prejudiced, because it came from an ostensibly unbiased source. This means that a bad deed is seen as less bad if a prior good deed preceded it. The credits versus credentials explanation of Miller and Effron (2010) could not be tested in our meta-analysis, as it was not possible to code participants' thoughts on this distinction.

Publication Bias

A meta-analysis allows one to examine publication bias, which is the tendency that significant results are more likely to get published than non-significant results. The regression model including the between-study moderators revealed that moral licensing studies in published articles tend to have larger effect sizes than studies that did not appear in published articles. The funnel plot in Figure 2 also points to publication bias. As with any publication bias, it can be caused by researchers only submitting positive results, by the tendency of journals to reject studies with negative results, and also by the design or the execution of specific studies (e.g., a part of the unpublished articles may suffer from methodological flaws, see Thornton & Lee, 2000). We hope that our estimate of the effect size of moral licensing helps other researchers to come up with well-powered tests, for which it should be easier to publish those also if there is no effect. Running studies with larger samples would thus not only make the effect size estimates more precise, it would hopefully also help to reduce the publication bias.

Possible Limitations

Although we believe that the current meta-analysis offers important insights for the moral licensing literature, several limitations of the analysis are noteworthy. First, a substantial amount of the included effect sizes is based on small sample sizes, which could lead to an overestimation of the true effect

size (Reynolds & Day, 1984). Another point is that as the moderator analyses are performed on 88 studies, we could select a limited number of moderator variables. Furthermore, as with all meta-analyses, the inclusion criteria are subjective. Although we formulated clear inclusion criteria in advance, the inclusion of some studies was debatable. For instance, we included both studies in which the moral licensing induction exists of (the recall of) actual good behavior as well as studies in which the moral licensing induction focuses on having good intentions. One could argue that both types of independent variables are different in nature and should therefore be analyzed separately. However, according to the current theorizing on moral licensing, both good behaviors and good intentions should lead to the behaviors associated with moral licensing (Khan & Dhar, 2006). Another example is Study 1 by Gneezy, Imas, Brown, Nelson, and Norton (2012), in which the independent variable existed of a donation to charity that was automatically deducted from the participants' payment or a donation to charity that was made by the experimenter on behalf of the participants. Gneezy et al. stated that the donation that was made by the experimenter on the behalf of the participants would create a self-license. We, however, did not include these studies in our meta-analysis, because we do not think participants necessarily perceive this kind of behavior as having done something good themselves, and therefore the studies did not meet our definition of moral licensing. Finally, although we explicitly searched for unpublished studies in an attempt to prevent file drawer effects, there are likely studies on moral licensing that we could not detect.

Important Aspects for Future Studies on Moral Licensing

Hofmann, Wineski, Brandt, and Skitka (2014) recently investigated everyday morality outside of the lab, through assessing moral and immoral acts in a large community sample ($N = 1,252$).[5] They found a moral licensing pattern, in that people who committed a moral act had a larger likelihood of committing an immoral act later that day. Thus, although the average effect size of moral licensing in experimental paradigms is small-to-medium, moral licensing seems to occur in daily life. The moral licensing effect and the way it is studied thus deserve further attention. In the remainder of this section, we will outline recommendations for scholars studying moral licensing.

The main advice is for researchers studying moral licensing to increase the power of the studies. As explained before, this is important because it helps the researcher by lowering the chance of a study not finding an effect. Note that with the average power of current moral licensing studies (28%), there is a 72% chance to find a non-significant moral licensing effect if there is an effect. Our power calculation reveals that one needs 165 participants per cell, to have 80% power to find an effect as large as the one established in our meta-analysis.

