

The Social Dynamics of Morality

Florien M. Cramwinckel

ISBN: 978-94-6182-708-1

Cover: Off Page, Amsterdam

Lay-out & printing by: Off Page, Amsterdam

Photo credit cover: Dennis Kuzee

Cover models: Bart van Benthem, Florian Cramwinckel, Joshua Cramwinckel, Ruth Rexwinkel, Tirza Cramwinckel, & Vincent van Ruijven

Copyright © 2016 by F. M. Cramwinckel

The Social Dynamics of Morality

De Sociale Dynamiek van Moraliteit

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector
magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor
promoties in het openbaar te verdedigen op vrijdag 14 oktober 2016 des middags te
12.45 uur

door

Florien Maite Cramwinckel

geboren op 13 maart 1986

te Voorschoten

Promotoren:

Prof. dr. C. van den Bos

Prof. dr. E. van Dijk

Dit proefschrift werd (mede) mogelijk gemaakt door een *Graduate Program* onderzoeksbeurs van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO, 022.003.040), uitgereikt aan het Kurt Lewin Instituut.

Beoordelingscommissie:

Prof. dr. Henk Aarts

Prof. dr. Vincent Buskens (voorzitter)

Prof. dr. Belle Derks

Prof. dr. Naomi Ellemers

Prof. dr. Paul van Lange

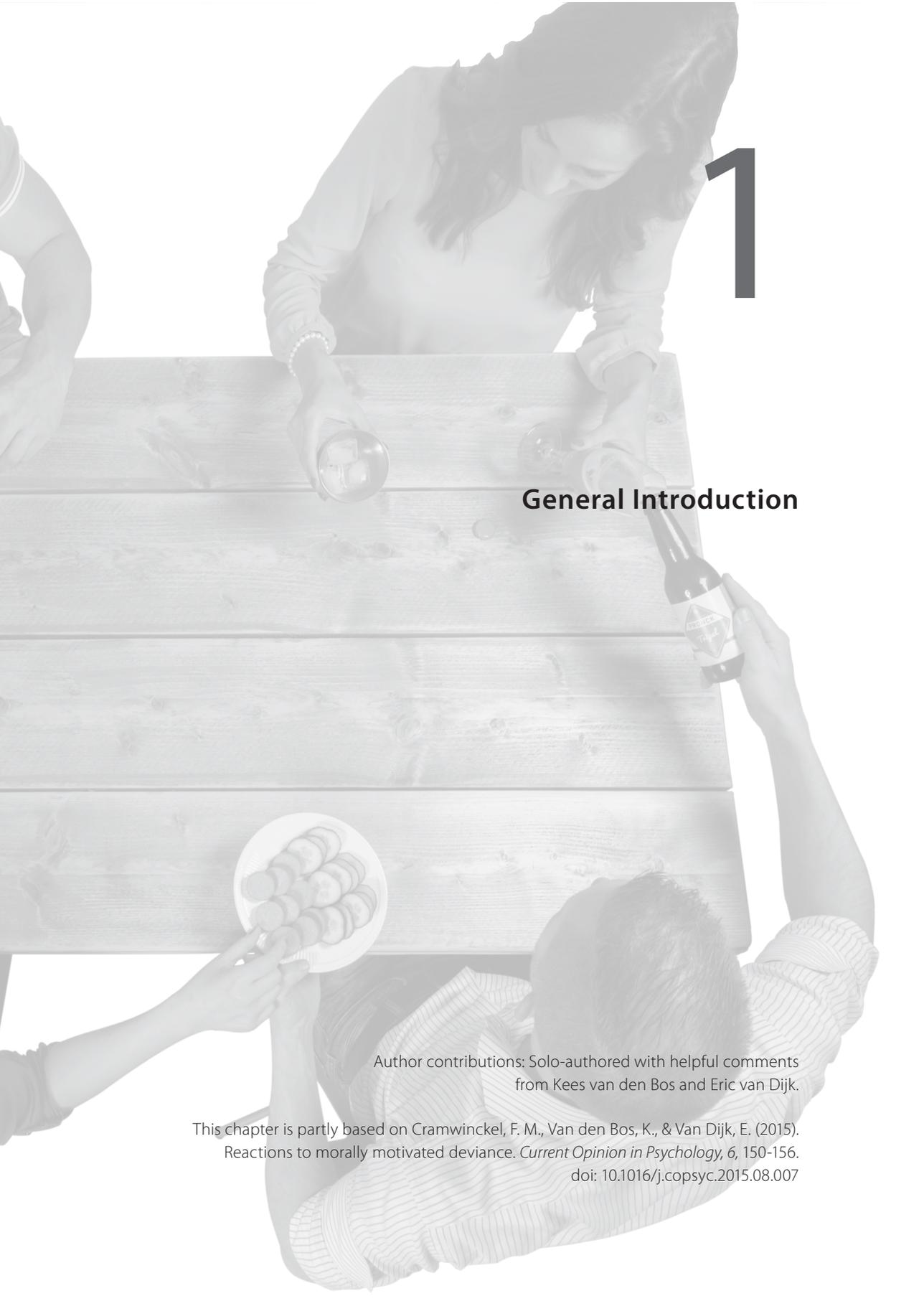
Voor mijn gezin,

Omdat 'echtgenote' en 'moeder' mijn meest waardevolle titels zijn

Table of contents

Chapter 1	General Introduction	11
Chapter 2	The Threat of Moral Refusers for One's Self-Concept and the Protective Function of Physical Cleansing	33
Chapter 3	Negative Reactions to Non-Drinkers: Self- and Other-Directed Aggression After Confrontations with Moral Refusers	61
Chapter 4	Derogating Benevolent Behavior of Deviant In-Group Members—Group Processes Within a Real-Life Sample of Heterosexual Christians	83
Chapter 5	Displaying Counter-Attitudinal Behavior Regarding Equal Rights for Homosexuals—Effects on Self-Evaluations and Evaluations of Others	99
Chapter 6	General Discussion	135
Addendum	References	155
	Dutch Summary Nederlandse Samenvatting	169
	Acknowledgements Dankwoord	171
	KLI dissertation series	175





1

General Introduction

Author contributions: Solo-authored with helpful comments from Kees van den Bos and Eric van Dijk.

This chapter is partly based on Cramwinckel, F. M., Van den Bos, K., & Van Dijk, E. (2015). Reactions to morally motivated deviance. *Current Opinion in Psychology*, 6, 150-156. doi: 10.1016/j.copsy.2015.08.007

It is very important to me to be a good person. For me, this means that I hope that when I reflect on my life and the choices I have made, I will conclude that I have done the right thing for both myself and others. I hope, therefore, that I will have displayed behavior that I can be proud of, even when the situation was difficult or challenging. I hope that others will also see me as a good person. Thus, I care deeply about being moral and being perceived as moral by others. And I am not the only one. In fact, most people place moral values among the most important guiding principles in life (see, e.g., Schwartz, 1992). This means that people will try to act in ways that demonstrate their morality, in their own eyes as well as in the eyes of others (Cohen & Sherman, 2014; Schaumberg & Wiltermuth, 2014). But how do you do that, and what is ‘morality’ exactly?

Morality, according to the Merriam-Webster dictionary, can be defined as “beliefs about what is right behavior and what is wrong behavior”. So morality has to do with what is right and wrong. But which acts are moral? What is the right or wrong thing to do? Philosophers have tried to provide satisfying answers to these questions for centuries. This has led to several normative theories of morality—which inform people what is right and wrong and thus tell people what they must and must not do.

Two prominent philosophical schools of thought are deontology and consequentialism (e.g., Beauchamp, 1991). Deontology comes from the greek word ‘deon’ which means ‘duty’. According to deontologists, acts are moral when they correspond to ‘moral obligations’. This means that acts are ‘wrong’ or ‘right’ insofar as they correspond with moral principles. One of the best known deontological theories was developed by Immanuel Kant (1785). Kant argued “*that an action is right if, and only if, it conforms to a moral rule that a rational agent (person) would necessarily follow if the agent were acting in accordance with reason*” (Beauchamp, 1991, p. 178). Thus, according to Kant, actions are only right when they are done for the right reasons (i.e., because they correspond to the universal moral law) and only wrong when they violate this universal moral law. This universal moral law is a *categorical imperative*, which means that it informs people what they *must* do regardless of whether or not they want to do it and regardless of the consequences. Importantly, according to this categorical imperative, people must only display behavior when they at the same time could also want their behavior to become an universal law (Kant, 1785).

Consequentialism, on the other hand, is a generic term for moral theories that are concerned with the outcomes of acts. According to consequentialist theories, behavior is moral when it produces the right outcomes. This means that no behaviors are inherently wrong or right, it all depends on the consequences. One of the most influential consequentialist

1

2

3

4

5

6

&

1

2

3

4

5

6

&

theories is utilitarianism, as formulated by Jeremy Bentham (1789) and John Stuart Mill (1863). According to utilitarianists, such as Bentham and Mill, acts are right when they provide the best outcomes for all persons affected by that act (Beauchamp, 1991). For example, Mill proposed the Greatest Happiness Principle which holds that “*actions are right in proportion as to which they tend to promote happiness, wrong as they tend to produce the reverse of happiness*” (Mill, 1863, Chapter 2).

These two philosophical moral theories—deontology and consequentialism—can perhaps be best explained with some examples. One well-known example is the trolley dilemma; a thought experiment originally designed by Philippa Foot (1978). In this dilemma, people need to make a choice whether or not to sacrifice one person in order to save five. Consider the following scenario: You are standing next to the train tracks, close to a switch. A trolley with broken brakes is speeding in your direction and there is no way to stop the trolley. The trolley is currently headed towards one of the tracks on which five people are stuck. They cannot get off the train tracks and will surely die when the trolley hits them. However, you can save them by pulling the switch. This would lead the trolley to switch to a different track. On this alternative track, there is only one person stuck. Pulling the switch would mean that the trolley would hit this one person and this person would surely die.

What is the right thing to do? Should you do nothing and let five people die? Or should you pull the switch and rescue five people by sacrificing one? According to a deontological line of reasoning, you should do nothing, because pulling a switch that would kill one person is inherently wrong even though you could save five people. In Kant’s words, killing someone would violate the universal moral law (i.e., the categorical imperative) and would therefore be wrong. However, to many people, this makes no sense in this scenario. Most people think that you should adhere to consequentialist thinking, which in this case implies that you should pull the switch and save five people while sacrificing one. Thus, in this scenario, people often adhere to a consequentialist approach and would argue that saving five people and sacrificing one is more moral than saving one person and sacrificing five.

Now consider an alternative example. Imagine that you are a doctor in a hospital and one of your patients is suffering from a life-threatening disease. You have the medicine available that will cure this patient. However, you also discover that this person is a perfect match for five patients in your hospital that are on the organ list. These five patients would surely die without a match and the organs from the person with the life-threatening disease would save their lives. Should you let this patient with the life-threatening disease die and use the organs to save five other patients? If you are like most people, you would be horrified by

this idea. In this case, the categorical imperative—it is wrong to let this patient die even if you can save five people by doing so—makes more sense to people and seems like the moral course of action for many if not most of us.

These and other moral dilemmas have often been used in psychological experiments to investigate moral reasoning—what do people find morally acceptable, why and under which circumstances (e.g., Broeders, Van den Bos, Müller, & Ham, 2011; Greene et al., 2009; Nichols & Mallon, 2006; Waldmann & Dieterich, 2007)? In other words, these moral dilemmas have been used to investigate the content of morality (i.e., what is moral behavior and what is immoral behavior?). And these studies have yielded some very important and interesting findings. For example, Cikara and colleagues (2010) investigated different types of moral dilemmas such as the ones described above to investigate whether stereotypes motivate people to value the lives of their in-group members over the lives of out-group members. They indeed demonstrated that people find it more morally acceptable to save an in-group member (in this case, a fellow American) and sacrifice an out-group member (e.g., a homeless person) than to save an out-group member while sacrificing an in-group member. These findings thus show that moral judgments are not made in a social vacuum but are also influenced by intergroup biases and stereotypes.

Although these and other moral dilemma studies have provided many important insights, there are some drawbacks. As most readers probably realize, the described scenarios are quite bizarre and extreme. Only a handful of people ever encounter a real-life situation that even remotely resembles making life or death decisions such as the ones described above. Morality in real life often does not take the form of life or death decisions but is centered around more mundane events. Furthermore, how people react to hypothetical situations may better reflect what they think they *should* do than what they would *actually* do in that situation. Indeed, people are often unable to correctly predict or assess how they would react in real-life situations (see e.g., Baumert, Halmburger, & Schmitt, 2013). This inability may be particularly pronounced in situations involving morality because people are strongly motivated to see and portray themselves as good human beings (e.g., Cohen & Sherman, 2014). For example, Baumert and colleagues (2013) investigated how willing people would be to intervene when they witnessed a thief. These researchers demonstrated that although people were extremely willing to intervene in a hypothetical theft scenario, only 27% actually intervened in a real-life theft situation. Because of these reasons, I think it is informative to investigate morality in *immersive settings* (e.g., Blascovich, Mendes, & Seery, 2002) where people actually experience the specific situation that one is interested in. In other words, investigating morality in situations that are relatively high

1

2

3

4

5

6

&

1

in mundane realism may lead to particularly relevant insights about human behavior in moral situations.

2

Another important drawback of focusing on the content of morality is that it neglects a very important observation: People often do not agree on what is moral and immoral.

3

In fact, diversity of moral beliefs exists between different people, different societies and different cultures (e.g., Beauchamp, 1991), which makes it hard to find an answer to the question what morality is.¹ However, although people may not agree on the content of

4

morality, they do agree that it is important to be moral (e.g., Schaumberg & Wiltermuth, 2014; Schwartz, 1992). Therefore, I think it is important to focus on answering different

5

questions. How does morality function? How does it influence people in their daily lives? Thus, in short, what does morality *do* to people?

6

&

These questions also highlight a crucial difference between philosophical and psychological approaches to morality. Whereas philosophers are predominantly interested in how people *should* behave, I—as an experimental social psychologist—am interested in how people *actually* behave. My strength is to investigate how people actually react when they are placed in situations where moral considerations play a role. In this social psychological dissertation, I² will therefore take a different and maybe less conventional approach: Rather than focusing on the prescriptive aspects of morality (i.e., what *is* morality and what *should* people do?), I focus on the descriptive aspects of morality (i.e., what *does* morality and what do people actually *do*?).

The Power of Morality

In my view, when investigating morality, it is important to consider its social function. How does the construct of morality shapes people's attitudes, thoughts and behavior? It has been argued that morality regulates people's behavior in a social context (Ellemers & Van den Bos, 2012; Rai & Fiske, 2011). Because people want to display moral behavior and try to avoid displaying immoral behavior (e.g., Zhong, Liljenquist, & Cain, 2009), labeling behaviors

¹ This does not mean that there is not a certain universal moral code of conduct that everyone should in principle agree with, but it does implies that people in reality often do not agree on what is right and wrong (see e.g., Beauchamp, 1991).

² Although the General Introduction and General Discussion are written from a first person perspective (i.e., "I", "me", "mine", etc.), the empirical chapters are written from a plural perspective (i.e., "we", "us", "our" etc.). This reflects the collaborative nature of this PhD project; all empirical research reported in this dissertation has been the result of joint efforts.

as wrong or right has motivational power. Especially labeling certain behaviors as ‘wrong’ may have a very powerful influence on the performers of those behaviors. Morality is thus a very powerful tool that influences how people think and feel about themselves. This is reflected in a growing body of literature demonstrating the importance of morality for people’s self-concepts.

Moral self-regulation theory describes that people constantly compare their moral behavior with their moral ideals (Zhong et al., 2009). People experience discomfort when their behavior falls short of their moral standard. This evokes a need to compensate moral shortcomings by displaying future moral behavior (Zhong et al., 2009). For example, Sachdeva and colleagues (2009) illustrated this process of *moral compensation* by instructing people to write a story about themselves that contained certain words. These words were either moral words (e.g., caring, fair), immoral words (e.g., disloyal, selfish), or neutral words (e.g., book, keys). Afterwards, people were asked to donate money to charity. People who included immoral words in this story about themselves donated about five times as much (\$5.30) to charity as people who included moral words in their story (\$1.07). Apparently, thinking about immoral aspects of oneself evoked the need to engage in moral behavior. This fits with moral self-concept maintenance theory (Mazar, Amir, & Ariely, 2008), which states that people balance their need to maintain a positive moral self-concept with the temptation to engage in immoral behavior. According to this theory, people will engage in immoral behavior whenever they can do so without repercussions for their moral self-concept. This means that people will display immoral behavior when they can easily categorize their behavior as morally acceptable or even unrelated to morality. For example, it is easier to take a pen that’s worth 50 cents from the office than to take 50 cents from the secretary’s office and put it in your pocket. According to self-concept maintenance theory, this is because it is easy to categorize taking a pen as morally acceptable behavior (after all, you may plan on using it to grade papers after hours) while it is not easy to categorize stealing money as morally acceptable. Both moral self-regulation theory and self-concept maintenance theory emphasize the importance of morality for people’s perceptions of themselves.

Morality in the Social World

Morality is thus deeply valued by people, and being and appearing moral is a very important goal in life for many of us (Cohen & Sherman, 2014; Effron, 2014; Schaumberg & Wiltermuth, 2014). Especially important for the work reported in this dissertation is that people’s sense of morality not only depends on their own behavior. We live in a social world where we are constantly exposed to other people. What these other people think, feel, and

1

2

3

4

5

6

&

1

do also influences us and our ideas about ourselves. This is, in fact, the core notion of social psychology. Therefore, people's moral self-esteem not only depends on their own actions and inactions, but also on how their behaviors relate to the behaviors of relevant others.

2

3

For example, Webber and colleagues (2013) investigated how people's reaction to their own behavior was influenced by the moral judgment and behavior of others. In two experiments, participants engaged in an 'extermination-task' where they killed bugs—or so they were led to believe. Webber and colleagues manipulated whether another participant ostensibly rejected this behavior on moral grounds, or ostensibly thought this behavior was morally acceptable. The most important findings were that people experienced higher levels of stress and negative self-directed emotions (e.g., guilt and shame) when their 'extermination behavior', was disapproved of by others.

4

5

6

&

This can be explained by social comparison theory (Festinger, 1954). According to social comparison theory, people's self-evaluations in a certain domain (e.g., morality) are based on how well they are doing in that domain. But how would you know how well you are doing? Sometimes, it is very easy to figure this out because there are objective standards available. For example, if you want to know whether you can run fast enough to qualify for the Olympic games, you time yourself while running and you know whether you meet the criteria proposed by the national Olympics committee to send people off to the Olympic games. Easy. (At least the timing part, not the actual running...). But there are also domains where objective standards are not available.

I argue that morality is such a domain. Although people in general agree that it is important to be moral (e.g., Schaumberg & Wiltermuth, 2014; Schwartz, 1992) it is not that clear *how* moral certain behaviors are exactly—as I explained above. So what do you do when you want to know how moral you are? According to social comparison theory, you compare yourself to others: or, in other words, you engage in processes of social comparison. When you engage in downward comparison, you compare yourself with someone who is acting in less moral ways than you and this can make you feel good about yourself because you are better than the other person. When you engage in upward social comparison, you compare yourself with someone who is acting in more moral ways than you and this can make you feel bad about yourself because you are lesser than the other person.

It is not always possible to choose with whom to compare yourself. Sometimes, you are involved in situations where the comparison partner is predetermined by the environment you are in. For example, imagine taking part in an experiment. You had some spare time and decided to earn some money by participating in an experiment. During this experiment,

you were instructed to taste a sausage and answer some questions. How would you feel about yourself after taking part in this experiment? Probably you would feel pretty good about yourself. After all, you did something good by helping to further scientific knowledge and made some money in the process as well.

But now imagine that you learn that another participant in the same experiment was also instructed to taste a sausage but she had refused because she thought eating meat was immoral. How would you now feel about yourself? How would her refusal reflect on your own behavior? Would you still think you are a good person?

In this situation, you may compare your own behavior to that of the refusing other participant which means you would engage in upward social comparison in the moral domain. This may lead you to question your morality—a quite aversive experience for people (see e.g., Monin, 2007). In fact, I argue that because of the tremendous importance placed on morality and the fact that people's moral self-evaluations can depend on the behavior of others, people's sense of morality can become threatened when they are exposed to the moral behavior of others.

All these deliberations led me to focus on the following topic in my dissertation: I investigate how people react to the moral behavior of others. More specifically, I focus on reactions towards *morally motivated deviance*, defined as behavior that is (1) displayed for moral reasons and (2) different from the average behavior. In four empirical chapters, I demonstrate how being confronted with morally motivated deviance influences people's evaluations of, and their behavior towards, morally motivated deviants. I also investigate how exposure to morally motivated deviance influences their (moral) self-concepts and behavior directed at themselves. Finally, I identify different moderators that influence people's reactions to morally motivated deviance.

The Derogation of Moral Behavior

My investigation starts with the observation that people sometimes react particularly negative towards others who display morally motivated behavior (e.g., Minson & Monin, 2012; Monin, 2007; Parks & Stone, 2010). For example, recent studies show that people dislike others who refuse to perform a racist task (Monin, Sawyer, & Marquez, 2008). People also have negative evaluations of activists that try to achieve equality (Bashir, Lockwood, Chasteen, Nadolny, & Noyes, 2013) and of people who donate to charity (Newman & Cain, 2014). We sometimes even punish those who perform pro-social behavior (Barasch, Levine, Berman, & Small, 2014). Why does this happen?

1

2

3

4

5

6

&

1

2

3

4

5

6

&

As already explained, people are eager to evaluate their moral standing. Comparing one's own behavior with that of more moral exemplars may be especially painful because this process of upward social comparison in the moral domain jeopardizes people's view of themselves as moral individuals (e.g., Monin, 2007). As a consequence, people may react very negatively towards morally motivated deviants (Jamieson, Valdesolo, & Peters, 2014; Monin, 2007; Monin et al., 2008). So according to this line of reasoning, people dislike morally motivated deviants because they feel threatened by them.