Moral licensing is typically seen in the temporal pattern that people who behaved in a good or moral way later feel justified to refrain from socially desirable or morally laudable actions (Miller & Effron, 2010; Monin & Miller, 2001). All studies on moral licensing that we are aware of investigated the phenomenon in terms of two consecutive behaviors or events, where good behavior "A" leads to less desirable behavior "B." However, there is no reason to assume that the process of moral licensing actually operates in the order of these two consecutive behaviors. For instance, people may not consciously feel that after displaying certain good behavior "A" they can now engage in undesirable behavior "B," because of the prior good behavior. Perhaps people who face a dilemma in which they would like to engage in undesirable behavior "B" (e.g., not donating money to charity) are more likely to find a reason why that is acceptable after having just performed a good action (e.g., performing voluntary work in a soup kitchen). In other words, moral licensing could also be a justification strategy that people deliberately use to excuse their morally questionable behaviors. This line of reasoning is illustrated by recent studies on moral credentialing. Merritt et al. (2012) found that people strategically attempt to earn moral credentials when they anticipate performing morally dubious behaviors. For instance, participants who expected that their future behavior could be regarded as prejudiced exaggerated how favorably they perceived a Black person in a previous job hiring task. In a similar vein, Effron (2014) found that participants who were worried that their future behavior could be regarded as prejudiced or unethical overestimated to which extent previous non-racist choices or ethical behaviors proved their morality to other persons. Moreover, in their research on counterfactual licensing, Effron et al. (2012) and Effron et al. (2013) showed that to justify future undesirable behaviors, people exaggerate negative counterfactuals of their foregone behavior.

Thus, instead of the reasoning being "I just did good deed A, so now I can do bad deed B," the reasoning might also be "I feel tempted to do bad deed B. Can I do that? Yes, because I just did good deed A." This might seem like a negligible nuance, but it also implies that not only aspects of the prior good deed (as many theories and moderators now focus on), but also tempting aspects of the morally questionable behaviors might be an important part of moral licensing theory. This reasoning fits with a justification-based account of self-regulation (De Witt-Huberts, Evers, & De Ridder, 2014). Further studies on moral licensing might benefit from focusing on aspects of the immoral behavior as well.

Conclusion

Going back to the main goal of conducting the meta-analysis, the best estimate we have of the moral licensing effect is a Cohen's *d* effect size of 0.31. The effect is somewhat smaller

than other typical effects in social psychology (Richard et al., 2003), but also relatively small effects can have large societal implications. An important consequence of our effect size estimate is that future studies on moral licensing need far more participants to allow for more solid conclusions. Especially when we want to delve further into the process of why this interesting phenomenon occurs, and what its boundary conditions are, the study power needs to be sufficient. The current meta-analysis gives researchers a good starting point to determine that power. The effects moral licensing has on less desirable and negative behaviors show the importance of continuing the research on this topic: A psychological process that helps to predict when people display ambiguously racist attitudes, engage in cheating behavior, and become more selfish, deserves further attention.

Acknowledgment

Declaration of Conflicting Interests

Funding

Notes

1. Variants on these keywords were also used.
2. All analyses were carried out with the statistical software R, using the metafor package (Viechtbauer, 2010). The data set and syntax for the R file can be found in Online Appendices III and IV.
3. Q_M is an omnibus test that analyzes whether the effect sizes for the two moderator conditions significantly differ (Zhang, 1999).
4. This power analysis is based on a two-sided test. We think this is appropriate, because the alternative hypothesis (consistency in behavior) would also be theoretically likely.
5. We did not include this study in our meta-analysis because we wanted to include effect sizes based on the comparison between an experimental licensing condition and a control condition. In this specific study, there was no control condition and no random assignment of participants to conditions.

Supplemental Material

The online supplemental material is available at http://pspb.sagepub.com/supplemental.

References

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

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (1968). *Theories of cognitive consistency: A sourcebook*. Chicago, IL: Rand McNally.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk costs. *Organizational Behavior and Human Decision Processes*, *35*, 124-140. doi:10.1016/0749-5978(85)90049-4

Batson, C. D., & Thompson, E. R. (2001). Why don't moral people act morally? Motivational considerations. *Current Directions in Psychological Science*, *10*, 54-57. doi:10.1111/1467-8721.00114

Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, *6*, 1-62. doi:10.1016/S0065-2601(08)60024-6

*Blanken, I., Van de Ven, N., & Zeelenberg, M. (2012). [A variety of studies on the self-licensing effect]. Unpublished raw data.