Some of the first empirical evidence to support this argument comes from Monin and colleagues (2008). In their research, people were instructed to pick the most likely culprit out of several crime suspects. After making their choice they learned that another participant had refused to engage in the task. This other participant had refused to do so because he/she stated the task "was obviously biased" given that the most likely suspect was also the only Black person (Monin et al., 2008, Experiment 2). In the Monin et al. experiment, only half of the participants engaged in the task themselves. The other half merely read the reaction of the refuser. Finally, participants indicated how they evaluated the refusing confederate. Interestingly, participants who engaged in the task themselves (i.e., actors) derogated the refuser, while mere observers had very positive evaluations of this refuser. Apparently, actors felt threatened by the superior moral behavior of the refuser and reacted defensively by derogating the refuser, while observers did not. This observation also demonstrates again that how people react in actual situations is different from how they react to hypothetical situations.

Further indirect support for threat as the underlying mechanism was provided by demonstrating that participants who engaged in self-affirmation did not derogate the refuser. Self-affirmations are reflections on important values and are argued to buffer people's self-concepts (e.g., Cohen & Sherman, 2014). Arguably, when people's self-concepts are buffered—or strengthened in some way—the confrontation with the refusing participant is not as threatening, and therefore does not evoke the need for defensive responding. This is exactly what was found by Monin and colleagues (2008, Experiment 4). Participants who were self-affirmed did not derogate the refusing confederate, while participants without self-affirmation did derogate the refusing confederate. Apparently, a threatened self-concept underlies the derogatory responses.

Although the work by Monin and colleagues (2008) was groundbreaking and inspirational, some important questions remain unanswered. For example, what does moral threat mean exactly and can it be measured more directly? One possibility is to operationalize

threat in line with the bio-psycho-social model (BPSM) of challenge and threat as proposed by Blascovich and colleagues (see e.g., Blascovich & Mendes, 2010). According to this perspective, threat is a motivational state that arises when people experience a self-evaluative situation (i.e., a situation where they will be evaluated by themselves or others on a dimension that is important to them), and feel that they do not have the capacity to cope with this situation. For instance, it is important for people to portray themselves as moral, both in their own eyes and in the eyes of others (see e.g., Cohen & Sherman, 2014). People experience threat when they do not feel they are able to convincingly portray themselves as a moral person (either to themselves or to others), for example, because they were involved in behavior that is morally questioned by someone else. Challenge, the motivational counterpart of threat, is experienced when people feel they are capable to convincingly portray themselves as a moral person.

Importantly, according to the BPSM, threat and challenge are motivational states that can be measured by assessing certain important physiological indices, such as blood pressure and total peripheral resistance (i.e., TPR, the resistance in people's veins). More specifically, when people experience threat, their physiological response is comparable to a lock-down. This means that their blood pressure increases combined with a heightened resistance in their veins. This makes it more difficult for the body to efficiently transport energy: a defensive response (Blascovich & Mendes, 2010). When people experience challenge, on the other hand, their physiological response is more adaptive: blood pressure increases while the resistance in the veins decreases. This makes it possible for energy to be transported through the body efficiently. Operationalizing threat as a motivational state where the situational demand outweighs people's resources makes it possible to obtain physiological indices of threat and to provide direct evidence that a confrontation with a morally motivated deviant induces threat. This is exactly what I did in Chapter 2, as I will explain in more detail below.

Furthermore, another question that needs to be answered is whether dislike of morally motivated deviants occurs because they are deviants or because they are *morally motivated* deviants. For example, in the work by Monin and colleagues (2008), the reactions towards confederates who refused the assignment out of moral reasons are compared with reactions towards compliant confederates. Although it has been argued that people's negative reactions occur because the refusing confederate had moral reasons for the refusal, it cannot be ruled out that the negative reactions are due to the refusal alone. In fact, there is an abundance of work demonstrating that people have negative reactions to deviants (Abrams, Marques, Bown, & Henson, 2000; Chekroun & Nugier, 2011; Hichy, Mari,

1

2

3

4

5

6

&

1

2

3

4

5

6

&

& Capozza, 2008; Mendoza, Lane, & Amodio, 2014; Parks-Stamm, 2013; Parks & Stone, 2010; Pinto, Marques, Levine, & Abrams, 2010; Travaglino, Abrams, De Moura, Marques, & Pinto, 2014). In order to provide solid evidence for the claim that negative reactions are caused by the moral character of the refusal (see e.g., Monin, 2007; Monin et al., 2008), it would be necessary to compare reactions towards behaviors that are equally deviant, but for different reasons. For example, by comparing people's reactions towards someone who refuses a task out of moral reasons with reactions towards someone who refuses the same task out of non-moral reasons.

In Chapter 2, I answer these questions using the sausage-tasting experiment described above. In two experiments, people were invited to participate in a tasting task and instructed to taste a sausage. Afterwards, they were confronted with the reaction of a confederate who either refused to taste the sausage out of moral reasons, or out of non-moral reasons (i.e., because he/she did not like the taste of meat). People's evaluations of the other participant were then assessed. In Experiment 2.1, threat is assessed on a physiological level by measuring the physiological indices of challenge and threat, as prescribed by the BPSM of challenge and threat (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003; Scheepers, 2009). In this experiment, I demonstrate that threat occurs only when people are confronted with the confederate who has moral reasons for his/her refusal and not when this confederate has non-moral reasons. Furthermore, participants dislike the confederate who provided moral reasons for not eating the sausage more than the confederate who provided non-moral reasons to refuse the sausage. In my view, this finding implies that a confrontation with a morally motivated deviant does indeed induce threat as well as dislike, but only when the deviant has moral reasons for his/her deviance. Thus, in Chapter 2, I provide direct empirical evidence that people experience threat when they are exposed to a morally motivated deviant and that threat only occurs when the deviance is morally motivated.

Consequences for the Self-Concept

Threat can also be conceptualized as a negative impact on one's self-concept (see e.g., Sherman & Cohen, 2002, 2006). In fact, it has been argued that the negative reactions to morally motivated deviants occur because these deviants pose a threat to people's self-concepts (see e.g., Monin, 2007; Monin et al., 2008). This implies that exposure to someone who makes a moral claim should lead people to feel less positive about themselves. There is some indirect evidence to support this line of thought. For instance, Monin and colleagues (2008) demonstrate that people's negative reactions towards others who have moral claims

can be explained by the fact that people expect these morally motivated others to dislike them. Thus, these authors found that the dislike of morally motivated deviants is mediated by anticipated dislike.

Other support comes from Minson and Monin (2012). These authors used a scenario where meat eaters were invited to evaluate vegetarians on several dimensions, such as kindness and intelligence. Half of the participants first thought about how they would be evaluated by vegetarians, and afterwards evaluated the vegetarians, whereas the other half immediately evaluated vegetarians without first thinking about how they would be evaluated by the vegetarians. It was found that participants had more negative evaluations of vegetarians when they first thought about how they would be evaluated by vegetarians. According to Minson and Monin, letting participants contemplate about how they would be evaluated by these vegetarians was threatening for participants, because the participants would expect their dietary choices to be seen as immoral by the vegetarians. This potential rejection led them to derogate the vegetarians in return. These studies demonstrate that negative reactions towards morally motivated deviants occur when people evaluate their own relative moral standing.

However, although there are indications that exposure to morally motivated others threatens people's self-concepts, there is little direct support that people's explicit-self concepts are affected by these situations. One way of obtaining direct evidence would be to assess people's explicit evaluations of themselves. For example, if people's self-concepts are indeed threatened by the exposure to the morally motivated deviant, you would expect them to have lowered self-evaluations in these situations. For instance, people would be less happy with themselves, more disappointed in themselves, etcetera. In both Chapters 2 and 5, I assess participants' evaluations of themselves and demonstrate that explicit self-evaluations are indeed lowered after exposure to morally motivated deviance. Thus, in this way, I provide direct support in this dissertation that exposure to the moral behavior of others threatens participants' explicit self-concepts.

Behavioral Reactions

So far we know that exposure to the moral behavior of others can lead people to dislike these others and feel threatened by them. But would people take this negativity one step further and engage in actual negative behavior directed at the other? This is an important question, especially in a time where actual behavior is understudied in psychological science (see e.g., Baumeister, Vohs, & Funder, 2007). Furthermore, we know from the literature that there can be a substantial gap between one's attitudes and one's behavior (see e.g., Baumert et al., 2013). This inconsistency between what one thinks and does has

1

2

3

4

5

6

&

1

been termed the ‘attitude-behavior gap’ or ‘intention-action gap’ (see e.g., Webb & Sheeran, 2006). Therefore, it is important to investigate whether the negative reaction people experience after exposure to morally motivated deviants translates into actual behavior.

2

3

4

5

6

&

This is what I investigate in Chapter 3. Here, I use a similar set-up as described above, but now in the domain of negative reactions to non-drinkers. What could be a behavioral measure that would indicate negative behavior towards a non-drinker? It is known that drinking alcohol is one of the biggest problems in society nowadays and that alcohol scores in the top 3 of most harmful substances to individuals and society (Dutch National Institute for Public Health and the Environment [RIVM], 2009). A very important reason why people drink alcohol is because of peer influence, as a result of which non-drinking peers are pressured into drinking (see e.g., Simons-Morton, Haynie, Crump, Eitel, & Saylor, 2001). Thus, providing non-drinkers with alcoholic beverages is harmful behavior that aggravates an important societal problem. Therefore, I use the allocation of beer to a non-drinker as a negative behavior directed at the other person in Chapter 3.

In Experiment 3.1, participants are first instructed to taste beer—in reality, non-alcoholic beer. After tasting, they are exposed to the bogus reaction of a peer who has refused to drink beer out of moral reasons (i.e., because drinking alcohol is wrong) or non-moral reasons (i.e., because the taste of alcohol is gross). Thus, participants are exposed to a moral or non-moral refuser. The most important outcome measure is how much beer participants allocate to this refuser. In Experiment 3.1, I demonstrate that participants allocate more beer to a moral refuser than to a non-moral refuser. I interpret this as evidence that people show more negative behavior towards a moral refuser. Viewed in this way, this observation extends earlier findings that participants negatively evaluate a moral refuser and shows that people engage in actual negative behavior directed at the moral refuser.

Another important question is whether negative self-directed attitudes (e.g., a lowered self-concept) also translate into actual behavior. According to the literature, people will engage in compensation behavior after having performed transgressions, in order to make up for their moral missteps (e.g., Zhong et al., 2009) and prove their moral worth (e.g., Cohen & Sherman, 2014). One way of achieving this may be by experiencing physical pain (e.g., Nelissen, 2012; Nelissen & Zeelenberg, 2009). It has been argued that inflicting and enduring pain can demonstrate one’s moral character (Bastian, Jetten, & Stewart, 2013). This may explain why people sometimes inflict pain on themselves after transgressions (Inbar, Pizarro, Gilovich, & Ariely, 2013; Nelissen & Zeelenberg, 2009). Following this reasoning, I argue that enduring pain can be an effective response to exposure to a morally motivated

deviant because it reaffirms one's moral character (e.g., Bastian et al., 2013; Cohen & Sherman, 2014). If this assumption is correct, this means that people would inflict and endure more pain to themselves when they are confronted with a morally motivated deviant rather than a non-morally motivated deviant.

I test this hypothesis in Experiment 3.2, that uses a similar set-up as Experiment 3.1: Participants are first instructed to drink beer—in reality, non-alcoholic beer—and are exposed to the (bogus) reaction of a peer who refused to drink beer out of moral or non-moral reasons. Hereafter, participants complete a cold-presser task where they are instructed to keep their hand submersed in ice-cold water (2 °C) for as long as possible. This cold presser test (CPT) induces physical pain (Bastian, Jetten, & Fasoli, 2011; Bastian et al., 2013; Chen, Poon, & DeWall, 2015; Jackson, Iezzi, Nagasaka, Fritch, & Gunderson, 2002). I demonstrate that people indeed keep their hands submersed for a longer period of time when they are exposed to a morally motivated deviant as opposed to a non-morally motivated deviant. Thus, in Chapter 3, I show that when people are placed in a situation where they are exposed to the morally motivated behavior of someone else, this can lead to negative behavior directed at the other person, as well as negative behavior directed at the self.

Threatened by the Identity of Morally Motivated Deviants

Thus far, I have described and investigated situations where people reacted negatively to morally motivated deviants because the deviants' behavior reflects poorly on people's own behavior. More specifically, in Chapter 2, the morally motivated deviant refuses to eat meat after participants tasted meat themselves; in Chapter 3, the morally motivated deviant refuses to drink beer after participants drank beer themselves. In these chapters, we will learn that morally motivated deviants can become the targets of derogation, dislike and negative behavior when they refuse to perform behavior that the target did perform. But people do not necessarily have to engage in certain behaviors themselves to feel threatened by the moral behavior of someone else.

Could it be that the positivity or negativity of the reaction to the moral behavior of someone else depends on characteristics of the person who performs the moral behavior? Perhaps people can also be threatened by the mere identity of the person showing the moral behavior. In Chapter 4, I will demonstrate that even behavior that would otherwise be perceived as highly moral and laudable can be derogated when it is displayed by in-group deviants.

According to the theory of Subjective Group Dynamics (SGD; Abrams et al., 2000; Marques, Abrams, Paez, & Martinez-Taboada, 1998), people want their group to be different from

1

2

3

4

5

6

&

1

2

3

4

5

6

&

other groups and want their group members to adhere to the group norms. Because of this, people negatively react to deviant group members, because these deviants undermine the distinction between their group and others, as well as the validity of their group norms. Furthermore, moral norms are especially important for the in-group (e.g., Ellemers, Pagliaro, Barreto, & Leach, 2008; Pagliaro, Ellemers, & Barreto, 2011). As a consequence, moral behavior may be derogated by the group when it is performed by in-group deviants.

For example, imagine how conservative Christians would evaluate a couple that wants to adopt an orphan child. Probably, to most Christians, the adoption of this orphan would seem like a very moral and praiseworthy thing to do. But what if the adoptive parents are lesbians? In this case, the lesbian couple would deviate from important norms that are part of some extreme Christian in-group values (e.g., “men and women are destined to be together”) and would thereby undermine the in-group’s positive identity. Importantly, the fact that this couple undermines important values for conservative Christians should be especially clear when Christian norms are made salient. One way of increasing the salience of these Christian norms could be when the couple identifies as Christian as well.

In Chapter 4, I investigate this scenario by asking conservative Christians in the Dutch bible belt how they would evaluate this prospective adoption. I manipulate whether the adoptive parents to be are a heterosexual or a lesbian couple and whether this couple explicitly identifies as Christian or not. I demonstrate that participants indeed derogate the adoption when it is performed by in-group deviants (i.e., a Christian lesbian couple), while they praise it when it is performed by normative in-group members (i.e., a Christian heterosexual couple). Thus, behavior that is considered as very moral can be derogated when it threatens people’s social identity and undermines norms and values.

Protecting and Restoring the Self-Concept

The picture I have painted thus far is somewhat dark and gloomy: Apparently, threats to people’s fragile sense of moral worth leads them to lash out against others who take a moral stance and stick to their beliefs. Would it be possible to somehow protect or restore people’s self-worth and prevent negative reactions?

According to self-affirmation theory, there are several ways in which people can respond to moral threat (e.g., Cohen & Sherman, 2014). One of these ways is defensive responding which leads people to derogate or discredit the source of the threat posed by the morally motivated deviant, and thereby distort the threatening event (e.g., Cohen & Sherman,

2014; Sherman & Cohen, 2002; 2006). This is the type of reaction that we have seen so far in the research that I have conducted and described. For example, disliking someone who refuses to eat meat out of moral concern (Chapter 2) or allocating beer to someone who explicitly does not want to drink alcohol (Chapter 3). However, it is also possible to protect one's self-concept in other ways and thereby prevent defensive responses. One of the ways to do this is by engaging in self-affirmations (see, e.g., Cohen & Sherman, 2014). This is often operationalized as reflecting on one's important values. This reflection leads people to realize that their self-worth does not solely depend on the situation they are in at that moment. This helps people to deal with the threatening situation in a more open way, and reduces the need to distort or disqualify the source of the threat (Sherman & Cohen, 2006). This means that self-affirmation or other ways to protect the self-concept should prevent negativity after exposure to a morally motivated deviant.

Thus, following this line of reasoning, exposure to a morally motivated deviant should not evoke negative behavior when people have engaged in self-affirmation. To test this prediction, I also include a self-affirmation manipulation in Experiment 3.1. So after drinking beer, but before reading the reaction of the other person, half of the participants engage in a self-affirmation task, whereas the other half of the participants engage in a neutral task. I indeed demonstrate that while participants in the neutral condition allocate more beer when they are exposed to a moral non-drinker than to a non-moral non-drinker, this is not the case for participants who are self-affirmed. Thus, protecting people's self-concepts by engaging in self-affirmation reduces their consequent negative behavior in a self-threatening situation.

There are also other ways to protect one's self-concept. One intriguing alternative is to engage in physical cleansing (e.g., Zhong & Liljenquist, 2006), which literally means cleansing (parts of) one's body. There are two reasons why physical cleansing could protect one's self-concept. The first is that abstract topics have become linked to concrete concepts that are present in the physical world (Landau, Meier, & Keefer, 2010). This means that the complex and abstract concept of morality has been linked to the concrete experience of cleanliness. Furthermore, this also implies that influencing the concrete concept changes the abstract topic as well. Thus, when people engage in the physical activity of cleansing themselves, they also increase their abstract sense of morality. There is some evidence to support this claim. For example, Zhong and colleagues (2010) demonstrate that when participants imagined being clean, they also felt more moral.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

Another reason why physical cleansing may protect one's self-concept is that it removes physical and psychological residues of past actions (Lee & Schwarz, 2010, 2011). Thus, physical cleansing can be used to 'wash away' one's past. Because of these reasons, engaging in physical cleansing after displaying a certain behavior should protect one's self-concept and thereby prevent negative consequences.

In Chapter 2, I investigate whether physical cleansing can indeed prevent negative consequences of exposure to moral vegetarians. More specifically, in Experiment 2.2, participants are asked to taste a sausage. Afterwards, half of the participants engage in physical cleansing by washing their hands with antibacterial soap while the other half does not. Hereafter, they are exposed to the reaction of a peer who has refused to eat the sausage out of moral or non-moral reasons. I demonstrate that when participants did not engage in physical cleansing, they show more dislike when they are exposed to a moral refuser rather than a non-moral refuser. However, when participants engage in physical cleansing, they do not show more dislike for a moral refuser rather than a non-moral refuser. Thus, engaging in physical cleansing reduces people's consequent negative evaluations in a threatening situation.

Finally, there may also be situational cues that increase or reduce the threat to one's self-concept when exposed to a morally motivated deviant. One of these situational cues may be the visibility of one's behavior. It is known that people are very motivated to present themselves favorably to others (for a review, see, e.g., Geen, 1991). This leads people to display praiseworthy behavior and refrain from displaying wrongful behavior when it is visible to others. For example, people are more inclined to engage in moral transgressions in private settings than in public settings (Greene & Lowe, 2014) and show more desirable behavior when they feel they are being watched (Bateson, Nettle, & Roberts, 2006). This could also mean that people's self-concepts would be most threatened when they perform wrongful behavior in a setting that is visible to others, but not so much when it is not so visible to others.

I test this prediction in Chapter 5. Here, I use a setting where participants are asked to write an anti-gay essay that explicitly denounces equal rights for same-sex couples. This is something that goes against the values of Dutch University students, who in general endorse equal rights for same-sex couples (The Netherlands Institute for Social Research [SCP], 2013, 2016). After writing their own anti-gay essay, they are exposed to the moral or non-moral refusal of someone else to do the same. Furthermore, I also manipulate whether participants thought their essay the experimenter would read their essay or not. I demonstrate that people's self-concepts are only lowered when they believe their essays

would be read by the experimenter and they are exposed to the morally motivated refusal of someone else. Apparently, a setting that reduces the visibility of one's behavior may also prevent negative consequences of one's behavior for one's self-concept.

Positive Reactions to Morally Motivated Deviance

So there are several ways in which people's self-concepts can be protected and that prevent negative reactions to the morally motivated behavior of others. But are there also situations where people not only refrain from negative reactions but actually appreciate moral character of those who stick to their beliefs?

Would it not be great if people would recognize and value the moral worth of those who do the right thing? There is some research that witnessing the moral exemplary behavior of others can inspire people to do good—a state called moral elevation (e.g., Aquino, McFerran, & Laven, 2011; Silvers & Haidt, 2008). For example, breastfeeding mothers who watched a morally elevating video were more likely to nurse and hug their babies as compared to breastfeeding mothers who watched an enjoyable comedy video (Silvers & Haidt, 2008). Aquino and colleagues (2011) demonstrated that watching a morally exemplary music video where money was donated to charitable causes (vs. watching a neutral music video) led people to donate their own money to a charitable cause as well.

But here is the catch. In these and other studies, people were unvested in the morally exemplary situation which means they had not acted (or failed to act) in this same situation. This is crucial because, when people themselves have failed to show moral behavior in the same situation, the morally exemplary behavior of the other person constitutes a source of threat rather than elevation (see e.g., Chapter 2 and Chapter 3; see also Franco, Blau, & Zimbardo, 2009; Monin et al., 2008). Thus, the real challenge lies in identifying situations where people can appreciate the moral character of others, even when their own behavior is not up to par.

When would people appreciate a morally motivated deviant in a situation where their behavioral 'failure' is clear? When you think about the situations that I study in Chapters 2 and 3, there is one common characteristic: The behavior the morally motivated deviant opposes (i.e., eating meat or drinking beer) is not blatantly 'wrong'. The same goes for the studies by Monin and colleagues (2008), where participants took part in a 'police task' and had to choose the most likely culprit out of a series of suspects. Again, in this study, participants probably did not think their behavior (i.e., picking a suspect) was clearly immoral. This implies that participants probably did not realize that there could be anything

1

2

3

4

5

6

&

1

wrong with the behavior they just performed. If so, the behavior of the morally motivated deviant may have served the purpose of delivering an unwelcome message to participants: That their behavior was immoral, or at least seen as immoral by someone else.

2

3

4

5

6

&

But what would happen when participants engage in behavior that they clearly realize is wrong, such as doing something that explicitly does not fit with their moral values? In that case, if they would be exposed to someone else who would have refused to perform this counter-attitudinal behavior, this would not send the same unwelcome new message about their own behavior because it would already be clear that people's own behavior was not up to standard. In this case, people may not react negatively to these morally motivated deviants, regardless of the effect of this situation on their self-concept. In fact, perhaps this type of situation may even evoke moral elevation and may lead people to appreciate those who stick to their moral beliefs.

This is what I investigate in Chapter 5. In three experiments, participants engage in a task that is clearly counter-attitudinal. More specifically, Dutch participants are instructed to write an essay denouncing equal rights for homosexuals. This is something that is explicitly counter-attitudinal for most Dutch people, given that around 90% of Dutch people endorses equal rights for same-sex couples (The Netherlands Institute for Social Research [SCP], 2013, 2016). I verified that the task was explicitly counter-attitudinal for the participants in these experiments: Agreement with this task was very low (e.g., between 11 and 18 on a 100-point scale ranging from 0 to 100). After writing these essays, participants are exposed to either the reaction of a person who refused to write the essay out of moral reasons (i.e., because it is wrong to write this essay and homosexuality is normal and moral) or a person who based his/her refusal on a non-moral reason. What I consistently find in these three experiments is that people have positive evaluations of the morally motivated deviant who refuses to write an anti-gay essay out of moral reasons. This occurs even in situations where people's explicit self-concepts are negatively affected, and they thus experience a threat to their self-concepts. This is important because it shows that even in situations where people are personally involved and implicated, they can appreciate the moral character of others who stick to their moral beliefs when their own behavior was blatantly counter-attitudinal.

Overview

The introduction of the different research projects that I have conducted in this dissertation is almost complete. In Chapter 2, you will learn that meat-eating participants dislike, and are threatened by, a moral vegetarian. Furthermore, participants also negatively evaluate themselves after exposure to a moral vegetarian. Protecting the self-concept by engaging in physical cleansing prevents these negative responses. In Chapter 3, I show that beer-drinking participants allocate more beer to a non-drinking peer who has moral reasons not to drink, and inflict and endure more pain to themselves after this confrontation with a moral non-drinker (vs. a non-drinking peer with non-moral reasons). Furthermore, bolstering the self-concept by engaging in self-affirmation prevents the negative effect on beer allocation. In Chapter 4, I demonstrate that people can even derogate behavior that would otherwise be seen as laudable when it is performed by an in-group deviant: Conservative Christians negatively evaluate the adoption of an orphan when the adoptive parents are a lesbian couple, especially when this couple explicitly identifies as Christian. In Chapter 5, different factors that influence the effect of exposure to morally motivated deviants on the self-concept are highlighted, such as actively engaging in counter-attitudinal behavior, the salience of one's behavior to others, and the salience of morality. Furthermore, I demonstrate that people can appreciate the morally motivated refusal of others in situations where the moral quality of this behavior is clear. Finally, in Chapter 6, I will reflect on all these findings.

After reading these chapters, I hope to have convinced you that there is something special about morally motivated deviance which makes people react strongly. I will try to show that morally motivated deviance (when compared to non-moral deviance) evokes threat, dislike, negative evaluations of oneself, negative behaviors directed at the other and the self, and in some cases also positive evaluations. I investigate this topic in a range of different studies where people are placed in actual, immersive situations—with Chapter 4 being the exception to this rule—and react to issues concerning daily morality. I hope you enjoy reading about this research as much as I have enjoyed conducting it.

1

2

3

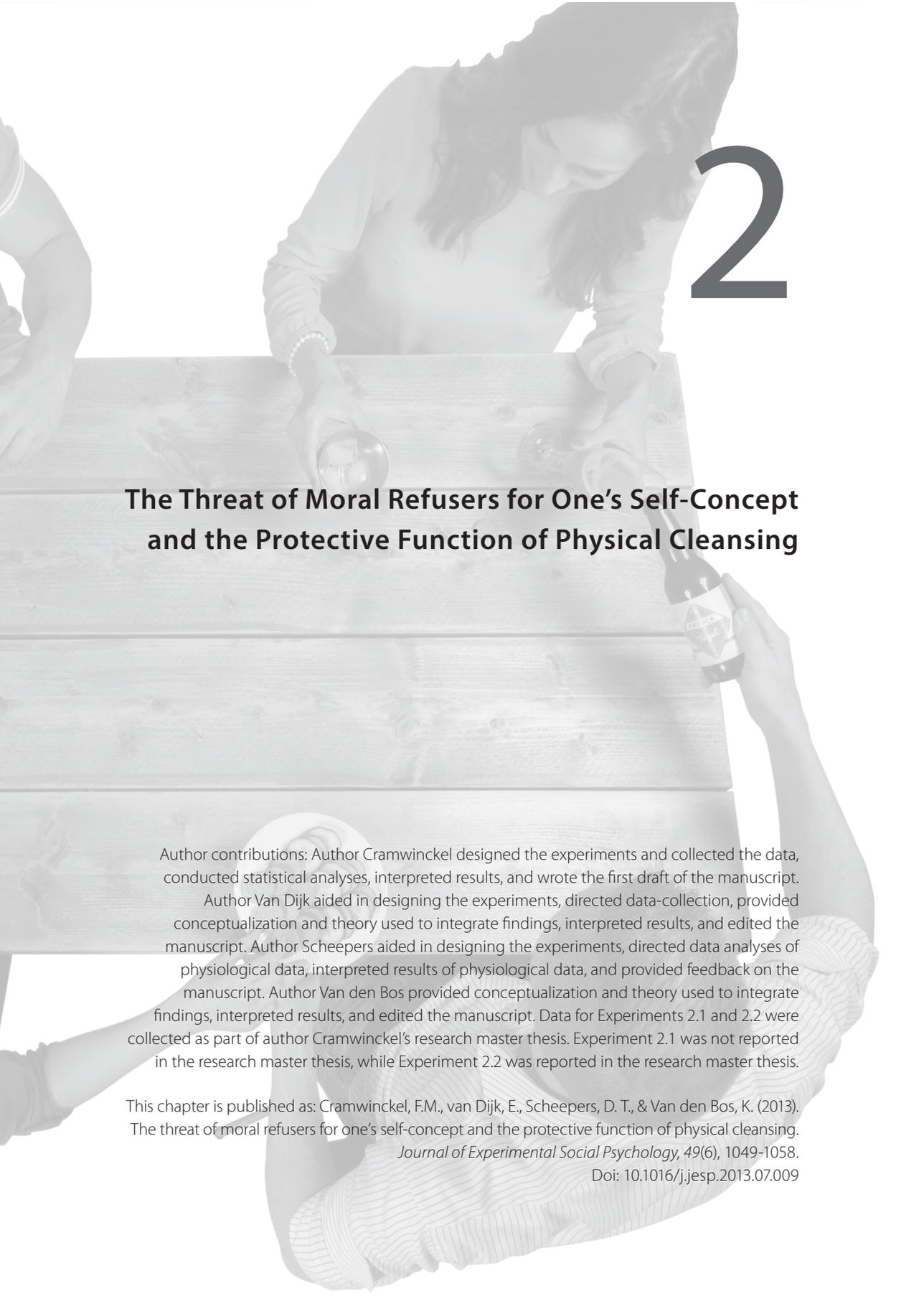
4

5

6

&





2

The Threat of Moral Refusers for One's Self-Concept and the Protective Function of Physical Cleansing

Author contributions: Author Cramwinckel designed the experiments and collected the data, conducted statistical analyses, interpreted results, and wrote the first draft of the manuscript.

Author Van Dijk aided in designing the experiments, directed data-collection, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Author Scheepers aided in designing the experiments, directed data analyses of physiological data, interpreted results of physiological data, and provided feedback on the manuscript. Author Van den Bos provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Data for Experiments 2.1 and 2.2 were collected as part of author Cramwinckel's research master thesis. Experiment 2.1 was not reported in the research master thesis, while Experiment 2.2 was reported in the research master thesis.

This chapter is published as: Cramwinckel, F.M., van Dijk, E., Scheepers, D. T., & Van den Bos, K. (2013). The threat of moral refusers for one's self-concept and the protective function of physical cleansing.

Journal of Experimental Social Psychology, 49(6), 1049-1058.

Doi: 10.1016/j.jesp.2013.07.009

Abstract

1

2

3

4

5

6

&

We investigated how people respond to moral threats and the consequences this has for one's moral self-concept. In two experiments, participants first tasted a sausage and were then confronted with a bogus participant who had refused to taste the sausage because of moral or non-moral reasons. People disliked the moral refuser more than the non-moral refuser. The self-threatening effect of having one's morals questioned was also reflected in specific patterns of cardiovascular responses and negatively affected participants' self-evaluations. We further showed that the negative effects of a moral threat can be prevented by a simple intervention of physical cleansing: Participants who had cleansed their hands before being confronted with a moral refuser did not show the negative effects on self- and refuser evaluations. Importantly, the protective effects of physical cleansing were most pronounced for people with a strong moral identity. Taken together, these results underline the importance of one's self-concept when confronted with a moral refuser, and introduce an effective intervention to prevent these negative consequences.

Morality is one of the most important dimensions of self-definition. People learn already at a young age that just and moral behaviors are desirable and admirable (Lerner, 1980). Being moral is often so important to people that they go to great lengths to preserve their image as moral beings. In fact, this can lead people to derogate others who pose a threat to the moral self. As a result, a confrontation with a *moral refuser* (someone who, out of moral concern, refuses to go along with a certain behavior) can have severe consequences for the way people evaluate themselves and this other person. For example, Minson and Monin (2012) demonstrated that when non-vegetarians were asked to think about how vegetarians would evaluate them, they responded by negatively evaluating vegetarians. In the current research, we aim to provide further support for the idea that being confronted with a moral refuser has negative consequences for how you evaluate yourself and the moral refuser.

More specifically, with two experiments we aim to better understand the negative consequences of a confrontation with a moral refuser and to provide a possible intervention to prevent these consequences. Hereby, we extend existing knowledge in three ways. First, we provide converging evidence that being confronted with a moral refuser is indeed threatening, and that this threat can be discerned at the cardiovascular level (Blascovich & Tomaka, 1996; see also Blascovich & Mendes, 2010). Second, we demonstrate how a relatively simple intervention, physical cleansing, can prevent the negative self- and other- directed consequences of a confrontation with a moral refuser. Third, we show the important role that one's moral self-concept plays in dealing with moral refusers. That is, we demonstrate that especially people whose self-concept is centered around morality (i.e., those with a strong moral identity) benefit from physical cleansing when being confronted with a moral refuser.

Dealing with Moral Threat

According to moral self-regulation theory (Zhong, Liljenquist, & Cain, 2009), people strive for a positive moral balance; that is, people want to view themselves as good and moral individuals (Sherman & Cohen, 2002, 2006). Morality is therefore a very important aspect of one's self-concept (Osgood, Suci, & Tannenbaum, 1957; Pronin, 2008). Besides being one of the most central aspects of the self-concept, morality is also fundamentally different from non-moral aspects of the self-concept (Skitka, Bauman, & Sargis, 2005). Moral statements are different from personal preferences partly because moral convictions are, or could be, universal (Haidt, Rosenberg, & Hom, 2003). This implies that moral statements tend to apply to everyone, and not only to the person who makes a statement about certain acts to be moral or immoral. This makes a confrontation with a moral refuser especially threatening, because the refuser's moral stance may imply that the behavior you engaged in was not as moral as you would like it to be.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

People can deal with this moral threat in several ways. One possible way is to change their perception of the person or event that threatens their moral balance, for example by derogating the source of the threat (Sherman & Cohen, 2002, 2006). This has been demonstrated in previous research (e.g., Minson & Monin, 2012; Monin, Sawyer, & Marquez, 2008). For example, Monin and colleagues (2008) asked participants to choose the most likely suspect for a crime from a set of possible subjects. Later, participants were confronted with another participant who had refused to do the same task because it ostensibly was racist. Afterwards, the participants evaluated this moral refuser. People reacted to moral refusers by derogating them and apparently felt threatened by the implicit stain on their own morality.

We performed two experiments that directly build upon the work of Monin and colleagues (2008). In both experiments, participants first ate a sausage and were then confronted with another participant who had refused to eat a sausage either based on moral grounds (the moral refuser condition) or on non-moral grounds (the non-moral refuser condition). In the experiments, we provide further empirical support for the notion that, in accordance with previous research (e.g., Minson & Monin, 2012; Monin, 2007; Monin et al., 2008), being confronted with a moral refuser leads to a negative evaluation of the refuser and has a negative impact on one's self-concept.

To obtain converging evidence that moral refusers elicit threat, in Experiment 2.1 we included cardiovascular indices of threat and challenge motivational states on the basis of the biopsychosocial model (BPSM; Blascovich, 2008; Blascovich & Mendes, 2010; Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003; Blascovich & Tomaka, 1996). The BPSM applies specifically to *motivated performance situations* such as giving a speech for a moral refuser. A core characteristic of a motivated performance situation is that such a situation is self-evaluative, i.e., an individual must believe that he/she will be evaluated, by others or by his/herself, on a domain that is important for the self, such as morality. Importantly, the same situation (e.g., giving a speech) can be challenging for some people, but threatening for other people, depending on how they evaluate themselves and the situation (Blascovich & Mendes, 2010).

According to the BPSM, people display a motivational state of threat when the demands of the motivated performance situation outweigh the resources the person brings into the situation to deal with these demands. By contrast, a motivational state of challenge emerges when personal resources outweigh situational demands. Because it is important for people to be and appear moral (e.g., Pronin, 2008; Zhong et al., 2009), we argue that

giving a speech to a moral refuser is more likely to lead to threat than to challenge. The confrontation with a moral refuser may lead people to realize that they have done something immoral (eating meat), which may make them feel ill equipped to portray themselves as moral, and thus to perform well in the speech-task. This confrontation with a moral refuser therefore creates a situation where the situational demands (giving a speech to a moral refuser) outweigh one's personal resources (such as the belief that one can convince others that one is, in fact, a moral person), and thus would constitute a motivational state of threat rather than challenge.

Apart from providing converging evidence that moral refusers elicit threat, it should be noted that the use of cardiovascular indices to capture threat has some additional advantages above traditional methodology (e.g., self-report measures). That is, the cardiovascular indices described by the BPSM (which will be described in more detail in the Method section of Experiment 2.1) can provide online, continuous, and unobtrusive measures of threat, and its motivational counterpart, challenge. These cardiovascular indices can also pick up on motivational processes that a person is not aware of, and are non-susceptible to self-presentation concerns, which could normally play a role in these types of settings that concern areas of the self that are important to people (e.g., morality).

Besides using physiological measures, we also extend previous research by including a non-moral refuser comparison group in our research designs. By doing so, we are able to rule out an alternative explanation for the rejection of moral refusers that has not been addressed in previous research, namely that people simply reject others who behave differently from them, regardless of the rationale for this deviant behavior. Previous research has typically compared people's reactions to moral refusers (who refused out of moral concern) and compliant others (who went along with the task; e.g., Monin et al., 2008). Although it has been presumed that people derogated the moral refusers because they were *moral* refusers, it is also possible that people reacted negatively to them because they were *refusers*. We aim to provide more empirical evidence for the process that underlies the negative consequences of being confronted with a moral refuser, by disentangling the effects of moral refusal and the effects of refusal alone.

Physical Cleansing

Physical cleansing is literally cleansing (parts of) one's body, such as one's hands. Physical cleansing provides an intriguing means to protect one's threatened morality (Zhong & Liljenquist, 2006). There are two ways in which physical cleansing can help to prevent the negative consequences of a confrontation with a moral refuser. First, physical cleansing

1

2

3

4

5

6

&

1

2

3

4

5

6

&

can create a moral buffer. That is, people link complex ideas and topics to easier concepts that are present in the physical world (Landau, Meier, & Keefer, 2010). In doing so, morality is often linked to physical cleanliness. This is, for example, clear in language: People have to 'come clean' and have 'a clean conscience.' Importantly, not only do people think about abstract topics in physical terms, these physical metaphors such as cleanliness can also influence people's judgments about, and behaviors related to, the abstract topic. This metaphoric transfer (Landau et al., 2010) implies that when people's physical sense of cleanliness is altered, their moral judgments and behaviors are also influenced. This would mean that if you literally feel clean (e.g., after cleaning your hands), you would also feel more moral. Empirical research provides support for this idea: Zhong, Strejcek, and Sivanathan (2010) demonstrated that when participants imagined being clean, they also felt more moral, which also translated into their moral judgment and decision-making.

The second reason why physical cleansing can prevent the negative consequences after a confrontation with a moral refuser is the 'clean-slate' logic, put forth by Lee and Schwarz (2010, 2011). According to these authors, physical cleansing not only creates a moral buffer but also removes the physical and psychological residues of past actions. For example, if one would clean one's hands after eating a greasy sausage, this act of cleansing would not only remove the physical residues of eating the sausage (e.g., grease and smell) but also the psychological residues of eating the sausage (such as feeling guilty after eating meat in a moral confrontation). Physical cleansing can thus be used to sever the link with one's past behavior.

In the current research, some of the participants cleansed their hands after a potentially objectionable behavior (eating a sausage), but before the confrontation with a moral refuser. In summary, physical cleansing could thus prevent the negative consequences of a confrontation with a moral refuser by providing a moral buffer and by removing the psychological residues of the potentially objectionable behavior.

Moral Identity

To further investigate the role of one's (threatened) moral self-concept after a confrontation with a moral refuser, we also investigate individual differences in the importance people attach to the moral self-concept. Although morality is central to the self-concept of most people (e.g., Pronin, 2008), there are important individual differences in the extent to which one's self-concept is organized around a set of moral traits (Aquino & Reed, 2002). Being a moral person and acting in line with this is more important for people with a strong moral identity than for people with a weaker moral identity. Research shows, for example, that

people with a strong moral identity are more likely to volunteer (Aquino & Reed, 2002) or to offer help to out-group members (Reed & Aquino, 2003; Shao, Aquino, & Freeman, 2008).

For the current research, it is also relevant to note that people with a strong moral identity are particularly sensitive to threats to their moral self (Reed & Aquino, 2003). Failure to live up to one's moral standards may be especially threatening to people with a strong moral identity (Aquino, Freeman, Reed, Lim, & Felps, 2009), because their self-concept is strongly built upon the notion that they are moral and good individuals. Therefore, we expect that being confronted with a moral refuser has the most severe consequences for the self-concept of people for whom morality is really important, while it may leave the self-concept of people with a weaker moral identity relatively unaffected. Physical cleansing may therefore have the most beneficial effects for people with a strong moral identity.

Experiment 2.1

In the first experiment, participants first tasted a sausage and were then confronted with a fellow participant who refused to partake in this part of the experiment out of moral or non-moral concerns. Afterwards, participants were asked to give a speech about their eating habits to the refuser, while we collected cardiovascular indices to assess threat and challenge following the BPSM (Blascovich & Mendes, 2010; Blascovich & Tomaka, 1996). Whether participants are challenged or threatened when giving the speech depends on whether they evaluate themselves as having the 'tools' to perform well in this situation, and thus on whether they will be able to come across as a good (and moral) person. The confrontation with a moral refuser may lead participants to question their own morality (e.g., Monin, 2007), because the moral refuser implicitly questions the morality of the participants' previous meat-eating in the experiment. Therefore, we expect that participants in the moral refuser condition are relatively more threatened than participants in the non-moral refuser condition, because they will have more difficulties portraying themselves as good and moral individuals, which is important to them (e.g., Pronin, 2008).

In addition to physiological indices of challenge and threat, we also measured self-reported evaluations of the refuser. We expected that a moral refuser would be evaluated more negatively than a non-moral refuser.

Method

Participants and design. Forty-eight participants ($M_{\text{age}} = 20.88$; 39 women) participated in an experiment on 'food preferences and personality' in which they allegedly had to judge

1

2

3

4

5

6

&

1

2

3

4

5

6

&

someone's personality on the basis of this person's food preferences. Participants received 6 Euros or course credit for their participation and were randomly assigned to the moral or non-moral refuser condition.¹

Procedure. Participants were seated in a cubicle, with a webcam and equipment for physiological recording. Following the procedures described by De Wit and colleagues (De Wit, Scheepers, & Jehn, 2012), four electrodes to measure impedance-cardiographic signals, two electrodes to measure electrocardiographic signals and a finger sensor to measure bloodpressure were attached to the participants. Then, participants were asked to answer some background questions (sex, age, their dietary status, etc.) and engaged in a short filler task.

¹ Please note that 13 additional participants completed this experiment but were excluded from the analyses; ten participants were excluded because of a priori criteria, and three participants were identified as multivariate outliers in the main analyses. We had two important a priori criteria to exclude participants from this research: not eating the sausage (e.g., being a vegetarian) and not realizing that the refuser had refused to taste the sausage. We therefore administered manipulation checks after the self-reported dependent variables to assess whether participants had tasted the sausage and correctly identified that the refuser had not tasted the sausage. Based on these a priori criteria, we thus excluded 4 participants who had not tasted the sausage, and 6 participants who thought the refuser had tasted the sausage, from the analyses. Although we reasoned that vegetarians were not eligible for this research, we did not use this a criterion in the recruitment of participants because we did not want to raise suspicion about the experiment's true purpose.

We calculated two distance measures in our main analyses to spot potential multivariate outliers: Cook's distance and Mahalanobis distance. Given that it can be difficult to detect multivariate outliers, it is useful to compute several measures to identify potential outliers and we therefore computed the two measures which can currently best detect multivariate outliers (Tabachnick & Fidell; 2007). Cook's distance is a measure of influence, which indicates the extent to which a case is in line with others and the extent to which a case is far from other cases. Mahalanobis distance *"is the distance of a case from the centroid of the remaining cases where the centroid is the point created at the intersection of the means of all the variables"* (Tabachnick & Fidell, 2007, p. 74). We used Cook's distance $> .10$ as the cut-off value. Following the guidelines by Tabachnick and Fidell (2007), we used a very conservative cut-off for the Mahalanobis value, which was $p < .001$ for the χ^2 value. We decided to exclude potential outliers whenever they exceeded either the Cook's cut-off value of .10, and/or the Mahalanobis distance of 10.13 ($p < .001$, based on 1 df). All Mahalanobis distances were well below the cut-off value of 10.13 (all $\chi^2 < 1.02$). Cook's distances revealed three outliers (Cook's distances $> .10$), which were excluded from the analyses.

Finally, we want to note that thirty-nine additional participants were initially recruited for Experiment 2.1, but we were not able to analyze their data because of technological difficulties (for thirty-one participants not all cardiovascular data were registered, for eight participants the cardiovascular data were impossible to score reliably because of poor impedance cardiograph quality; because of a loose electrode).