*Blanken, I., Van de Ven, N., Zeelenberg, M., & Meijers, M. H. C. (2014). Three attempts to replicate the moral licensing effect. *Social Psychology*, *45*, 232-238. doi:10.1027/1864-9335/a000189

*Bradley-Geist, J. C., King, E. B., Skorinko, J., Hebl, M. R., & McKenna, C. (2010). Moral credentialing by association: The importance of choice and relationship closeness. *Personality and Social Psychology Bulletin*, *36*, 1564-1575. doi:10.1177/0146167210385920

*Brown, R. P., Tamborski, M., Wang, X., Barnes, C. D., Mumford, M. D., Connely, S., & Devenport, L. D. (2011). Moral credentialing and the rationalization of misconduct. *Ethics and Behavior*, *21*, 1-12. doi:10.1080/10508422.2011.537566

*Cain, D. M., Loewenstein, G., & Moore, D. A. (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *Journal of Legal Studies*, *34*, 1-25. doi:10.1086/426699

*Cain, D. M., Loewenstein, G., & Moore, D. A. (2010). When sunlight fails to disinfect: Understanding the perverse effects of disclosing conflicts of interest. *Journal of Consumer Research*, *37*, 836-857. doi:10.1086/656252

*Cascio, J., & Plant, E. A. (2015). Prospective moral licensing: Does anticipating doing good later allow you to be bad now? *Journal of Experimental Social Psychology*, *56*, 110-116. doi:S0022103114001450

Chiou, W. B., Yang, C. C., & Wan, C. S. (2011). Ironic effects of dietary supplementation: Illusory invulnerability created by taking dietary supplements licenses health-risk behaviors. *Psychological Science*, *22*, 1081-1086. doi:10.1177/0956797611416253

*Choi, B., Crandall, C. S., & La, S. (2014). Permission to be prejudiced: Legitimacy credits in the evaluation of advertisements. *Journal of Applied Social Psychology*, *44*, 190-200. doi:10.1111/jasp.12217

*Clot, S., Grolleau, G., & Ibanez, L. (2013a). Self-licensing and financial rewards: Is morality for sale? *Economics Bulletin*, *33*, 2298-2306. Retrieved from https://ideas.repec.org/a/ebl/ecbull/eb-13-00385.html

*Clot, S., Grolleau, G., & Ibanez, L. (2013b). *Anti-social behavior and self-licensing: Experimental evidence from a gangster game*. Presented at 30. Journées de Micraoeconomie Appliquée, Nice, FRA (2012-06-06- 2013-06-07). Retrieved from: http://prodinra.inra.fr/record/264362

*Clot, S., Grolleau, G., & Ibanez, L. (2014). Smug alert! Exploring self-licensing behavior in a cheating game. *Economics Letters*, *123*, 191-194. doi:10/1016/j.econlet.2014.01.039

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155.

*Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates

whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, *38*, 907-919. doi:10.1177/0146167212442394

*Cornelissen, G., Bashshur, M. R., Rode, J., & Le Menestrel, M. (2013). Rules or consequences: The role of ethical mindsets in moral dynamics. *Psychological Science*, *24*, 482-488. doi:10.1177/0956797612457376

*Cornelissen, G., Karelaia, N., & Soyer, E. (2013, September). *Clicktivism or slacktivism? The ironic effects of symbolic prosocial behavior*. Paper presented at Society for the Advancement of Judgment and Decision Making Studies, Granada, Spain.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*, 67-80. doi:10.1007/s00199-006-0153-z

De Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition & Emotion*, *21*, 1025-1042. doi:10.1080/02699930600980874