Hereafter, baseline measures of the cardiovascular indices would be taken (see below) while participants were asked to sit still and relax while watching a five-minute neutral aquatic movie (De Wit et al., 2012).

After baseline measures were collected, participants engaged in the food tasting task which consisted of tasting a slice of cucumber and a slice of sausage. Participants were instructed to taste the slice of cucumber, take their time to taste it, and to write down what they experienced while tasting it by answering the following open questions: *"How does it taste?"*, *"what sensation comes to mind?"*, *"what kind of thoughts appear?"*, *"how do you feel?"*. After answering the questions for the slice of cucumber, participants were instructed to taste the slice of sausage, take their time to taste it, and to answer the same questions about the slice of sausage. Afterwards, participants engaged in a filler task.

Next, the refuser manipulation was administered. Participants were instructed to carefully read the (bogus) answers of another participant on the food tasting task and to evaluate the character of this alleged participant based on his/her answers. They first read the answers about the cucumber, which were presented in exactly the same way as when the participants answered these questions themselves. After this, participants saw the response of this alleged other person on the sausage tasting questions.

In the moral refuser condition, the response of the other person indicated that s/he refused to taste the sausage because s/he thought that eating meat was unethical, and therefore based his/her refusal on moral grounds. In the non-moral refuser condition, this person refused to taste the sausage because of personal taste; s/he did not like the taste of meat and therefore based his/her refusal on non-moral grounds. In both conditions, the phrasing of the reactions was exactly the same, with the exception of a few target words. For example, in the moral refuser condition participants read *"I have not eaten meat for years because it is unethical"* while in the non-moral condition participants read *"I haven't eaten meat for years because I don't like the taste of it"*.

After completing a series of self-report measures of the evaluation of the refuser (see below) participants were instructed to deliver a short speech about their eating habits, which constituted the motivated performance situation we focused on to record the cardiovascular indices of challenge and threat motivational states. Participants were told that they had to give a one-minute speech which would allegedly be sent to the other participant. Participants were instructed to talk about their eating habits, for instance what they eat, when they eat, and whether they eat alone or with others. They could first practice the speech for a maximum of three minutes and saw some questions that they could use to

1

2

3

4

5

6

&

1

2

3

4

5

6

&

prepare their speech. These example questions were: “*What kind of dishes are your favorite?*”, “*How often do you eat those dishes?*”, “*How would you describe your eating habits?*”, “*How often do you go out for dinner?*”, “*What kind of things do you focus on while grocery shopping?*”, “*Do you usually eat alone or with others?*”, “*How healthy or unhealthy are your eating habits?*”, and “*How regular do you eat?*”. After preparing the speech, participants were instructed to look into the webcam and to deliver the speech. Physiological measures were recorded during the actual delivery of the speech. Finally, the participants were thanked, thoroughly debriefed, and paid for their participation.

Likability of refuser. Based on Monin et al. (2008), *likability of refuser* was measured using 12 items asking to what extent they thought the other participant was smart, kind, warm, pleasant, generous, reasonable, stupid (reverse coded), cold (reverse coded), dishonest (reverse coded), stingy (reverse coded), obnoxious (reverse coded) and to what extent they thought it would be pleasant to cooperate with the refuser. Participants’ responses were given on 7-point Likert scales with 1 (*not at all*) to 7 (*completely*) as endpoints. The likability items were assessed among filler items and averaged to form a reliable scale (Cronbach’s $\alpha = .89$).

Cardiovascular measures. As outlined in the introduction, the BPSM describes specific cardiovascular markers of the motivational states of threat and challenge during so-called motivated performance situations. During motivated performance, a challenge motivational state is reflected in high cardiac output (CO; the amount of blood being pumped-out by the heart in a minute), coupled with low total peripheral resistance (TPR; a measure of vascular resistance) which enables the efficient mobilization and transportation of energy during motivated performance. Threat, by contrast, is indicated by relatively high vascular resistance (TPR) and low CO which leads to a less efficient mobilization and transportation of energy during motivated performance (e.g., Blascovich & Tomaka, 1996). In addition to CO and TPR, we also measured pre-ejection period (PEP, a measure of ventricular contractility), and heart rate (HR). Although these latter measures do not distinguish between challenge and threat, a shorter PEP and higher HR in comparison to baseline levels indicate task engagement, which is an essential aspect of motivated performance (Seery, Blascovich, Weisbuch, & Vick, 2004).

Impedance cardiographic signals, electrocardiographic signals, and blood pressure were continuously measured throughout the experiment using a Biopac MP150 system, using the same apparatus and following the procedures described by De Wit and colleagues (2012). In line with standard practice (e.g., Blascovich, Seery, Mugridge, Norris, & Weisbuch, 2004; Seery et al., 2004), mean scores for HR, PEP, CO, and TPR for the last minute of the baseline and the first

minute of the speech task were calculated and then reactivity scores were created by subtracting baseline responses from task responses. Following Seery et al. (2004), reactivity scores exceeding 3 *SDs* from the mean were identified as extreme and were recoded by assigning them a value 1% higher than the next-highest non-extreme value in the same condition, thereby decreasing the influence of the extreme value while maintaining the rank order of the distribution. For each dependent variable, between zero and five values were recoded.

Results²

Likability of the refuser. An analysis of variance on the 1-7 scale that measured likability of the refuser showed a significant effect of the refuser manipulation, $F(1, 47) = 4.22, p = .046, \eta^2 = .08$. Participants liked a moral refuser less ($M = 4.38, SD = .82$) than a non-moral refuser ($M = 4.88, SD = .84$).

Cardiovascular indices of threat and challenge. In both conditions, PEP decreased significantly and HR increased significantly from baseline levels, $t_s > |6.04|, p_s < .001$, indicating task engagement and paving the way for a further interpretation of cardiovascular responses in terms of challenge and threat. A threat–challenge index was created by converting CO and TPR reactivity scores into z-scores. After multiplying the z-score for TPR with -1, it was added to the z-score of CO. Higher values on the resulting index indicate a stronger challenge motivational state whereas lower values indicate a stronger threat motivational

² The interested reader may wonder whether there were any differences in the content of the speeches between the non-moral and moral refuser conditions. For example, one could imagine that people who were confronted with a moral refuser would try to defend their meat-eating, while people who were confronted with a non-moral refuser would not. We rated to what extent the participant seemed nervous, confident, and proud to be a meat-eater, to what extent the participant excused him/herself for eating meat, mentioned good and bad things about his/her eating habits, highlighted that he/she liked to eat meat and ate meat for health reasons, and explicitly mentioned that he/she eats meat. Furthermore, we also counted how often participants mentioned meat and/or meat products, fish, vegetarian products, healthy products, unhealthy products and the number of words that participants used. Analyses of the speech-content only revealed a significant effect of the number of times meat (eating) and/or meat products were mentioned. People in the moral refuser condition mentioned meat more often ($M = 3.21, SD = 2.21$) than people in the non-moral refuser condition ($M = 1.71, SD = 1.55$), $F(1, 46) = 7.441, p = .009, \eta^2 = .14$. There were, however, no other differences in content of the speeches between the different conditions. We attribute this high similarity in speech-content in the two conditions to the fact that we provided participants with several specific guiding questions which they could discuss during their speech (e.g., how healthy their diet is, what is important to them while grocery shopping, whether they eat alone or with others, etc.; see method section) and virtually no-one veered 'off-script' by discussing things that were not related to these guiding questions.

1

2

3

4

5

6

&

state (Blascovich et al., 2004). Following the rationale of Seery et al. (2004), we controlled for task engagement (HR and PEP reactivity) in the analysis of challenge and threat. A regression analysis was performed where the covariates (reactivity scores of PEP and HR) were added in the first step, and the refuser manipulation was added in the second step. The experimental conditions were effect-coded; the non-moral refuser condition was coded -1, and the moral-refuser condition was coded 1. Results showed a significant model $F(3, 44) = 8.48, p < .001, R^2 = .37$. Supporting our hypothesis, when participants were confronted with a moral refuser, they were relatively more threatened than those who were confronted with a non-moral refuser, $t(44) = -2.50, p = .019$. Table 2.1 presents the regression coefficients.

Table 2.1.
Summary of Regression Analysis Predicting TCI (Experiment 2.1)

Variable	<i>b</i>	<i>SE b</i>	β	<i>p</i> -value
Step 1				
Reactivity HR	.06	.02	.37	.019
Reactivity PEP	-.02	.01	-.23	.137
Step 2				
Reactivity HR	.06	.02	.38	.011
Reactivity PEP	-.02	.01	-.23	.113
Refuser Manipulation	-.44	.18	-.30	.016

Note. HR = heart rate; PEP = pre-ejection period, a measure of ventricular contractility. The refuser conditions were effect-coded with the non-moral refuser condition coded as -1 and the moral refuser condition coded as 1.

In Table 2.2, the correlations between all variables (refuser manipulation, refuser likeability, HR and PEP reactivity and TCI scores) are reported. The correlations between the physiological measures (HR and PEP reactivity and TCI scores), and between the refuser manipulation and likability ratings of the refuser were significant, and the correlation between the refuser manipulation and TCI was marginally significant.

Discussion

The results of Experiment 2.1 showed that participants responded more negatively to a moral refuser than to a non-moral refuser. This negative reaction was expressed in negative evaluations of the moral refuser (by disliking him/her), and was also visible in terms of physiological signs of a motivational state of threat. That is, participants who were confronted with a moral refuser demonstrated more cardiovascular signs of threat

Table 2.2.

Correlations between Refuser Manipulation, Likability of the Refuser, TCI, HR and PEP reactivity scores (Experiment 2.1)

	Refuser manipulation	Likability of Refuser	TCI	Reactivity HR	Reactivity PEP
Refuser Manipulation	1	-.290*	-.277 ⁺	-.043	-.029
Likability of Refuser		1	.125	.080	.100
TCI			1	.490***	-.425**
Reactivity HR				1	-.538***
Reactivity PEP					1

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

than participants who were confronted with a non-moral refuser. It would be interesting to investigate whether a confrontation with a moral refuser also impacts explicit areas of self-evaluation, which is what we focused on in our second experiment.

An issue that should be discussed concerns the absence of a significant relation between threat responses and the evaluation of the moral refuser. Theoretically, a significant correlation between TCI and negative evaluations of the refuser would be plausible. However, we think the absence of this correlation in our experiment could be due to either a lack of power in our experiment or the fact that we measured the physiological indices after the self-reported evaluations of the refuser, which could have reduced the correlation. Some participants could have devalued the moral refuser in the self-reported measures, and as a consequence, shown less physiological signs of threat. This would have reduced the correlation between the self-reported dislike of the refuser and the physiological TCI.

Experiment 2.2

In our second experiment, we used a similar set-up as in the first experiment, with some adaptations in order to investigate whether physical cleansing could be used to prevent the negative consequences of the confrontation with a moral refuser. Because we reasoned that the negative consequences of being confronted with a moral refuser are especially visible in areas that concern one's self-concept, we measured self-evaluation. Furthermore, we measured moral identity to examine whether, as our reasoning implies, the effects of the moral refuser on affect and evaluations are more pronounced for people with a strong moral identity. In addition, we used refuser agency ratings (e.g., how strong, intelligent, etc. the refuser seems) to see whether the dislike for moral refusers would extend to all

1

2

3

4

5

6

&

1

evaluations of the refuser. People could, for example, acknowledge the character strength of the moral refuser, while at the same time disliking him/her.

2

Method

3

Participants and design. One-hundred-and-eight participants ($M_{\text{age}} = 21.9$; 74 women) received 3 Euros or course credit for their participation. Participants were randomly assigned to one of the four conditions in a 2 (Refuser: Moral vs. Non-Moral) X 2 (Cleansing: Absent vs. Present) design. Moral identity was assessed for all participants and served as a continuous independent variable in the analyses.³

4

5

6

&

Procedure. At the start of the experiment, participants were seated in separate cubicles with a computer and completed the 5-item internalization subscale of the Aquino and Reed (2002) moral identity scale to measure the importance of morality to the self-concept.⁴ This measure consists of the description of a person who possesses certain personality traits (e.g., caring and honest) and five questions about how it would make

³ Please note that thirty-five additional participants completed this research but were excluded from the analyses; 33 participants were excluded due to a priori criteria, and two participants were identified as multivariate outliers in the main analyses. As in Experiment 2.1, we had two important a priori criteria to exclude participants from this research: not eating the sausage (e.g., being a vegetarian) and not realizing that the refuser had refused to taste the sausage. We therefore administered manipulation checks after the self-reported dependent variables and, based on these a priori criteria, excluded 24 participants who had not tasted the sausage, and 9 participants who thought the refuser had tasted the sausage from the analyses. Furthermore, as in Experiment 2.1, we calculated both Cook's and Mahalanobis distances in the main analyses to spot potential multivariate outliers. Again, we used Cook's distance $> .10$ as the cut-off value. Similar to Experiment 2.1, we followed the guidelines by Tabachnick and Fidell (2007), and used a very conservative cut-off for the Mahalanobis value. As in Experiment 2.1, we decided to exclude potential outliers whenever they exceeded either the Cook's cut-off value of $.10$, and/or the Mahalanobis distance threshold (cut-off value = 24.32; $p < .001$, based on 7 df). In contrast to Experiment 2.1, all Cook's values were well below the cut-off of $.10$ (all Cook's distances $< .08$) but the Mahalanobis distances of two participants strongly exceeded the threshold (values of 46.61 and 38.60, respectively), and were therefore excluded from the analyses.

Please note that although we used the same criteria for excluding outliers in Experiment 2.1 and Experiment 2.2, there were no participants exceeding the threshold for Mahalanobis distances in Experiment 2.1, and no participants exceeding the threshold for Cook's distances in Experiment 2.2. This led us to exclude three participants based on extreme Cook's distances in Experiment 2.1, and two participants based on extreme Mahalanobis distances in Experiment 2.2.

⁴ Following Aquino and colleagues (2009; Aquino, McFerran, & Laven, 2011) we use only the internalization subscale of the moral identity scale, because this subscale appears to be the most robust predictor of morally relevant behavior (Aquino & Reed, 2002; Reed & Aquino, 2003; Aquino et al., 2007) and is most consistent with our definition of moral identity.

them feel to be a person that would possess such traits (e.g., “*It would make me feel good to be a person who has these characteristics*”). These items were completed using 7-point Likert-type scales, ranging from 1 (*not at all*) to 7 (*completely*), participants' answers to these items were averaged to yield a scale with sufficient reliability for the current theory-testing purposes (Cronbach's $\alpha = .63$; see Nunnally & Bernstein, 1994). After completing the items, participants engaged in the same task as in Experiment 2.1, where they were asked to taste a slice of cucumber and sausage and describe their experiences while tasting these foods.

Subsequently, in the cleansing present condition, the experimenter entered the cubicle and let the participants cleanse their hands by administering antibacterial hand soap. After this, the experimenter started the next part of the experiment and left.⁵ In the cleansing absent condition, the experimenter entered the cubicle, started the next part of the experiment and left.

After this, similar to Experiment 2.1, participants were confronted with the reaction of either someone who refused to participate in the sausage tasting task because he/she thought eating meat was immoral (moral refuser condition) or because he/she did not like meat (non-moral refuser condition). Afterwards, the dependent variables were assessed. Finally, participants were thanked, thoroughly debriefed, and paid for their participation.

Dependent variables. Responses were given using 7-point Likert scales with 1 (*not at all*) to 7 (*completely*) as endpoints. Items to measure evaluations of the refuser and the self were based on Monin et al. (2008), were assessed among other items and averaged to form reliable scales. As a measure of *dislike of the refuser*, participants responded to five items asking to what extent they thought the other participant seemed pleasant (reverse-coded), unreasonable, an obnoxious person, annoying, and to what extent they despised the refuser (Cronbach's $\alpha = .85$).

Participants also rated the extent to which they perceived the refuser as being *agentic* by indicating how intelligent, strong, confident, honest, dependent (reverse coded), independent, the other participant seemed, and to what extent they thought this refuser had high self esteem (Cronbach's $\alpha = .81$).

Self-evaluation was measured by asking participants to answer fourteen items asking them to what extent at that moment they felt happy with themselves, satisfied with themselves,

⁵ Due to the restrictions of the hand-washing manipulation, we could not include cardiovascular measures in Experiment 2.

1

2

3

4

5

6

&

good, happy, comfortable, confident, determined, disappointed with themselves (reverse coded), annoyed with themselves (reverse coded), disgusted with themselves (reverse coded), angry with themselves (reverse coded), dissatisfied with themselves (reverse coded), self-critical (reverse coded), and guilty (reverse coded), (Cronbach's $\alpha = .94$).

Results

The data were analyzed in two steps. In the first step we tested our hypothesis about the influence of refuser type (moral vs. non-moral) and cleansing (absent vs. present) on self- and refuser evaluations. In the second step we examined whether these patterns were moderated by moral identity. To this end, in the first step, regression analyses with the combined influence of the refuser and cleansing manipulations on evaluations of the refuser and the self were conducted (West, Aiken, & Krull, 1996). The refuser and cleansing conditions were effect coded (the non-moral refuser and non-cleansing conditions were coded -1, the moral refuser and cleansing conditions were coded 1) and entered in the first block. The two-way interaction effect (Refuser x Cleansing) was entered in the second block. The results of all these regression analyses are displayed in Table 2.4, under Step 1.

Dislike of the refuser. Results showed a significant model, $F(3, 104) = 3.94, p = .01, R^2 = .10$. A significant effect of the refuser manipulation was observed, $b = .28, t(104) = 2.71, p = .008$, which indicated that participants disliked a moral refuser more than a non-moral refuser. A significant Refuser x Cleansing interaction was also observed, $b = -.22, t(104) = -2.09, p = .04, \Delta R^2 = .04$. To aid in interpretation, the simple slopes were inspected (Dawson & Richter, 2006; West et al., 1996). As can be seen in Figure 2.1, when participants had not cleansed their hands, they disliked the moral refuser more than the non-moral refuser, $b = .49, t(104) = 3.19, p = .002$, similar to Experiment 2.1. When participants had cleansed their hands, however, they no longer disliked the moral refuser more than the non-moral refuser, $b = 0.06, t(104) = 0.45, p = .65$.

Perceived agency of the refuser. Results showed a significant model, $F(3, 104) = 3.28, p = .02, R^2 = .06$. Only a significant Refuser x Cleansing interaction was observed, $b = -.22, t(104) = -2.75, p = .007, \Delta R^2 = .07$. Figure 2.2 shows this interaction effect. Inspecting the simple slopes showed that when participants had not cleansed their hands, they thought the moral refuser was more agentic than the non-moral refuser, $b = 0.33, t(104) = 2.81, p = .006$. When participants had cleansed their hands, however, they no longer considered the moral refuser to be more agentic than the non-moral refuser, $b = -0.10, t(104) = -1.02, p = .31$.

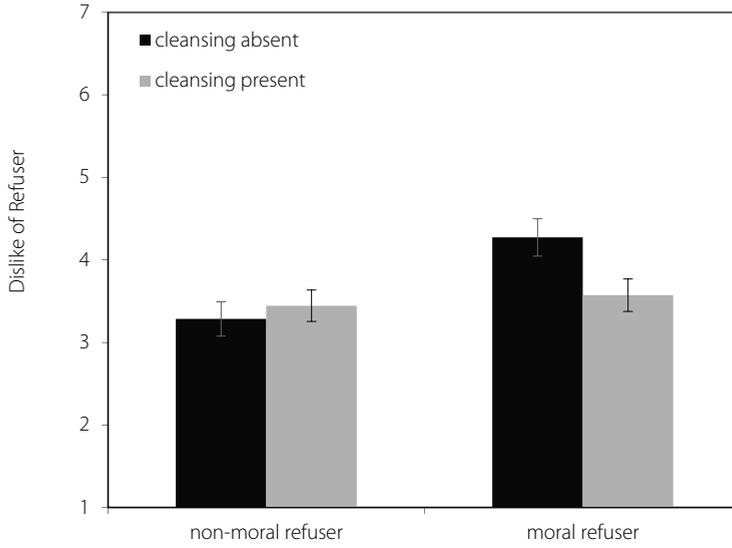


Figure 2.1. Dislike of refuser following the presence or absence of a cleansing opportunity (Experiment 2.2). Error bars are plotted one standard error above and below the mean

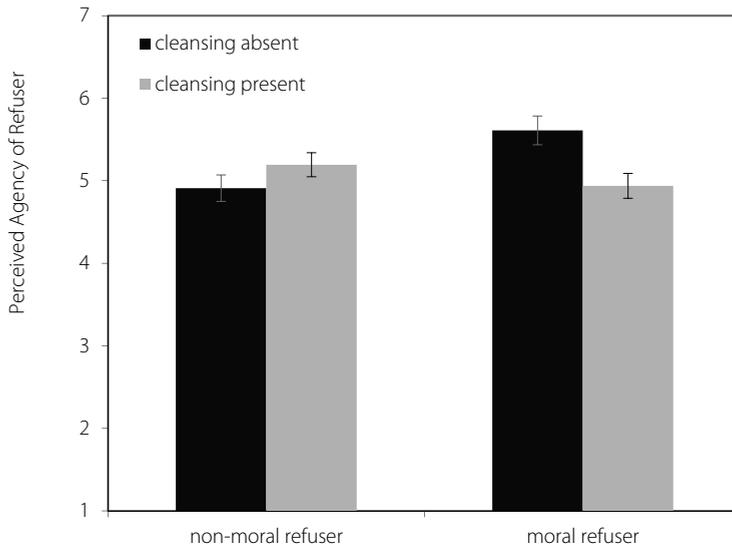


Figure 2.2. Perceived agency of refuser following the presence or absence of a cleansing opportunity (Experiment 2.2). Error bars are plotted one standard error above and below the mean

- 1
- 2
- 3
- 4
- 5
- 6
- &

1

2

3

4

5

6

&

Self-evaluation. Results showed a significant model, $F(3, 104) = 2.99, p = .03, R^2 = .08$. A significant effect of the cleansing manipulation was observed, $b = -.15, t(104) = 2.00, p = .048$, which indicated that participants evaluated themselves more positively when they had cleansed their hands than when they had not cleansed their hands. A significant Refuser x Cleansing interaction was also found, $b = .16, t(104) = 2.11, p = .037, \Delta R^2 = .04$. Simple slope analyses demonstrated that when participants had not cleansed their hands, they evaluated themselves more negatively when they were confronted with a moral refuser than when they were confronted with a non-moral refuser, $b = -0.26, t(104) = -2.16, p = .03$. When participants had cleansed their hands, however, they no longer evaluated themselves more negatively when they were confronted with a moral refuser rather than with a non-moral refuser, $b = 0.06, t(104) = 0.62, p = .54$.