DerSimonian, R., & Laird, N. (1986). Meta analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177-188. doi:10.1016/0197-2456(86)90046-2

DeVoe, S. E., & Pfeffer, J. (2007). Hourly payment and volunteering: The effect of organizational practices on decisions about time use. *Academy of Management Journal*, *50*, 783-798. doi:10.5465/amj.2007.26279171

De Witt-Huberts, J. C., Evers, C., & De Ridder, D. T. (2012). License to sin: Self-licensing as a mechanism underlying hedonic consumption. *European Journal of Social Psychology*, *42*, 490-496. doi:10.1111/j.1751-9004.2010.00263.x

De Witt-Huberts, J. C., Evers, C., & De Ridder, D. T. (2014). "Because I am worth it": A theoretical framework and empirical review of a justification-based account of self-regulation failure. *Personality and Social Psychology Review*, *18*, 119-138. doi:10.1177/1088868313507533

Effron, D. A. (2014). Making mountains of morality from molehills of virtue: Threat causes people to overestimate their moral credentials. *Personality and Social Psychology Bulletin*, *40*, 972-985. doi:10.1177/0146167214533131

*Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites. *Journal of Experimental Social Psychology*, *45*, 590-593. doi:10.1016/j.jesp.2009.02.001

*Effron, D. A., Monin, B., & Miller, D. T. (2012). Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *Journal of Personality and Social Psychology*, *103*, 916-932. doi:10.1037/a0030008

Effron, D. A., & Monin, B. (2010). Letting people off the hook: When do good deeds excuse transgressions? *Personality and Social Psychology Bulletin*, *36*, 1618-1634.

Effron, D. A., Monin, B., & Miller, D. T. (2013). The unhealthy road not taken: Licensing indulgence by exaggerating counterfactual sins. *Journal of Experimental Social Psychology*, *49*, 573-578. doi:10.1016/j.jesp.2012.08.012

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629-634. doi:10.1136/bmj.315.7109.629

Eskine, K. J. (2012). Wholesome foods and wholesome morals? Organic foods reduce prosocial behavior and harshen moral judgments. *Social Psychological and Personality Science*, *4*, 251-254. doi:10.1177/1948550612447114

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. doi:10.3758/bf03203630

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research*, *32*, 370-377. doi:10.1086/497548

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*, 195-220. doi:10.1037/h0023552

Gawronski, B., & Strack, F. (2012). *Cognitive consistency: A fundamental principle in social cognition*. New York, NY: Guilford Press.

Gneezy, A., Imas, A., Brown, A., Nelson, L. D., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science*, *58*, 179-187. doi:10.1287/mnsc.1110.1437

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486-504. doi:10.1037//1082-989X.3.4.486

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, *21*, 107-112. doi:10.1080/00223980.1946.9917275

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effect. *Journal of Experimental Social Psychology*, *19*, 323-340. doi:10.1016/0022-1031(83)90026-4

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539-1558. doi:10.1002/sim.1186

Hofmann, W., Wineski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*, 1340-1343. doi:10.1126/science.1251560

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*, e124. doi:10.1371/journal.pmed.0020124

*Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, *37*, 701-713. doi:10.1177/0146167211400208

*Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research*, *43*, 259-266. doi:10.1509/jmkr.43.2.259

*Leonard, B. (2012). Unpublished raw data.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE.

*Mann, N. H., & Kawakami, K. (2012). The long, steep path to equality: Progressing on egalitarian goals. *Journal of Experimental Psychology: General*, *141*, 187-197. doi:10.1037/a0025602

Marascuillo, L. A., Busk, P. L., & Serlin, R. C. (1988). Large sample multivariate procedures for comparing and combining

effect sizes within a single study. *Journal of Experimental Education*, *58*, 69-85. doi:10.1080/00220973.1988.10806497

*Mazar, N., & Zhong, C. B. (2010). Do green products make us better people? *Psychological Science*, *21*, 494-498. doi:10.1177/0956797610363538