Moderation by moral identity. In a second step in the regression analyses we examined whether the effects of the moral versus non-moral refuser manipulation and the physical cleansing manipulation would be more pronounced for people with a strong moral identity. To this end, we performed a new set of regression analyses where we added moral identity as a continuous predictor. Following West et al. (1996), the refuser and cleansing conditions were effect-coded (the non-moral refuser and cleansing absent conditions were coded -1, the moral refuser and cleansing present conditions were coded 1). The moral identity scale was mean-centered. The three main effects were included in Block 1. All two-way interaction effects (refuser x cleansing, refuser x moral identity, and cleansing x moral identity) were entered in Block 2. The three-way interaction (refuser x cleansing x moral identity) was entered in Block 3. The results of all these regression analyses are displayed in Table 2.4, under Step 2.

Table 2.4.

Summary of Regression Analyses Predicting Dislike of Refuser, Agency of Refuser, and Self-Evaluation (Experiment 2.2)

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
<i>Dislike of Refuser</i>				
Step 1				
Refuser Manipulation	.28	.10	.25	.01
Cleansing Manipulation	-.14	.10	-.12	.19
Refuser x Cleansing Interaction	-.22	.10	-.20	.04

Table 2.4.
 Summary of Regression Analyses Predicting Dislike of Refuser, Agency of Refuser, and Self-Evaluation
 (Experiment 2.2) (continued)

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
Step 2				
Refuser Manipulation	.29	.11	.26	.01
Cleansing Manipulation	-.14	.11	-.13	.18
Moral Identity	-.30	.27	-.11	.26
Refuser x Cleansing Interaction	-.20	.11	-.18	.06
Refuser x Moral Identity Interaction	-.02	.27	-.01	.95
Cleansing x Moral Identity Interaction	-.12	.27	-.04	.66
Refuser x Cleansing x Moral identity Interaction	-.02	.27	-.01	.95
<i>Agency of Refuser</i>				
Step 1				
Refuser Manipulation	.12	.08	.14	.15
Cleansing Manipulation	-.09	.08	-.11	.24
Refuser x Cleansing Interaction	-.22	.08	-.26	.01
Step 2				
Refuser Manipulation	.14	.08	.16	.09
Cleansing Manipulation	-.10	.08	-.11	.22
Moral Identity	.45	.20	.21	.03
Refuser x Cleansing Interaction	-.23	.08	-.28	.00
Refuser x Moral Identity Interaction	.30	.20	.15	.13
Cleansing x Moral Identity Interaction	-.29	.20	-.14	.14
Refuser x Cleansing x Moral identity Interaction	.17	.20	.08	.40
<i>Self-Evaluation</i>				
Step 1				
Refuser Manipulation	-.10	.08	-.12	.20
Cleansing Manipulation	.15	.08	.19	.05
Refuser x Cleansing Interaction	.16	.08	.20	.04
Step 2				
Refuser Manipulation	-.08	.07	-.11	.26
Cleansing Manipulation	.16	.07	.20	.03
Moral Identity	-.25	.19	-.13	.19
Refuser x Cleansing Interaction	.16	.07	.21	.03
Refuser x Moral Identity Interaction	-.31	.19	-.15	.11
Cleansing x Moral Identity Interaction	.05	.19	.03	.78
Refuser x Cleansing x Moral identity Interaction	.50	.19	.25	.01

Note. The refuser and cleansing conditions were effect coded with the non-moral and non-cleansing conditions coded -1, and the moral and cleansing conditions coded 1.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

Refuser evaluations. Apart from the significant effects on the refuser evaluations already described above, only a significant main effect of moral identity was observed on the agency refuser ratings, $b = .45$, $t(100) = 2.26$, $p = .03$. This effect indicated that the stronger people's moral identity was, the more agentic they considered the refuser to be. This could be the case because people for whom morality is important might appreciate it more when people follow their own internal guidelines, even when these internal guidelines are not necessarily related to morality. No effects of moral identity were found on the dislike refuser ratings.

Self-Evaluation. Results showed a significant model, $F(7, 100) = 2.83$, $p = .01$, $R^2 = .17$. Besides the lower-order effects already described above, we also observed a significant refuser \times cleansing \times moral identity interaction, $b = .50$, $t(100) = 2.64$, $p = .01$, $\Delta R^2 = .06$. Figures 2.3a and 2.3b illustrate this effect. Inspecting the simple slopes indicated that for participants with a weak moral identity (see Figure 2.3a) their negative self-directed affect was not significantly different when they were confronted with a moral refuser versus a non-moral refuser, regardless of whether they had cleansed their hands ($b = 0.003$, $t[100] = 0.02$, $p = .98$), or not ($b = 0.08$, $t[100] = 0.44$, $p = .66$). However, for participants with a strong moral identity (see Figure 2.3b), cleansing their hands proved beneficial. Participants with a strong moral identity who had not cleansed their hands had higher negative self-directed affect

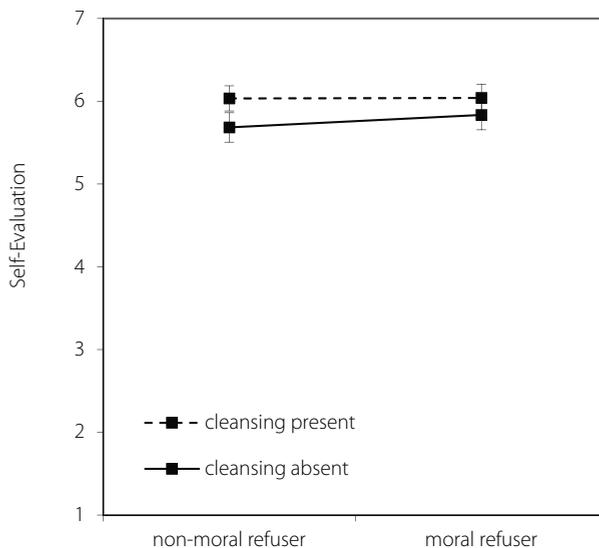


Figure 2.3a. Self-evaluation for participants with a weak moral identity, as a function of the refuser and cleansing manipulations (Experiment 2.2). Slopes are plotted one SD below the mean. Error bars are plotted one standard error above and below the mean. Higher numbers indicate more positive self-evaluations.

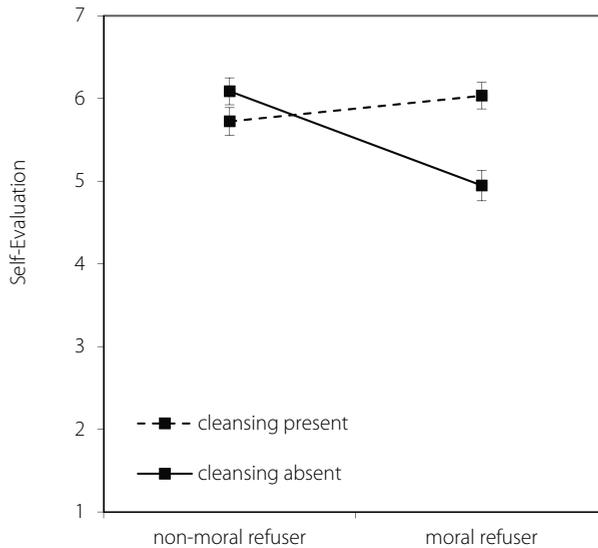


Figure 2.3b. Self-evaluation for participants with a strong moral identity, as a function of the refuser and cleansing manipulations (Experiment 2.2). Slopes are plotted one SD above the mean. Error bars are plotted one standard error above and below the mean. Higher numbers indicate more positive self-evaluation.

when they were confronted with a moral refuser than with a non-moral refuser ($b = -.57$, $t[100] = -3.57$, $p = .001$). This resembled the interaction reported earlier. For participants with a strong moral identity who had cleansed their hands however, their negative self-directed affect was no longer heightened when they were confronted with a moral refuser ($b = .16$, $t[100] = 1.08$, $p = .29$).

Discussion

Our results show that the negative implications of a confrontation with a moral refuser for self evaluations and refuser evaluations are reduced by physical cleansing. More specifically, participants who had cleansed their hands before being confronted with a moral refuser no longer showed the negative impact of this confrontation on their self- and refuser-evaluations. Participants who had not cleansed their hands, however, disliked a moral refuser more than a non-moral refuser, but also considered this moral refuser to be more agentic. Furthermore, participants experienced more negative self-evaluations when they were confronted with a moral refuser than with a non-moral refuser. Moreover, the positive influence of physical cleansing on one's self-evaluations was especially pronounced for people for whom morality was an important aspect of the self-concept.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

It is particularly noteworthy that we only found effects of moral identity on self-evaluations and not on the scales referring to the evaluation of the refuser. The fact that physical cleansing has a strong influence on self-evaluations of people with a strong moral identity is perfectly in line with the notion that a moral refuser poses a threat to the image of oneself as a moral being. The more important morality is to you, the more you experience a negative backlash regarding your self-image when you are confronted with someone who (implicitly) questions your morality. The asymmetry between the effect of moral identity on self- and other evaluations is thus in line with the notion that the strength of one's moral identity (Aquino & Reed, 2002) is about the importance of being moral to oneself and is not related to evaluations of other people.

This could further imply that, although a confrontation with a moral refuser impacts both one's evaluations of this refuser and of oneself, the underlying process of these altered self- and other-evaluations might be at least partly different. It has been found that a threatened moral self-concept underlies the derogation of these moral refusers (e.g., Monin et al., 2008), and we have found some tentative evidence to suggest that this is indeed the case. More specifically, self-evaluations and dislike of the refuser were correlated (see Table 2.3), which seems to indicate that disliking the refuser is indeed related to more negative self-evaluations, as one would expect. Furthermore, we found some very tentative evidence that suggests that dislike of the refuser could be (partially) mediated by self-evaluations: More specifically, the refuser x cleansing interaction, that was a significant predictor of refuser dislike ($p = .037$) when self-evaluation was not included as a predictor, was marginally significant ($p = .075$) only when self-evaluation was included as a predictor of refuser dislike. However, the indirect effect was not significant (90% CI: LL -.0946, UL .0014), so these findings should be interpreted with extreme caution.

Table 2.3.

Correlations Between Dislike of Refuser, Agency of Refuser, and Self-Evaluation (Experiment 2.2)

	Dislike Refuser	Agency Refuser	Self-Evaluation
Dislike Refuser	1	-.214*	-.192*
Agency Refuser		1	-.006
Self-Evaluation			1

Note. * $p < .05$

As can also be seen in Table 2.3, there were no significant relations between agency ratings of the refuser and self-evaluations, which indicated that the process via which moral

refuser are appreciated seems to be independent from the influence these moral refusers have on the self-concept. This also speaks to the fact that participants had a somewhat nuanced evaluation of the moral refuser: Although they dislike these moral refusers more than non-moral refusers, they do also seem to acknowledge the somewhat courageous nature that moral refusers have, by viewing them as more agentic than non-moral refusers. These findings also imply that a confrontation with a moral refuser does not lead to a more negative mindset in general (i.e., about the self, about the other, etc.). It could be a fruitful area for future research to further investigate when and why people's evaluations of the moral refusers differ from their self-evaluations. This is interesting because moral conflicts are very hard to solve and escalate easily, so investigating when and why people do not derogate, but rather appreciate, others who take a moral stance could be potentially helpful to solve these types of conflict. Furthermore, given that agency ratings of the refuser and self-evaluations are not related, it might be possible to prevent the self-concept against moral threats while remaining (or becoming) appreciative of the moral character of someone else.

One could wonder whether the effects of physical cleansing are due to an interruption of the experiment and to the distraction this provides, rather than to the metaphorical relationship between bodily- and moral purity. Our experiment provides two reasons to refute the former possibility. First, in both the cleansing absent and the cleansing present condition there was a short interruption when the experiment leader entered the room, said something to the participant, and started the next part of the experiment. This makes it unlikely that the observed effect was due to the distraction alone. Furthermore, if the effect of cleansing was only due to distraction, then the effect should be similar in all cleansing conditions. However, we found that the beneficial effects of physical cleansing were especially pronounced for participants with a strong moral identity, and not significant for participants with a weak moral identity. Distraction alone cannot explain why participants with a strong moral identity would benefit from this distraction when cleansing one's hands, while participants with a weak moral identity would not. Together, these results thus reveal physical cleansing as a simple, elegant and effective tool to prevent negative consequences that could arise from having one's morality questioned.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

General Discussion

Our first experiment indicated that people are threatened when they are confronted with a moral refuser (vs. a non-moral refuser), which was visible at the cardiovascular level. Furthermore, being confronted with a moral refuser led to negative evaluations of the refuser: Participants disliked a moral refuser more than a non-moral refuser. Our second experiment showed that being confronted with a moral refuser also has a negative impact on people's self-concept: Participants experienced more negative self-directed affect and less global positive affect after being confronted with a moral refuser (vs. a non-moral refuser). Although participants disliked a moral refuser more than a non-moral refuser, they also viewed moral refusers as more agentic than non-moral refusers, which suggests that participants have a differentiated view of moral refusers. Furthermore, we demonstrated that physical cleansing can be used as an effective tool to prevent the (negative) consequences for self- and other-judgments when being confronted with a moral refuser. When participants had cleansed their hands, the evaluations of themselves and the refuser were no longer significantly different when they were confronted with a moral refuser than when they were confronted with a non-moral refuser. Finally, we also showed that the beneficial effects of physical cleansing on self-related affect were especially pronounced for participants to whom morality was an important aspect of the self-concept (i.e., people with a strong moral identity).

Physical Cleansing

One implication of our work is that physical cleansing can be used to prevent the negative consequences of a future moral threat. In the present experiment, physical cleansing occurred before the introduction of the moral threat. Most previous work has primarily focused on how to restore one's moral self-concept *after* a moral threat (e.g., Zhong & Liljenquist, 2006). Our current work extends the literature on physical cleansing by demonstrating that physical cleansing can also be used to prevent against the consequences of future moral threats (see also Cramwinckel, De Cremer, & Van Dijke, 2013; Zhong et al., 2010), rather than only 'washing away' previous threats (e.g., Zhong & Liljenquist, 2006; see also Lee & Schwarz, 2011).

We also chose to implement the physical cleansing manipulation before the confrontation with the refuser because it fitted more neatly with our experimental set-up. We expected that participants would be suspicious of the physical cleansing manipulation when it would be administered after being confronted with the refuser, because there would be no obvious reason to cleanse one's hands at that point in the experiment. However,

cleansing your hands directly after eating a slice of greasy sausage (and thus before being confronted with the refuser), is a very sensible thing to do. Indeed, none of the participants reported any suspicion about the experimenter's request to cleanse their hands after eating.

One might wonder whether the observed effects of the physical cleansing manipulation are driven primarily by the moral buffer it provides (e.g., Zhong & Liljenquist, 2006; Zhong et al., 2010) or by 'the clean slate' it creates, that is, by removing the residual guilt etcetera after eating meat; Lee & Schwarz (2010, 2011). Our research does not provide a definitive answer to this question, because we aimed to offer a way to prevent the negative consequences of a confrontation with a moral refuser, rather than to provide a definitive answer on the exact workings of physical cleansing. Investigating the underlying process of physical cleansing, with experimental designs specifically designed to disentangle these separate theoretical accounts, could be a fruitful area for future research.

A related issue that deserves discussion concerns the boundary conditions of physical cleansing. Because physical cleansing is a relatively new phenomenon in social psychological research, there is little research on the limitations of its effects (for exceptions, see Cramwinckel et al., 2013; Schnall, Haidt, Clore, & Jordan, 2008). Our research indicates that the effects of physical cleansing may be most pronounced in the area of morality. First, our research indicates that physical cleansing has only a significant effect on evaluations of the self and the other when people are confronted with a moral, rather than a non-moral, refuser. This indicates that physical cleansing influences people's evaluations related to a moral confrontation but not to a non-moral confrontation. Second, physical cleansing has the most pronounced effect for people with a strong, rather than a weak, moral identity. Together, this suggests that physical cleansing has most power in the moral domain: Not only can it harness people against a moral threat; it also provides the best protection for people to whom morality is very important. Our research indicates that, although the workings of physical cleansing and morality may be extended to non-moral domains (Lee & Schwarz, 2011), the strongest effects of physical cleansing seem to be within the moral domain.

It might also be argued that the negative impact of being confronted with a moral refuser could be due to a difference in the formulation of the reactions of the non-moral and moral refuser, rather than the threat it poses to one's moral self. More specifically, one might argue that the moral refuser reaction is somewhat more judgmental than the non-moral refuser condition. However, we would argue that a moral statement typically involves a certain

1

2

3

4

5

6

&

1

2

3

4

5

6

&

degree of judgment and rejection, while a non-moral statement does not necessarily do so. Moral statements are, or could be, universal and apply to everyone (e.g., Haidt et al., 2003) and this is what makes moral statements different from non-moral statements. Being confronted with someone who takes a moral stance while you did not uphold this moral standard, will therefore typically involve (imagining) being implicitly or explicitly rejected and negatively evaluated by this moral refuser. Because it is so important for people to be moral (e.g., Ellemers, Pagliaro, Barreto, & Leach, 2008; Haidt et al., 2003; Pronin, 2008), this is exactly what makes a confrontation with a moral refuser so threatening.

The Moral Self-Concept

Being confronted with a moral refuser led to a threat to one's self-concept, which was visible at the cardiovascular level (Blascovich & Tomaka, 1996; Blascovich & Mendes, 2010; Blascovich et al., 2003). We thereby provided more direct evidence for the notion that people are threatened rather than challenged by others who take a moral stance (Monin, 2007; Monin et al., 2008). Theoretically, a confrontation with a moral refuser could elicit either challenge or threat, depending on whether an individual feels that he/she has the tools to perform well in this situation, and to portray him/herself in a positive way (Blascovich & Mendes, 2010). However, our results show that giving a speech to a moral refuser brings people in a motivational state of threat, rather than challenge. We argue that this is the case because (the behavior of) the moral refuser implicitly questions the morality of the participant's previous behavior (i.e., eating the sausage), which makes participants feel less equipped to portray themselves as moral and good during the speech. Therefore, these participants may have experienced that the demands of the situation (e.g., giving a speech to a moral refuser) outweighed their personal resources (e.g., being able to come across as a moral individual), which led to physiological signs of threat (Blascovich & Mendes, 2010; Blascovich & Tomaka, 1996). Interestingly, these cardiovascular measures were measured continuously and unobtrusively, which made them less susceptible to conscious processes such as social desirability and/or self-preservation concerns. We believe that these cardiovascular indices provide more direct evidence for the notion that moral refusers are indeed threatening for people (e.g., Monin, 2007; Monin et al., 2008), because these physiological challenge and threat responses operate outside awareness and individuals therefore didn't have to be consciously aware or willing to report that they felt threatened by the moral refuser.

Furthermore, we found, in line with previous research, that being confronted with a moral refuser leads to disliking this moral refuser, and evaluating oneself negatively

(Minson & Monin, 2012; Monin et al., 2008). Although the self-evaluation implications of a confrontation with a moral refuser might seem like a negative consequence, questioning your own morality might actually be a good thing. Since people strive for moral balance (Zhong et al., 2009), and want to feel like good and moral individuals, a negative experience in this area may lead people to behave more morally in the future. So although in the short term it might hurt to be confronted with your own moral shortcomings, it can also have the longer term benefit of paving the way for an improved moral self, especially if you care about morality.

Finally, we extended the literature on the negative consequences of being confronted with moral refusers by further highlighting the role of one's moral self-concept, that is, of people's moral identity. People to whom being moral is important experience the most severe consequences for their self-concept when they are confronted with a moral refuser. Taken together, these findings suggest that the more important it is for you to be moral, the more easily you start to question your own morality when you are confronted with someone who takes a moral stance, and the more severe the consequences for your self-evaluation are.

1

2

3

4

5

6

&



3

Negative Reactions to Non-Drinkers: Self- and Other-Directed Aggression After Confrontations with Moral Refusers

Author contributions: Author Cramwinckel designed the experiments and directed data-collection, conducted statistical analyses, interpreted results, and wrote the first draft of the manuscript. Author Van den Bos aided in designing the experiments, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Author Van Dijk aided in designing the experiments, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Author Roumen aided in designing Experiment 3.1 and collected data for Experiment 3.1. Author Rebergen aided in designing Experiment 3.2 and collected data for Experiment 3.2. Author Teitink aided in designing Experiment 3.2 and collected data for Experiment 3.2.

This chapter is based on: Cramwinckel, F. M., Van den Bos, K., Van Dijk, E., Roumen, D., Rebergen, L., & Teitink, D. (Revise and Resubmit in *the European Journal of Social Psychology*). Negative reactions to explicit non-drinkers: Distributing beer to non-drinkers and inflicting self-punishment after confrontations with moral refusers

Abstract

1

2

3

4

5

6

&

We investigate whether confrontations with moral refusers have behavioral consequences. In two studies, participants drank beer and were confronted with another participant who refused to taste beer out of moral or non-moral concern (the moral refuser and non-moral refuser conditions, respectively). In Experiment 3.1 ($N = 146$), we focus on aggressive behavior aimed at the moral refuser: We demonstrate that participants allocate more beer to a moral refuser (vs. a non-moral refuser). This difference did not occur for participants who engaged in self-affirmation. In Experiment 3.2 ($N = 143$), we focus on aggressive behavior aimed at the self. Our key finding is that participants kept their hands submerged in ice-water for a longer period of time when they were confronted with a moral refuser (vs. a non-moral refuser). Combined, our findings show that confrontations with moral refusers can lead to aggressive behavior aimed at the moral refuser and the self.

Seeing someone else display moral behavior can be threatening. Especially when the other person refuses to perform a behavior you just performed and has moral reasons for this refusal (Chapter 2; Minson & Monin, 2012; Monin, 2007; Monin, Sawyer, & Marquez, 2008). Those who refuse to display certain behaviors out of moral concerns are called *moral rebels* (Monin et al., 2008) or *moral refusers* (Chapter 2). A growing body of research shows that those who take a principled stance are often negatively evaluated (Bashir, Lockwood, Chasteen, Nadolny, & Noyes, 2013; Chapter 2; Minson & Monin, 2012; Monin, 2007; Monin et al., 2008; Parks & Stone, 2010). For example, people negatively evaluate vegetarians (Minson & Monin, 2012), especially when these vegetarians indicate moral reasons to refuse to eat meat (Chapter 2). Also, people negatively evaluate others who refuse to partake in an ostensibly racist task, especially when targets themselves did engage in this task (Monin et al., 2008).

Although research shows that people have negative evaluations of moral refusers, it is not clear whether confrontations with moral refusers also have actual behavioral consequences. In the current research, we demonstrate that confrontations with moral refusers can evoke aggressive behavior, aimed at the moral refuser (Experiments 3.1 and 3.2) and at the self (Experiment 3.2). Furthermore, we obtain further support for the notion that self-threat underlies these negative reactions. In agreement with this view, we demonstrate that self-affirmation (Cohen & Sherman, 2014; Sherman & Cohen, 2002) buffers against aggressive reactions following confrontations with moral refusers (Experiment 3.1).

We investigate these issues in a paradigm focused on reactions to non-drinking peers. Drinking alcohol has a range of negative consequences for people themselves and society as a whole. Indeed, alcohol scores in the top 3 of a ranking of the most harmful substances for individuals and society (Dutch National Institute for Public Health and the Environment [RIVM], 2009). One of the key reasons why people engage in drinking behavior is the influence of their peers (Ali & Dwyer, 2010; Choo & Shek, 2013; Godbold & Pfau, 2000; Goode, Balzarini, & Smith, 2014; Kim & Kim, 2012; Knee & Neighbors, 2002; McKay & Cole, 2012; Poelen, Engels, Van Der Vorst, Scholte, & Vermulst, 2007; Simons-Morton, Haynie, Crump, Eitel, & Saylor, 2001). For example, people exert pressure on non-drinking peers to conform (Simons-Morton et al., 2001). In the current research, we investigate how people respond to a non-drinking peer who has moral or non-moral reasons to refrain from drinking beer. We argue that people feel more threatened by refusers with a moral reason (the moral refuser conditions) than by refusers with a non-moral reason (the non-moral refuser conditions).

1

2

3

4

5

6

&

1

The threat of moral refusers

2

3

4

5

6

&

It has been argued that people derogate moral refusers in response to self-threats (Monin, 2007). People have a fundamental need to protect and maintain their self-integrity, meaning that they need to be able to see and portray themselves as moral and good individuals (Aquino & Reed, 2002; Cohen & Sherman, 2014; Sherman & Cohen, 2002, 2006). Confrontations with moral refusers can threaten this need because—by their refusal—moral refusers implicitly condemn the behavior targets displayed (Minson & Monin, 2012; Monin, 2007; Monin et al., 2008). Thus, moral refusers may lead people to feel that their behavior and morality are questioned. In other words, moral refusers may threaten people's self-integrity.

Cramwinckel and colleagues (Chapter 2) showed that confrontations with moral refusers can be perceived as threatening, as defined by the Biopsychosocial Model of challenge and threat (e.g., Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003; Blascovich & Mendes, 2010; Blascovich & Tomaka, 1996). Cramwinckel et al. (Chapter 2) examined how meat-eating participants reacted to other participants who refused to eat meat. Results showed that people experience a motivational state of threat when they are confronted with someone who refuses meat out of moral reasons (a moral refuser), and experience a motivational state of challenge when they are confronted with someone who refuses to eat meat out of non-moral reasons (a non-moral refuser). Related to this, Frings and colleagues (2012) suggested that people feel threatened when they do not have the psychological resources to deal with a confrontation with another participant, and showed that people feel challenged when they do have the psychological resources to deal with this confrontation.

Building on the above mentioned literature, we argue that moral refusers can undermine people's psychological resources, and hence can be threatening (Chapter 2; Cohen & Sherman, 2014). In the current two studies, we investigate what the implications of this self-threat can be for aggressive reactions to targets who refuse to drink beer out of moral concern.

Aggressive reactions to moral refusers

People often react defensively to self-threats, for example, by derogating the source of the threat (Cohen & Sherman, 2014; Sherman & Cohen, 2002, 2006). For example, Minson and Monin (2012) investigated how people evaluated vegetarians. In the critical condition, participants first thought about how their own moral character would be evaluated by vegetarians, before they evaluated the vegetarians. People in this condition evaluated

vegetarians more negatively than people who did not first think about how they would be evaluated by the vegetarians. Apparently, thinking about how they would be evaluated by vegetarians invoked self-threat, to which people responded by derogating the vegetarians in return.

In the current paper, we take this line of reasoning one step further, and investigate whether confrontations with moral refusers also lead to actual negative behavior directed at the other person. More specifically, in our first experiment, we aim to demonstrate that people display physical aggression aimed at moral refusers of alcoholic beverages.

Aggression is defined as any behavior intended to harm another person who does not want to be harmed (DeWall, Anderson, & Bushman, 2011). In experimental social psychology, aggression has often been operationalized by measuring how willing people are to administer unwanted physical experiences to someone else (Ayduk, Gyurak, & Luerssen, 2008; Begue et al., 2009; Lieberman, Solomon, Greenberg, & McGregor, 1999; Ritter & Eslea, 2005). For example, in the hot sauce paradigm (Lieberman et al., 1999), aggression is measured by the amount of hot chili sauce participants allocate to another person who allegedly does not like spicy foods (Ayduk et al., 2008; Begue et al., 2009; Lieberman et al., 1999; Ritter & Eslea, 2005).

We developed our research paradigm on the basis of this literature. In our experimental setting, participants first tasted beer themselves, and were later confronted with the reactions of another participant who ostensibly had refused to taste beer because he or she thought drinking alcohol was wrong (the moral refuser conditions), or because he or she disliked drinking alcohol (the non-moral refuser conditions). Afterwards, we assessed how much beer participants allocated to the refusing participant.

With this setup, we employed what could be seen as a behavioral measure of aggression: pouring beer for a person who explicitly refused to drink beer. We thus operationalized the amount of beer poured as a continuous measure of aggressive behavior against a non-drinking person. We predicted that participants would allocate more beer to a moral refuser than to a non-moral refuser.

The buffering effect of self-affirmation

People do not always react defensively to threats. Self-affirmation is a well-known method to buffer the self against integrity threats and reduce defensive responses to threats (for reviews, see e.g., Cohen & Sherman, 2014; Sherman & Cohen, 2002; 2006). Self-affirmation

1

2

3

4

5

6

&

1

2

3

4

5

6

&

is an act that demonstrates one's adequacy as a good person (Sherman & Cohen, 2014). Self-affirmation provides a buffer against threats, because it assures people that they are good people. Importantly, engaging in self-affirmation does not eliminate the threat, but it places the threatening situation in a larger context of the things that are truly important for one's sense of self and self-integrity (Cohen & Sherman, 2014). Furthermore, engaging in self-affirmation can not only buffer one's self-concept against threats, but also reduce defensiveness. Self-affirmation is often induced by asking people to reflect on values that are important to them (Cohen & Sherman, 2014; Sherman & Cohen, 2002; 2006). For example, people who engaged in self-affirmation showed more helping behavior after being confronted with a moral exemplar than people who did not engage in self-affirmation (Schnall & Roper, 2012).

In our first experiment, we build on this literature and investigate whether engaging in self-affirmation reduces aggression to moral refusers. More specifically, we expected that participants who tasted beer themselves would allocate more beer to a moral refuser than to a non-moral refuser, unless they engaged in self-affirmation before this confrontation.

Experiment 3.1: Allocating beer to non-drinkers

In this experiment, participants first tasted beer and then half of the participants had a chance to affirm important values with a writing task (the *self-affirmation conditions*) while the other half engaged in a neutral writing task (the *no-affirmation conditions*). Afterward, participants were confronted with another participant who had refused to drink this beer either out of moral reasons (the *moral refuser conditions*) or out of non-moral reasons (the *non-moral refuser conditions*).

Method

Participants and design. A total of 146 students at Utrecht University ($n = 144$) and Utrecht University of Applied Sciences ($n = 2$) took part in this experiment ($M_{\text{age}} = 21.8$ years, $SD_{\text{age}} = 2.44$ years; 105 women). Participants received 4 euros or course credit for their participation and were randomly distributed to one of the four conditions of a 2 (Affirmation: self-affirmation vs. no-affirmation) \times 2 (Refuser: moral vs. non-moral) between subjects design.¹

¹ We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in our experiments. All raw data and syntax files are available from the first author upon request. All original stimulus materials can be found in the online supplementary materials.

Procedure and instruments. Participants were recruited to participate in a ‘taste perception task’ in which they could try out new drinks. Upon arrival in the lab, participants were welcomed by the experimenter and guided to private cubicles. Participants received a plastic cup with non-alcoholic beer (i.e., Amstel bier 0.0%), which weighed 165 grams. Participants were unaware of the fact that the beer contained no alcohol and did not receive any information about the brand or type of beer they received. Participants were invited to “taste as much beer as they needed”. Hereafter, participants filled out a short questionnaire, consisting of 3 open questions and 9 questions that could be answered on 7-point Likert scales, ranging from 1 (*not at all*) to 7 (*very much*). These questions were merely intended to support the cover story that participants took part in a tasting task and to convey that we were interested in their evaluations of the drinks they tasted during the experiment. Example questions were “*How much did you like this beer?*” and “*To what extent would you be inclined to drink this beer again?*”. In a different room (and thus without participants’ knowing), the experimenter weighed the plastic cup to determine how much beer participants had drunk in this first round of drinking beer.

Self-affirmation manipulation. Hereafter, the self-affirmation manipulation was implemented by means of a short writing assignment. Participants in the *self-affirmation condition* wrote down three positive aspects about themselves and three instances where these positive aspects were important to them (Liu & Steele, 1986; Steele & Liu, 1983; Wiesenfeld, Brockner, & Martin, 1999). Participants in the *no-affirmation condition* wrote about a normal day in their lives, and what they usually do on such a normal day (Van den Bos et al., 2011).

Refuser manipulation. Based on the moral and non-moral refuser manipulations of Cramwinckel and colleagues (Chapter 2), participants were confronted with the fictitious response of an earlier participant, who ostensibly refused to taste the beer. This fictitious participant had not

In order to compute the required sample sizes for our studies a priori, we performed a power analysis with the program G*Power (Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009). We used the observed effect size obtained by Cramwinckel et al. (Chapter 2) as our estimation of the effect size of the negative reaction towards the moral refusers, because our experimental procedure resembled their experimental design. Therefore, we used an η^2 of .08 (as obtained by Cramwinckel et al., Chapter 2, Experiment 2.1), a power level of .90, and the standard α -level of .05. According to this power analysis, we would need a total sample of 123 to achieve a power of .90. We expected some participants to have extreme scores on our dependent variables, because we used behavioral measures (i.e., amount of beer distributed or submersing one’s hand in ice-water for a certain period of time) that we expected to be sensitive to outliers. Therefore, in both experiments, we decided to collect as many participants as possible within the reserved lab time, with a minimum of 123 participants per experiment.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

filled out any of the questions on the questionnaire, with the exception of one hand-written sentence on the bottom of the page. In the *moral refuser condition*, this sentence read “*I don’t drink beer because I think drinking alcohol is wrong*”. In the *non-moral refuser condition* this sentence read “*I don’t drink beer because I don’t like drinking alcohol*”. After reading the response of the refuser, participants answered eight questions on 7-point Likert scales ranging from 1 (*not at all*) to 7 (*very much so*), and two open questions. These questions were intended to support the cover story and to make sure that participants paid attention to the reaction of the refuser. Example questions were “*The other participant and I agree about the taste of the drink*” and “*The other participant and I would probably have a great time hanging out*”.

Hereafter, participants received a second plastic cup of (non-alcoholic) beer, identical to the first round. Again, participants were invited to “*taste as much as they needed*”. After tasting the second beer, participants answered the same questions as in the first round. One extra question was added where participants could indicate which of the two drinks was their favorite. In another room (and thus without participants’ knowing), the experimenter weighed the plastic cup to measure how much beer the participants had drunk in this second round of drinking beers.

Dependent variables. Participants received an empty plastic cup (weight: 6 grams) and two cans of beer. These were ostensibly the two beers that participants tasted in Rounds 1 and 2. It was impossible for participants to see the original beer cans, because the cans were covered by can coolers that were taped to the cans. These can coolers were marked ‘1’ and ‘2’, ostensibly to identify which beer participants tasted in the first and second round, respectively. Participants were informed that they could pick one of these two beers and pour it for the participant of whom they just read the responses (i.e. the moral or non-moral refuser). Participants noted which beer (i.e., ‘1’ or ‘2’) they would allocate to the other participant. They were instructed to “*pour as much as they thought the other participant needed*”. Our main dependent variable was how much beer participants poured for the other participant.

For explorative reasons, we also measured participants’ self-evaluations at that moment with 23 self-evaluation items used by Cramwinckel and colleagues (Chapter 5; Cronbach’s $\alpha = .91$). Items could be answered on 7-point Likert scales, ranging from 1 (*not at all*) to 7 (*very much so*). Example items were “*To what extent do you feel happy with yourself?*” and “*To what extent do you feel disappointed in yourself?*” (reverse coded), see the supplemental materials for an overview of all questions. Finally, participants provided demographic variables such as age, sex, living situation, religion, and the amount of alcoholic beverages they typically consumed within one week. Afterwards, participants were thanked, debriefed, paid and dismissed. The whole procedure lasted about 25 minutes.

Results

Main analyses. To investigate whether participants would distribute more beer to moral refusers than non-moral refusers, possibly moderated by our self-affirmation manipulation, we performed an analysis of covariance (ANCOVA) with the refuser and affirmation manipulations as the independent variables, and the amount of beer distributed as the dependent variable. To control for error variance in this experimental paradigm with real societal implications, we included several covariates in these analyses that we expected to influence the main dependent variable. These covariates were the amount of beer participants drank in Rounds 1 and 2, their average number of alcoholic beverages participants typically consume within one week, and participants' gender.

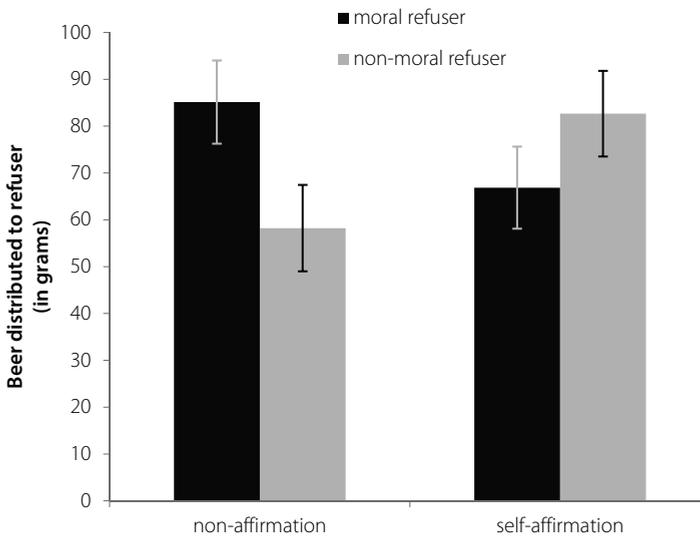


Figure 3.1. Amount of beer handed out to refuser in Experiment 3.1, as a function of Refuser and Affirmation manipulations. Amount of beer drank in rounds 1 and 2, sex, and the number of alcoholic beverages per week were added as covariates in the analyses. Errors bars are depicted one standard error above and below the mean.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

Results showed a significant Refuser x Affirmation interaction, $F(1, 133) = 5.61, p = .019, \eta_p^2 = .040$.² As can be seen in Figure 3.1, simple main effects showed that participants in the no-affirmation condition poured more beer for the moral refuser ($M = 85.14, SD = 59.58, 95\% \text{ CI } [67.56; 102.72]$) than for the non-moral refuser ($M = 58.22, SD = 57.63, 95\% \text{ CI } [40.89; 75.56]$), $F(1, 133) = 4.68, p = .032, \eta_p^2 = .034$. For participants in the self-affirmation conditions, this difference attenuated. More specifically, the amount of beer participants poured for a moral refuser ($M = 66.86, SD = 78.08, 95\% \text{ CI } [48.62; 85.10]$) did not differ significantly from the amount of beer participants poured for a non-moral refuser ($M = 80.75, SD = 60.15, 95\% \text{ CI } [64.58; 100.75]$), $F(1, 133) = 1.48, p = .225, \eta_p^2 = .011$. These findings were in line with our expectations that participants would pour more beer for a moral refuser than a non-moral refuser, unless they engaged in self-affirmation.

Furthermore, the covariates amount of beer drank in the second round ($F[1, 133] = 9.82, p = .002, \eta_p^2 = .069$), and the number of alcoholic beverages each week ($F[1, 133] = 4.90, p = .029, \eta_p^2 = .035$) were also significantly related to the amount of beer distributed to the refuser. This indicated that (as expected), the more participants drank themselves, the more beer they distributed to the refuser. The effects of beer drank in round 1 and gender were not significant (p 's $> .104$), nor were there any significant main effects of the refuser and affirmation manipulations (p 's $> .536$).³

² In order to identify potential outliers in our dataset, we calculated Cook's (1977) distance in this analysis (Cohen, Cohen, West, & Aiken, 2003). This revealed that 4 of the 146 participants of Experiment 3.1 (2.7% of the sample) indicated a Cook's distance score of more than 3 SDs above the mean. Closer inspection revealed that these participants distributed either extremely low amounts (i.e., 30-44 grams; 2 participants) or extremely high amounts (i.e., 231-285 grams; 2 participants) of beer to the refuser, compared to the overall average of 75.40 grams. Furthermore, there was 1 participant with missing data. These 5 participants were excluded from the main analyses of Experiment 3.1, leaving a total of 141 participants, with 34 to 37 participants in each cell of the design. Importantly, including the outlying participants in our analyses still yielded a significant two-way interaction effect between refuser and affirmation on the amount of beer distributed to the refuser, $F(1, 137) = 4.32, p = .040, \eta_p^2 = .03$, attesting to the robustness of the predicted effect.

³ Additionally, simple main effects showed a marginally significant effect: Participants in the non-moral refuser condition poured somewhat more beer for the refuser when they engaged in self-affirmation ($M = 82.66, SD = 60.15, 95\% \text{ CI } [64.58; 100.75]$) than when they did not engage in self-affirmation ($M = 58.22, SD = 57.63, 95\% \text{ CI } [40.89; 75.56]$), $F(1, 133) = 3.68, p = .057, \eta_p^2 = .027$. For participants in the moral refuser conditions, this difference was not significant, as the amount of beer participants poured did not differ significantly regardless of whether they engaged in self-affirmation ($M = 66.86, SD = 78.08, 95\% \text{ CI } [48.62; 85.10]$) or not ($M = 85.14, SD = 59.58, 95\% \text{ CI } [67.56; 102.72]$), $F(1, 133) = 2.00, p = .160, \eta_p^2 = .015$.

Additional analyses. Readers might also be interested in whether some participants refused to allocate beer to the earlier participant (i.e., the moral or non-moral refuser), and whether this differed as a function of our manipulations. Therefore, we recoded the continuous dependent variable (the amount of beer distributed, measured in grams) into a dichotomous variable indicating whether or not participants allocated beer to the refuser (i.e., 0 = no allocation of beer, 1 = allocation of beer).

As can be seen in Table 3.1, when participants were confronted with a moral refuser, they more often refused to allocate beer to this moral refuser when they engaged in self-affirmation (27.8% did not allocate beer) than when they did not engage in self-affirmation (10.8% did not allocate beer), although this difference was marginally significant only, $\chi^2(1) = 3.389, p = .066$, Cramer's $V = .215$. When participants were confronted with a non-moral refuser, participants did not more often refuse to allocate beer when they engaged in self-affirmation (16.2%) as opposed to no self-affirmation (13.9%), $\chi^2(1) = 0.077, p = .781$, Cramer's $V = .033$.

Table 3.1
Cross Tables of Beer Allocation to Refuser in Experiments 3.1 and 3.2

		No allocation	Allocation
Experiment 3.1			
Moral Refuser	Non-affirmation	10.8% (4)	89.2% (33)
	Affirmation	27.8% (10)	72.2% (26)
	Total	19.2% (14)	80.8% (59)
Non-Moral Refuser	Non-affirmation	16.2% (6)	83.8% (31)
	Affirmation	13.9% (5)	86.1% (31)
	Total	15.1% (11)	84.9% (62)
Total		17.1% (25)	82.9% (121)
Experiment 3.2			
Moral Refuser	Beer first	36.1% (13)	63.9% (23)
	Ice-water first	25.0% (9)	75.0% (27)
	Total	30.6% (22)	69.4% (50)
Non-Moral Refuser	Beer first	11.8% (4)	88.2% (30)
	Ice-water first	13.5% (5)	86.5% (32)
	Total	12.7% (9)	87.3% (62)
Total		21.7% (31)	78.3% (112)

1

2

3

4

5

6

&

1

2

3

4

5

6

&

To explore whether explicit self-evaluations would be influenced by the refuser and affirmation manipulations, we performed an ANCOVA with the refuser and affirmation manipulations as independent variables, and the self-evaluation scale as a dependent variable. As in the main analyses, we included the amount of beer participants drank in Rounds 1 and 2, their average number of alcoholic beverages participants typically consume within one week, and participants' gender as covariates. Results yielded no significant main- or interaction effects, all p 's $> .125$.

However, inspection of the simple main effects suggested that when participants were confronted with a moral refuser, their self-evaluations were somewhat higher when they were self-affirmed ($M = 5.57, SD = 0.59, 95\% CI [5.34; 5.80]$) than when they were not self-affirmed ($M = 5.26, SD = 0.76, 95\% CI [5.04; 5.49]$), $F(1, 132) = 3.35, p = .069, \eta_p^2 = .025$, although this effect was marginally significant only. When they were confronted with a non-moral refuser, self-evaluations did not differ as a function of whether participants were self-affirmed ($M = 5.57, SD = 0.67, 95\% CI [5.19; 5.65]$) or not ($M = 5.36, SD = 0.66, 95\% CI [5.14; 5.59]$), $F(1, 132) = 0.12, p = .728, \eta_p^2 = .001$. These patterns correspond with our line of reasoning that confrontations with moral refusers threaten people's self-evaluations, which can be reduced by self-affirmation. We hasten to note that these exploratory findings should be interpreted cautiously, because we did not observe a significant interaction effect between type of refuser and self-affirmation.