*Meijers, M. H. C., Noordewier, M. K., Verlegh, P. W. J., & Smit, E. G. (2014). *Identity relevance moderates the licensing effect* (Chapter from doctoral dissertation). Retrieved from http://dare.uva.nl/record/1/432499

*Merritt, A. C., Effron, D. A., Fein, S., Savitsky, K. K., Tuller, D. M., & Monin, B. (2012). The strategic pursuit of moral credentials. *Journal of Experimental Social Psychology*, *48*, 774-777. doi:10.1016/j.jesp.2011.12.017

Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and Personality Psychology Compass*, *4*, 344-357. doi:10.1111/j.1751-9004.2010.00263.x

Miller, D. T., & Effron, D. A. (2010). Psychological license: When it is needed and how it functions. *Advances in Experimental Social Psychology*, *43*, 115-155. doi:10.1016/S0065-2601(10)43003-8

*Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*, 33-43. doi:10.1037//0022 3514.81.1.33

Mukhopadhyay, A., & Johar, G. V. (2009). Indulgence as self-reward for prior shopping restraint: A justification-based mechanism. *Journal of Consumer Psychology*, *19*, 334-345. doi:10.1016/j.jcps.2009.02.016

Mukhopadhyay, A., Sengupta, J., & Ramanathan, S. (2008). Recalling past temptations: An information-processing perspective on the dynamics of self-control. *Journal of Consumer Research*, *35*, 586-599. doi:10.1086/591105

Reynolds, S., & Day, J. (1984). *Monte Carlo studies of effect size estimates and their approximations in meta-analysis*. Retrieved from http://catalogue.nla.gov.au/Record/5450328

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331-363. doi:10.1037/1089-2680.7.4.331

Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin*, *99*, 400-406. doi:10.1037/0033 2909.99.3.400

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses and electric shocks: On the affective psychology of risk. *Psychological Science*, *12*, 185-190. doi:10.1111/14679280.00334

*Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, *20*, 523-528. doi:10.1111/j.1467-9280.2009.02326.x

*Schüler, S., Lehnhardt, N., & Huber, M. (2012). *Can priming pro-social values eliminate the moral licensing effect?* Unpublished manuscript, Technische Universität Dresden, Germany.

Schwarz, N., Strack, F., Bless, H., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195-202. doi:10.1037//0022-3514.61.2.195

Spektor, M. (2014). Unpublished raw data.

*Susewind, M., & Hoelzl, E. (2014). A matter of perspective: Why past moral behavior can sometimes encourage and other times discourage future moral striving. *Journal of Applied Social Psychology*, *44*, 201-209. doi:10.1111/jasp.12214

Thaler, R. H. (1985). Mental accounting and consumer choice. *Marketing Science*, *4*, 199-214. doi:10.1287/mksc.1070.0330

*Thomas, J. S., & Showers, C. (2012). Unpublished raw data.

Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology*, *53*, 207-216. doi:10.1016/S0895 4356(99)00161-4

Viechtbauer, W. (2010). Conducting meta-analysis in R with the metafor package. *Journal of Statistical Software*, *36*, 1-48. Retrieved from http://www.jstatsoft.org/v36/i03/paper

*Young, Y., Chakroff, A., & Tom, J. (2012). Doing good leads to more good: The reinforcing power of a moral self-concept. *Review of Philosophy and Psychology*, *3*, 325-334. doi:10.1007/s13164-012-0111-6

Zhang, P. (1999). Omnibus test of normality using the Q statistic. *Journal of Applied Statistics*, *26*, 519-528. doi:10.1080/02664769922395

*Zhong, C. B., Ku, G., Lount, R. B., & Murnighan, J. K. (2009). Compensatory ethics. *Journal of Business Ethics*, *92*, 323-339. doi:10.1007/s10551-009-0161-6

Zhong, C. B., & Liljenquist, K. A. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, *313*, 1451-1452. doi:10.1126/science.113072