Discussion

The main focus of this first experiment was to investigate whether participants who tasted beer would react aggressively to someone else who refused to drink out of moral concerns. As expected, we observed that people distributed more beer to refusers with a moral reason ("*I don't drink beer because I think drinking alcohol is wrong*"), than to refusers with a non-moral reason ("*I don't drink beer because I don't like drinking alcohol*"). Importantly, this difference attenuated when participants had a chance to affirm some important aspects of themselves before the confrontation with a non-drinking peer.

We also observed that participants more often refused to allocate beer to moral refusers than to non-moral refusers, but only when participants engaged in self-affirmation. Combining these two insights lead us to conclude that self-affirmation not only leads people to allocate less beer to a moral non-drinker, but may also lead them to more often completely refrain from allocating beer to moral non-drinkers.

Although not the main focus of this experiment, we explored whether people's self-evaluations would be influenced by the confrontation with a moral refuser. We observed some tentative support for this notion. One may wonder why we did not observe stronger effects of the moral refuser on explicit self-evaluations. The main focus of this first experiment was to investigate aggressive behavior aimed at the refuser. Therefore, we first measured aggression (indicated by the allocation of beer to a non-drinking earlier participant). As a consequence, we measured self-evaluations second, and thus after the allocation of beer. It may be the case that allocating beer to the refuser adequately protected participants' self-concepts, at least for some of the participants (for a similar line of reasoning, see e.g., Chapter 2).

The pattern that participants who were confronted with a non-moral refuser tended to distribute somewhat more beer to this refuser after engaging in self-affirmation is in line with the workings and limitations of self-affirmation, as described by Cohen and Sherman (2014). Self-affirmation is a way to buffer against self-threats and to reduce defensive responding to those threats (e.g., Cohen & Sherman, 2014; Sherman & Cohen, 2002; 2006). As such, self-affirmation has the most beneficial consequences when self-threat is present (as is the case in a confrontation with a moral refuser). In these instances, self-affirmation has beneficial effects. These beneficial effects can be indicated by less defensive responses to threats, for example, by allocating less beer to a moral refuser when one is self-affirmed. This is also what we found in this experiment. However, in the absence of threat (as is the case in a confrontation with a non-moral refuser), self-affirmation may sometimes even backfire. In the case of the current experiment, this fits with the marginally significant finding that people distributed more beer to a non-moral refuser when they were self-affirmed than when they were not self-affirmed.

Experiment 3.2: Enduring pain

People may also have other reactions to self-threat, rather than reacting negatively towards moral refusers. For example, people can display behavior that restores their personal sense of adequacy (Cohen & Sherman, 2014). One type of behavior that is especially relevant for the current research is self-punishment, or inflicting and enduring pain to oneself (e.g., Nelissen, 2012; Nelissen & Zeelenberg, 2009). People sometimes inflict pain on themselves after transgressions (Inbar, Pizarro, Gilovich, & Ariely, 2013; Nelissen & Zeelenberg, 2009). For example, people who first wrote about one of their past ethical transgressions (i.e., a guilt-inducing event) have been found to inflict more intense electric shocks on themselves than did those who wrote about feeling sad or a neutral event (Inbar et al., 2013). Furthermore,

1

2

3

4

5

6

&

1

2

3

4

5

6

&

people also inflicted more intense electrical shocks on themselves in the presence of the victim of their transgression, rather than someone they have not encountered before (Nelissen, 2012). Inflicting pain to oneself and enduring pain can demonstrate one's moral character (Bastian, Jetten, & Stewart, 2013). In this sense, enduring pain may be especially effective as a response to self-threat, because it can affirm a person's character strength and virtue (Bastian et al., 2013; Cohen & Sherman, 2014). Inflicting pain to someone else is often seen as a form of aggression (Pedersen, Vasquez, Bartholow, Grosvenor, & Truong, 2014), and, by the same token, we argue that self-inflicted pain could therefore be used as a measure of self-directed aggression.

Therefore, in our second experiment, we investigate whether people endure more pain after a confrontation with someone who explicitly refused to drink alcohol out of moral concern (i.e, a moral refuser). We operationalized this as the amount of time participants held their hand submersed in a circulated ice-cold water bath (2° Celsius or 35.6° Fahrenheit). This *cold presser test* (CPT) induces pain (Bastian, Jetten, & Fasoli, 2011; Bastian et al., 2013; Chen, Poon, & DeWall, 2015; Jackson, Iezzi, Nagasaka, Fritch, & Gunderson, 2002). We expected that participants would submerge their hands for a longer time when they were confronted with a moral refuser (vs. a non-moral refuser). Furthermore, we also expected that participants would distribute more beer to moral refusers than to non-moral refusers.

We counterbalanced the order of our two dependent variables: how long participants held their hands in ice-cold water and the amount of beer distributed to a non-drinking peer. Thus, we randomly distributed participants to either distribute beer first, and submerge their hands in ice-cold water afterwards, or to do the reverse.

Method

Participants. A total of 143 students at Utrecht University took part in this Experiment ($M_{\text{age}} = 21.8$ years, $SD_{\text{age}} = 3.13$ years; 109 women). Participants received 4 euros or course credit for their participation and were randomly assigned to one of the four conditions of a 2 (Refuser: moral vs. non-moral) x 2 (Order of dependent variables: beer allocation first vs. ice-water first) between-subjects design. Type of dependent variable (beer allocation vs. ice-water) was included as a within-subjects variable in the analyses. We used the same criteria as in Experiment 3.1 to determine the sample size.

Procedure and instruments. The procedure of our second study was similar to Experiment 3.1. We therefore only describe the differences between the two studies. The most important difference was the implementation of a second dependent variable; a cold presser test

where participants submersed their hand into a circulating bath filled with ice-cold water (Bastian et al., 2011; Bastian et al., 2013; Rutchick & Slepian, 2013). Another difference with Experiment 3.1 was that we did not implement a self-affirmation manipulation here, nor did we measure explicit self-evaluation with self-reported statements.

Order of dependent variables. Half of the participants first allocated beer to the refuser (see the Method section in Experiment 3.1 for a complete description), and engaged in the cold presser test afterwards. The other half of the participants first engaged in the cold presser test, and allocated beer afterwards.

Cold presser test. Participants were led to a different cubicle, and were seated on a chair next to a bath filled with circulating ice-cold water of 2 °C. They received written instructions that informed them they were required to submerge their non-dominant hand into cold water (i.e., 2 °C), as part of a test measuring endurance. Participants further learned that the test would start when they submersed their hand completely into the water, and stopped when they retracted their hand completely from the water. They were instructed to keep their hand submersed for as long as possible, but to retract their hand whenever it became too sensitive or painful to submerge their hand any longer (Bastian et al., 2011; Bastian et al., 2013; Rutchick & Slepian, 2013). Finally, participants provided the same demographic variables as in Experiment 3.1. Afterwards, participants were thanked, paid and dismissed. The whole procedure took about 25 minutes. After the data collection was completed, participants were debriefed via email.

Results

We performed General Linear Model analyses (GLM) with the refuser manipulation and the order of the dependent variables as between subjects variables, type of dependent variable as a within-subjects variable, and the amount of beer distributed and the time spent in ice-water as the dependent variables. We also included the amount of beer participants drank in the first and second rounds, their average number of alcoholic beverages in one week, and participants' gender as covariates. Because the dependent variables were measured on different scales (i.e., beer allocation was measured in grams, time in ice-water was measured in seconds), we performed the analysis on the standardized variables. To ease interpretation, we report the means and SD's in the original units of measurement (i.e., beer allocation in grams and time in seconds).

Main analyses. Multivariate results on the between-subjects factors showed only a significant main effect of Refuser, $F(1, 130) = 5.916, p = .016, \eta_p^2 = .044$. As we predicted,

1

2

3

4

5

6

&

1

participants had higher scores on the dependent variables (holding their hand in ice-water for longer time and allocating more beer) when they were confronted with a moral refuser rather than a non-moral refuser.⁴

2

3

Inspecting the dependent variables separately by investigating the simple main effects demonstrated that only the effect of refuser on time in ice-water was significant, $F(1, 130) = 5.261, p = .023, \eta_p^2 = .039$. As can be seen in Figure 2a, participants who were confronted with a moral refuser kept their hand longer in ice-water ($M = 44.31, SD = 26.99$) than participants who were confronted with the non-moral refuser ($M = 34.75, SD = 18.65$), which was in line with our expectations.

4

5

6

&

The effect of refuser on beer allocated to the refuser was not significant, $F(1, 130) = 1.937, p = .166, \eta_p^2 = .015$. However, Figure 3.2b demonstrates that the means were in the expected direction with participants distributing more beer to moral refusers ($M = 65.90, SD = 63.81$) than to non-moral refusers ($M = 52.65, SD = 39.86$). No other significant between-subjects or within-subjects effects were observed, all multivariate $ps > .101$.

Additional analyses. As in Experiment 3.1, we explored whether some participants refused to allocate beer to the earlier participant (i.e., the moral or non-moral refuser), and whether this differed as a function of our manipulations. Following the same procedure as in Experiment 3.1, we recoded beer allocated to the refuser (measured in grams) into a dichotomous variable, with 0 indicating no allocation of beer and 1 indicating allocation of beer. As can be seen in Table 3.1, participants more often refused to allocate beer to a moral refuser (30.6% did not allocate beer) than to a non-moral refuser (12.7% did not allocate beer), $\chi^2(1) = 6.731, p = .009$, Cramer's $V = .217$. However, this effect was significant only when participants allocated beer first (36.1% vs. 11.8% refusal to allocate beer to a moral vs. a non-moral refuser, respectively, $\chi^2[1] = 5.637, p = .018$, Cramer's $V = .284$), and not significant when participants first submersed their hand into ice-cold water (25.0% vs. 13.5% refusal

⁴ Similar to Experiment 3.1, inspecting Cook's distance measure in this analysis revealed that 5 of the 143 participants of Experiment 3.2 (3.5% of the sample) indicated a distance score of more than 3 SDs above the mean. Closer inspection revealed that these participants distributed either nothing (i.e., 6 grams, 1 participant) or extremely high amounts of beer to the refuser (i.e., 240-273 grams; 2 participants), compared to the overall average of 60.98 grams; or they submersed their hands in ice-cold water for an exceptionally long time (i.e., 120-160 seconds; 2 participants), compared to the overall average of 41.4 seconds. These five participants were excluded from the main analyses of Experiment 3.2, leaving a total of 138 participants, with 34 to 37 participants in each cell of the design. Including the outlying participants in our analyses also yielded a significant main effect of refuser, $F(1, 135) = 4.09, p = .045, \eta_p^2 = .029$, attesting to the robustness of the predicted effect.

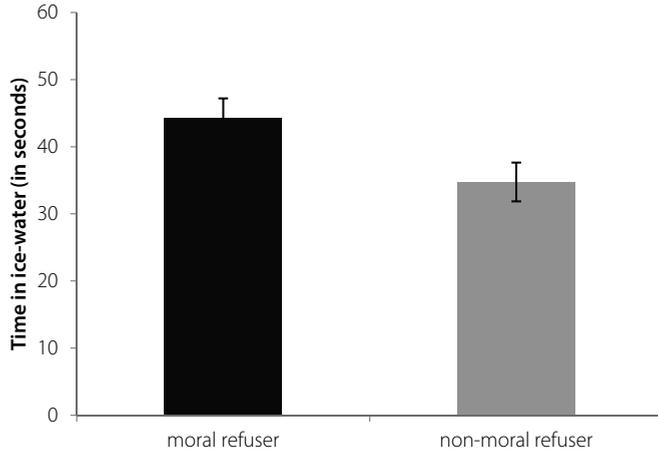


Figure 3.2a. Time spent in ice-water in Experiment 3.2, as a function of the refuser manipulation. Amount of beer drank in rounds 1 and 2, sex, and the number of alcoholic beverages per week were added as covariates in the analyses. Errors bars are depicted one standard error above and below the mean.

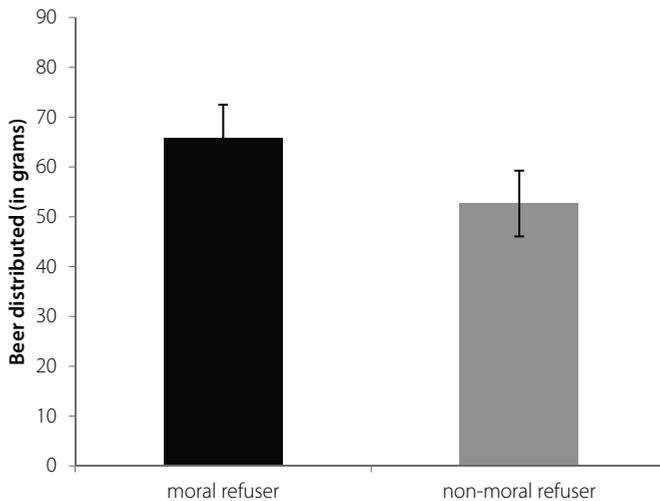


Figure 3.2b. Amount of beer distributed to refuser in Experiment 3.2, as a function of the refuser manipulation. Amount of beer drank in rounds 1 and 2, sex, and the number of alcoholic beverages per week were added as covariates in the analyses. Errors bars are depicted one standard error above and below the mean.

1

2

3

4

5

6

&

1

to allocate beer to a moral vs. a non-moral refuser, respectively, $\chi^2 [1] = 1.553, p = .213$, Cramer's $V = .146$).

2

Discussion

3

We expected and found that people display more aggression after a threatening confrontation with a moral refuser than after a less threatening confrontation with a non-moral refuser. More specifically, participants had higher combined scores on the two dependent variables (submersing their hand in ice-cold water and allocating beer to a non-drinking earlier participant) when they were confronted with a moral refuser. Closer inspection demonstrated that the effect of our refuser manipulation was most pronounced on self-directed aggression. More specifically, participants held their hands submersed in ice-cold water for a longer period of time when they were confronted with an earlier participant who refused to drink beer out of moral reasons, rather than out of non-moral reasons.

4

5

6

&

Additional analyses demonstrated that participants more often refused to allocate beer to a moral refuser than a non-moral refuser (but only when allocation of beer was the first dependent variable we measured), which resembled the observed pattern obtained in Experiment 3.1.

The main aim of Experiment 3.2 was to investigate whether moral refusers could lead to self-directed aggression, operationalized by how much pain participants inflicted on themselves by submersing their hand into ice-cold water. However, it would have been in line with our expectations to find that people also display aggression towards moral refusers, as we did in Experiment 3.1. Although we observed a significant combined effect of our refuser manipulation on time in ice-water and the amount of beer allocated, the effect on the beer allocation measure separately failed to reach significance (although the means were in the expected direction). It could be that Experiment 3.2 was underpowered to replicate the findings of the first study (Simonsohn, 2015).⁵ However, it would strengthen confidence in the robustness of the beer allocation measure if our main results from Experiment 3.1 would replicate. Therefore, we are currently in the process of running a pre-registered (close) replication of Experiment 3.1, following the guidelines by Van 't Veer and Giner-Sorolla (2016). Finally, we like to explicitly note that the main aim of our second experiment was to demonstrate that participants inflict more pain on themselves after a confrontation with a moral refuser, and we found convincing support that this is the case.

⁵ Post-hoc power analyses showed that observed power to replicate the effect on the beer allocation measure was low ($\beta_{\text{observed}} = .311$).

General Discussion

In two experiments, we investigate how people respond to self-threatening situations. More specifically, we show that a threatening confrontation with a moral refuser can lead to behavioral aggression, aimed at the moral refuser, and at the self. In Experiment 3.1, we focused on aggression aimed at the refuser. Here, we demonstrated that participants allocated more beer to an earlier participant who had moral reasons not to drink beer (i.e., a moral refuser) than to an earlier participant who had non-moral reasons not to drink beer (i.e., a non-moral refuser). Importantly, these differences attenuated when participants had the chance to protect their self-integrity with a self-affirmation task before the threatening confrontation, thereby highlighting the role of self-threat.

In Experiment 3.2, we focused on aggression aimed at the self. In this study we demonstrated that participants endured more pain (by submersing their hands in ice-cold water for a longer period of time) when they were confronted with a moral refuser rather than a non-moral refuser.

Our findings build on, and extend, the moral refuser literature (e.g., Chapter 2; Minson & Monin, 2012; Monin et al., 2008). In the existing literature, derogation of moral refusers is often demonstrated by self-reported evaluations of the targets. For example, people who ate meat disliked moral vegetarians more than non-moral vegetarians, and people who participated (vs. people who were merely observers) in an ostensibly racist task experienced less interpersonal attraction to someone who refused to participate out of moral reasons (Chapter 2; Monin et al., 2008). We extend these findings by showing that derogation of moral refusers can also be observed in actual behavior: displaying aggression by pouring beer for someone who does not want to drink alcohol.

Our findings highlight the important role of the threatened self-concept in explaining negative reactions to moral refusers. Apparently, the confrontation with a moral refuser threatens people's self-integrity: When people have the chance to affirm important values before being confronted with a refuser, they do not react more aggressively towards moral refusers than towards non-moral refusers. This corroborates with the finding that people experience a motivational state of threat when they are confronted with moral refusers (Chapter 2).

The current research has implications for research and theorizing on self-affirmation (e.g., Cohen & Sherman, 2014; Sherman & Cohen, 2002; 2006). First, we demonstrate that self-affirmation can buffer against aggressive reactions aimed at moral refusers. Self-affirmation

1

2

3

4

5

6

&

1

2

3

4

5

6

&

thus operates as a relevant strategy to prevent actual negative behavior: aggression aimed at the moral refuser. Second, although it was not an explicit aim of the current research, we demonstrated a potential boundary condition of self-affirmation. More specifically, when threat is absent (i.e., when people are confronted with non-moral refusers), engaging in self-affirmation can backfire and lead to more aggressive behavior. In order to reap the benefits of self-affirmation, it may thus be wise to only employ self-affirmation in threatening situations.

Furthermore, our findings may be relevant for the literature on peer pressure and drinking behavior (Ali & Dwyer, 2010; Choo & Shek, 2013; Godbold & Pfau, 2000; Goode et al., 2014; Kim & Kim, 2012; Knee & Neighbors, 2002; McKay & Cole, 2012; Poelen et al., 2007; Simons-Morton et al., 2001). In this literature, it is often found that peer influences determine to a large extent why and how much people drink (Ali & Dwyer, 2010). However, much less is known about why and when people apply pressure onto others to drink. This direct peer pressure can consist of offering drinks to others, and/or encourage them to drink (Simons-Morton et al., 2001). Our main dependent variable, pouring beer for someone who explicitly refused to drink, can be seen as a form of behavioral peer pressure. We demonstrated that people pour more beer for a moral refuser than for a non-moral refuser, which can be interpreted as applying more pressure to drink on moral refusers. The current research is thereby—to our knowledge—the first to provide systematic empirical support to demonstrate what characteristics of a non-drinker evoke more pressure from drinkers.

Related to this, in both studies, we observed that people were more likely to refrain from allocating beer to moral refusers than to non-moral refusers (although we observed this effect only for self-affirmed participants in Experiment 3.1). More specifically, around one third of participants did not allocate beer to moral refusers. This seems to indicate an interesting paradox: On the one hand, people are less likely to allocate beer to someone who has moral reasons not to drink. On the other hand, those who *do* allocate beer to a non-drinker, they allocate more when the refuser has moral reasons.

We can connect our findings to literature on self-punishment and aggression (Anderson & Bushman, 2002; Ayduk et al., 2008; Begue et al., 2009; DeWall et al., 2011; Inbar et al., 2013; Lieberman et al., 1999; Nelissen & Zeelenberg, 2009; Nelissen, 2012; Pedersen et al., 2014; Ritter & Eslea, 2005; Rutchick & Slepian, 2013) by demonstrating that people display self-directed aggression when they are confronted with moral refusers. Enduring self-inflicted pain may be an effective way to demonstrate one's moral character (Bastian et al., 2011; Bastian et al., 2013; Bastian, Jetten, Hornsey, & Leknes, 2014). Thus, it may be a successful

way to restore one's threatened self-concept. Our findings speak to this line of reasoning. More specifically, we demonstrate that people endure more self-inflicted pain when they are confronted with someone who refuses to drink beer out of moral reasons than when they are confronted with someone who refuse to drink beer out of non-moral reasons. Apparently, the confrontation with a moral refuser threatens participants' self-integrity to which they respond with inflicting pain on themselves.

Finally, although not the main focus of our research, our findings speak to the question of whether people can substitute one type of responding to threats with another type of responding (e.g., Cohen & Sherman, 2014). For example, if people already derogated the source of the threat (e.g., by acting aggressively towards the target), would they then also still need to demonstrate their personal adequacy in another way? Our results seem to suggest they do. More specifically, in Experiment 3.2, people displayed self-directed aggression regardless of whether they first had a chance to derogate the refuser (by allocating beer to the refuser) or not.

To conclude, people can react aggressively to confrontations with moral refusers. This aggression can be directed at the refuser, as well as themselves. People pour more beer for moral non-drinkers than for non-moral non-drinkers and inflict more pain on themselves after a confrontation with a moral non-drinker than with a non-moral non-drinker. Thus, moral refusers not only evoke psychological derogation (e.g., lower evaluations), but also actual behavior that has societal relevance.

1

2

3

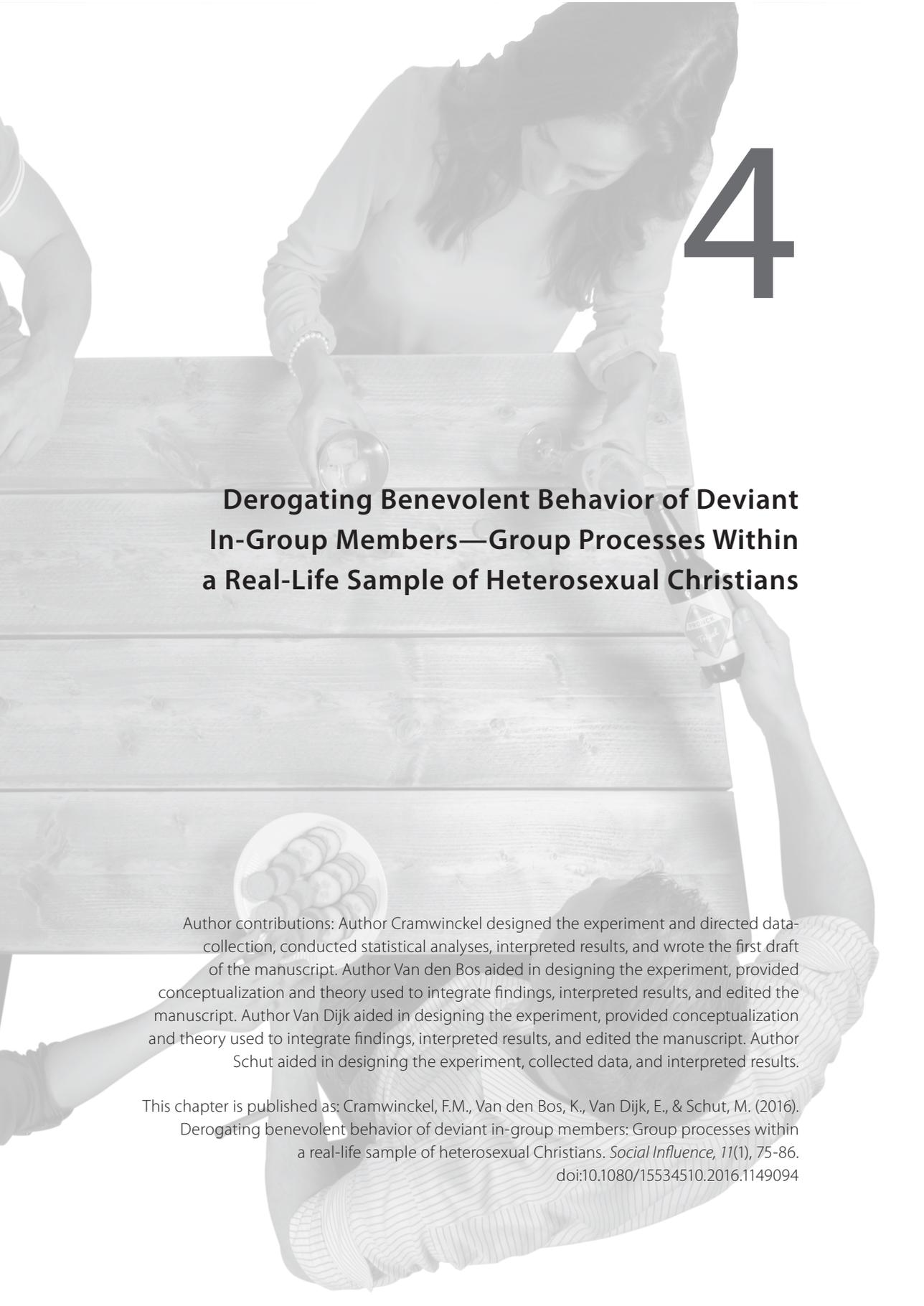
4

5

6

&





4

Derogating Benevolent Behavior of Deviant In-Group Members—Group Processes Within a Real-Life Sample of Heterosexual Christians

Author contributions: Author Cramwinckel designed the experiment and directed data-collection, conducted statistical analyses, interpreted results, and wrote the first draft of the manuscript. Author Van den Bos aided in designing the experiment, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Author Van Dijk aided in designing the experiment, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Author Schut aided in designing the experiment, collected data, and interpreted results.

This chapter is published as: Cramwinckel, F.M., Van den Bos, K., Van Dijk, E., & Schut, M. (2016). Derogating benevolent behavior of deviant in-group members: Group processes within a real-life sample of heterosexual Christians. *Social Influence, 11*(1), 75-86. doi:10.1080/15534510.2016.1149094

Abstract

How do people evaluate potentially good and desirable behavior by others? We investigate how participants ($N = 154$) evaluate a couple that wants to adopt an orphan that would otherwise die. We collected data from heterosexual Christians in two cities in the Dutch Bible belt. We manipulated whether the adoptive-parents-to-be were a heterosexual couple or a lesbian couple and whether the couple self-identified as devout Christians or not. Using a subjective group dynamics account, we predicted and found that participants evaluated the lesbian couple in more negative terms than the heterosexual couple, especially when this couple was also Christian. These findings illustrate how positive behavior is derogated when displayed by in-group deviants.

1

2

3

4

5

6

&

People's values and religious beliefs can influence how they feel about others that do not conform (strictly) to these beliefs. For example, Christians often have more negative evaluations of homosexuals than non-Christians, and many Christians reject homosexuals who have romantic relationships or raise children within those relationships (e.g., Ford, Brignall, Van Valey, & Macaluso, 2009; Fulton, Gorsuch, & Maynard, 1999). But why do (some) Christians reject homosexuality? The most obvious answer would be that it is prescribed in the Bible—concisely summarized by the well-known conservative Christian slogan "*God created Adam and Eve, not Adam and Steve*".

How would conservative Christians respond to lesbian couples that also identify as Christians? Following the rationale above, one would expect Christians to reject them regardless of whether or not these couples identify as Christians. This can be described as in-group favoritism; people have more positive evaluations of individuals who belong to the same group.

However, people also sometimes negatively evaluate their in-group members, especially when these in-group members deviate from important group norms (e.g., Marques, Abrams, Paez, Martinez-Taboada, 1998; Marques, Abrams, & Serôdio, 2001; Marques, Yzerbyt, & Leyens, 1988). In fact, in the current paper, we will demonstrate that conservative Christians have the *most negative* evaluations of a lesbian couple that also identifies as Christian, and that these evaluations are even lower than their evaluations of a non-Christian lesbian couple. More specifically, we demonstrate that behavior that would otherwise be considered as desirable (i.e., adopting an orphan child) is evaluated very negatively when displayed by in-group deviants (i.e., lesbian Christians). We argue that these negative reactions are a consequence of threats to the conservative Christians' in-group (e.g., Abrams, Rutland, Cameron, & Marques, 2003; Marques et al., 1998). We focus specifically on a well-known social psychological theory—*subjective group dynamics* (e.g., Marques et al., 1998)—to explain this observation.

The theory of Subjective Group Dynamics (SGD; e.g., Abrams, Marques, Bown, & Henson, 2000; Marques et al., 1998) explains when and why people derogate in-group deviants, and is based on two premises: First, people want to maximize and sustain intergroup differentiation. This means that people want their group to be clearly distinct from other groups. Second, people want to maximize and sustain "*the relative validity of prescriptive in-group norms through intragroup differentiation*" (Abrams et al., 2000, p. 906). This means that people want their in-group members to adhere to the prescriptive group norms. These two premises explain that although people usually favor in-group members over out-group

1

2

3

4

5

6

&

1

2

3

4

5

6

&

members, they also differentiate between in-group members. In-group members who adhere to group norms are favored over deviant members who do not follow the norms (Abrams et al., 2000; Marques et al., 1998). People even favor deviant out-group members over deviant in-group members, a phenomenon called ‘the black sheep effect’ (Marques et al., 1988).

In-group deviants are derogated because they reduce the distinction between people’s own group and relevant out-groups, and also undermine the norms of the in-group (Abrams et al., 2000). Thus, deviants can undermine a group’s positive distinctiveness, as well as the positivity of the group. This can lead other group members to *“express hostility towards deviants, redefine the deviants’ status in the group, and ultimately, ostracize them”* (cf. Marques et al., 1998, p. 437). For example, Chekroun and Nugier (2011) demonstrate that French students felt more ashamed and embarrassed when a fellow French student lit up a cigarette in a non-smoking area than when a Belgian student did so. Furthermore, French students felt a higher need to persuade the French student to change his/her behavior than to persuade the Belgian student. The black sheep effect is a way in which people reconcile their need to see their in-group in a positive light with the fact that one of their fellow group members did not adhere to the norm (e.g., Pinto, Marques, Levine, & Abrams, 2010).

Derogation of deviant in-group members may be particularly present in morally-laden situations, and deviation from moral norms may lead to especially harsh evaluations (e.g., Cramwinckel, Van den Bos, & Van Dijk, 2015). Moral norms are prescriptive norms, and people tend to discriminate against others who deviate from moral norms (Abrams, Rutland, Ferrel, & Pelletier, 2008; Marques et al., 1998; Rutland, Killen, & Abrams, 2010). Furthermore, in-group morality is very relevant for one’s self-concept, and people have a strong need to belong to moral groups (Brambilla, Sacchi, Pagliaro, & Ellemers, 2013; Ellemers, Pagliaro, Barreto, & Leach, 2008; Pagliaro, Ellemers, & Barreto, 2011). In-group members are evaluated on the moral quality of their behaviors, while morality seems to be less important for the evaluation of out-group members. Apparently, this difference occurs because out-groups do not constitute a big part of one’s identity, and the moral quality of the behavior of out-group members is therefore less relevant for the self. As a consequence, people may be less tolerant of norm deviations in the moral domain, in particular performed by an in-group member (e.g., Cramwinckel et al. 2015; Marques et al., 1998).

This notion has interesting consequences for how people evaluate benevolent behavior by deviant in-group members. One implication is that even when people perform behavior that could be considered as moral, their simultaneous deviant status may undermine the

positive distinctiveness and the prescriptive norms of the in-group (Abrams et al., 2003; Marques et al., 1998). For example, imagine a couple that wants to adopt an orphan whose family died during a natural disaster. This orphan would die if it would not be adopted. Most people probably agree that adopting this orphan is a good and desirable thing to do. Christians may consider the adoption to be especially praiseworthy, because it fits with the Christian values of taking care of people in need.

Now consider the following: The couple adopting the orphan is a lesbian couple, consisting of two Christian women. How would you expect conservative Christian participants to react now? The women in this couple could be devout Christians, pray every day, frequent church and try to follow the Ten Commandments. At the same time, their mere existence as a same-sex couple deviates from the Christian value that men and women are destined to be together. As such, the lesbian Christian couple deviates from important norms of conservative Christians and thereby undermines the group's positive social identity from within the in-group (Marques et al., 1998). Not in spite of their being Christian, but exactly because of it. Furthermore, the fact that this lesbian couple deviates from important norms of the Christian in-group should be especially clear when Christian values are salient (Marques et al., 1998). Thus, the negative response to a lesbian couple wanting to adopt an orphan should be stronger when this couple explicitly identifies as Christian. We therefore hypothesize that heterosexual Christians will evaluate lesbian couples more negatively than heterosexual couples, and will have the most negative evaluations of lesbian couples that explicitly identify as Christians.

We test this hypothesis using the scenario outlined above, where we focus on reactions of heterosexual Christians and measure their evaluations of a couple wanting to adopt a baby that would otherwise die. We manipulate whether this couple is heterosexual or lesbian, and whether this couple explicitly identified as Christian or not.

Method

Participants and design

Participants were approached in the city center of Ede and Lunteren, two cities in the Netherlands that are located in the 'Bible Belt', a strip of land where a relatively high number of conservative Christians live. A sample of 155 heterosexual Christian participants (86 women) completed the questionnaire. Ages ranged from 17 to 67 years, with a mean of 34.16 years ($SD = 10.93$). Education levels ranged from primary education (44 participants) to secondary education (98 participants) and post-secondary education (13 participants). This

1

2

3

4

5

6

&

1

2

3

4

5

6

&

was representative of the Dutch population (Statistics Netherlands [CBS], 2012). Participants were randomly allocated to one of four conditions of our 2 (religiosity of couple: Christian vs. non-religious) x 2 (sexual orientation of couple: lesbian vs. heterosexual) between-participants design, with 37 to 40 participants per cell.¹

Procedure

Potential participants were handed a flyer with the link to the online study. The study was created via the website www.EnqueteMaken.nu. On this website, participants read an informed consent form, after which they were presented one of four scenarios about a couple that was planning to adopt an orphan child from the Philippines. All participants read that in 2013 an earthquake had hit the Philippines and tens of thousands of lives were lost during this disaster. This disaster left many children as orphans, who had no chance of survival unless they would be adopted by foreign couples. After this general introduction, our manipulations followed.

In the *heterosexual couple condition*, participants read about a married couple, ‘Simon and Mariëlle’ (i.e., a typical Dutch male and female name, respectively), that was planning to adopt one of the children that were orphaned. For example, in this condition, participants read: “*I’m confident that we will be an amazing mommy and daddy*”. In the *lesbian couple condition*, participants read the same scenario but about a different married couple: ‘Simone and Mariëlle’ (i.e., two typical Dutch female names) that was planning to adopt one of the children that were orphaned. In this condition, participants read sentences such as: “*I’m confident that we will be two amazing mommies*”.

Furthermore, in the *religious couple condition*, participants read “*We are both Christians and we, as devout Christians, feel it is our duty to help one of these children*”. In the *non-religious couple condition*, participants read “*We feel it is our duty to help one of these children*”. Importantly, the phrasing of all scenarios was similar apart from the names of the couple and the two target sentences. Hereafter, the first of three manipulation checks followed, which asked participants to write down the names of the couple.

¹ We report how we determined sample size, data exclusions, manipulations, and all measures in the experiment. We aimed to collect data from two hundred participants. We stopped the recruitment of participants after we achieved this goal. Because our hypotheses were explicitly about heterosexual Christian participants, we had to exclude the data from 45 non-heterosexual participants and/or non-Christian participants from our analyses, leaving a final sample of $N = 155$ heterosexual Christian participants.

Subsequently, the main dependent variables were assessed. Participants indicated their evaluation of the adopting couple by stating their agreement with 27 statements on 7-point Likert scales ranging from 1 (*not at all*) to 7 (*very much so*). We combined these 27 items into an reliable *adoptive parents evaluation scale* ($\alpha > .99$). An overview of all items, as well as means and standard deviations, can be found in Table 4.1. Subsequently, two manipulation checks followed: Participants were asked whether the future adoption parents were (i) a man and a woman or (ii) two women, and whether the future adoption parents were (i) Christians or (ii) this was unknown/not mentioned. One participant failed one or more of these manipulation checks, and was therefore excluded from the main analyses, leaving a final sample of 154 participants.

Hereafter, demographic variables such as age, gender, education level, religion, and nationality were collected. We measured *Religiosity* with four items ($\alpha = .91$; Sethi & Seligman, 1993): “How important is religion to you?”, “How often do you read holy scripture?”, “How often do you pray”, and “How often do you attend religious activities or services?”. Answers to the first item could be given on a 7-point Likert scale ranging from 1 (*not important at all*) to 7 (*very important*). Answers to the last three items could be given on an 8-point categorical scale ranging from 1 (*more than once a day*) to 8 (*never*). Because these items were measured on different answering scales, they were standardized before being combined into the religiosity scale. The last three items were reverse coded, so that higher numbers represented a higher importance of religion. *Lesbian/Gay attitudes* were measured with 16 items of the Homosexuality Attitude Scale ($\alpha = .98$; Kite & Deaux, 1986). An example item is “I wouldn’t mind if one of my friends was gay or lesbian”. Answers could be given on 7-point Likert scales ranging from 1 (*do not agree at all*) to 7 (*agree completely*). Finally, participants indicated their own sexual orientation, could write down remarks about the study and were thanked for their participation.

Results

Checks on religiosity of participants

Only participants that self-identified as Christians were included in our final sample (see Note 1). Because we collected data in the Bible Belt, we assumed that Christians in our final sample would be relatively conservative. The religiosity measures support this assumption. Participants considered religion as important in their lives, and rated it well above the midpoint of the 7-point scale ($M = 5.66$, $SD = 0.97$). Furthermore, 86% percent of participants visited religious activities or services for once a month or more, 82% of participants prayed once a day or more, and almost 45% of participants read holy scripture once a month or more.

1

2

3

4

5

6

&

1

Main analyses

2

3

4

5

An analysis of variance (ANOVA) with our sexual orientation and religiosity manipulations as independent variables, and the adoptive parents evaluation scale as the dependent variable yielded a significant main effect of sexual orientation, $F(1, 150) = 1104.83, p < .001, \eta_p^2 = .88$, indicating that the heterosexual couple was rated more positively ($M = 5.81, SD = 0.49$) than the lesbian couple ($M = 2.30, SD = 0.93$). As expected, this effect was qualified by a significant sexual orientation \times religiosity interaction, $F(1, 150) = 45.03, p < .001, \eta_p^2 = .23$. We did not observe a significant main effect of religiosity of the couple, $p = .443$. Figure 4.1 shows these effects.

6

&

To interpret the interaction, we inspected the simple main effects. This demonstrated that when participants read about the adoption by a heterosexual couple, their evaluations were more positive when the couple was religious ($M = 6.22, SD = 0.28, 95\% \text{ CI } [6.01, 6.43]$) rather than non-religious ($M = 5.43, SD = 0.30, 95\% \text{ CI } [5.22, 5.63]$), $F(1, 150) = 27.96, p < .001, \eta_p^2 = .16$. In line with subjective group dynamics and our hypothesis, the reversed pattern was observed when participants read about the prospective adoption by a lesbian couple: Now their evaluations were *less* positive when the couple was religious ($M = 1.99, SD = 1.09, 95\% \text{ CI } [1.78, 2.20]$) rather than non-religious ($M = 2.62, SD = 0.61, 95\% \text{ CI } [2.41, 2.83]$), $F(1, 150) = 17.66, p < .001, \eta_p^2 = .11$.

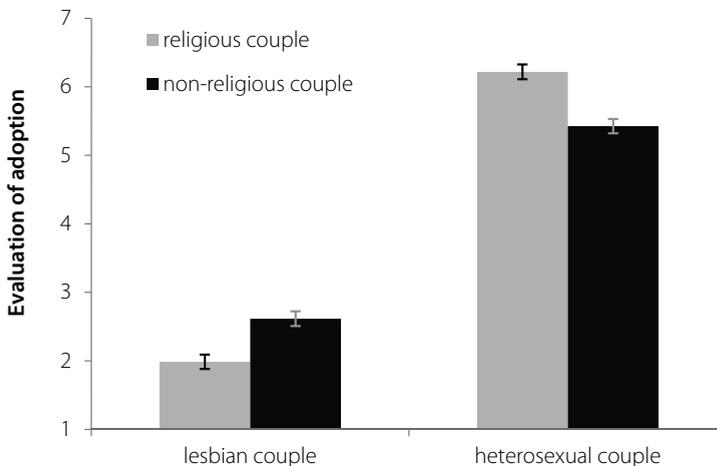


Figure 4.1. Evaluation of adoptive parents, as a function of sexual orientation and religiosity of the couple. Higher numbers indicate more positive evaluations; error bars are plotted one standard error above and below the mean.

Simple main effects also demonstrated that when people read about the adoption by a Christian couple, they had more positive evaluations of the couple when the couple was heterosexual ($M = 6.22, SD = 0.28, 95\% \text{ CI } [6.01, 6.43]$) rather than lesbian ($M = 1.99, SD = 1.09, 95\% \text{ CI } [1.78, 2.20]$), $F(1, 150) = 787.74, p < .001, \eta_p^2 = .84$. When people read about the adoption by a non-Christian couple, the pattern was similar (but less strong), with participants having more positive evaluations of the couple when the couple was heterosexual ($M = 5.43, SD = 0.30, 95\% \text{ CI } [5.22, 5.63]$) rather than lesbian ($M = 2.62, SD = 0.61, 95\% \text{ CI } [2.41, 2.83]$), $F(1, 150) = 356.52, p < .001, \eta_p^2 = .70$. These findings also fit with subjective group dynamics because deviants (i.e., the lesbian couple) were evaluated more harshly when they are in-group members (i.e., Christians) rather than out-group members (i.e., non-Christians).²

Additional analyses

Covariates. To control for the possible influence of people's attitudes towards homosexuals and their religiosity, and thereby reduce error variance, we performed a GLM analysis with our two manipulations (sexual orientation and religiosity of the couple) as independent variables, importance of religion and lesbian/gay attitudes as covariates, and the adoptive parents evaluation scale as the dependent variable. Results mimicked those reported under Main Analyses, but were somewhat stronger than the results reported earlier. More specifically, we observed a significant main effect of sexual orientation, $F(1, 148) = 1341.48, p < .001, \eta_p^2 = .90$, indicating that our participants evaluated the heterosexual couple more positively ($M = 5.73, SD = 0.49, 95\% \text{ CI } [5.61, 5.86]$) than the lesbian couple ($M = 2.39, SD = 0.93, 95\% \text{ CI } [2.27, 2.52]$). We also observed a significant main effect of religion, $F(1, 148) = 3.99, p = .048, \eta_p^2 = .03$, indicating that our participants evaluated the religious couple

² One might wonder whether there were sex differences or gender differences in our results, because men and women sometimes differ in their attitudes towards homosexuals (e.g., Kite & Whitley, 1996). Running a GLM on the evaluation scale with sex added as an extra between subject factor demonstrated that sex did not qualify the results. Although there was a main effect of sex, with women having somewhat more positive evaluations ($M = 4.46, SD = 1.91, 95\% \text{ CI } [4.05, 4.33]$) than men ($M = 3.54, SD = 1.80, 95\% \text{ CI } [3.73, 4.06]$), $F(1, 146) = 7.07, p = .009, \eta_p^2 = .05$, there were no significant two-way or three-way interactions between our experimental manipulations and participants' sex (all p 's $> .846$). Furthermore, running the main analyses for men and women separately yielded similar results for both sexes, with a significant interaction between sexual orientation and religiosity of the couple, for men, $F_{\text{men}}(1, 65) = 16.05, p < .001, \eta_p^2 = .20$, and women, $F_{\text{women}}(1, 81) = 25.12, p < .001, \eta_p^2 = .24$. This effect seemed somewhat stronger for women than for men, but this may have also been due to the fact that there were more women ($n = 85$) than men ($n = 69$) in our final sample. These results are in line with the results of a meta-analysis which demonstrated that although men have more negative attitudes about homosexuals than women do, men and women have similar attitudes about lesbians (Kite & Whitley, 1996). We did not measure participants' gender identities, so we could not include those in the analyses.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

more positively ($M = 4.16$, $SD = 2.27$, 95% CI [4.03, 4.28]) than the non-religious couple ($M = 3.97$, $SD = 1.49$, 95% CI [3.84, 4.10]). These effects were qualified by a Sexual orientation x Religiosity interaction, $F(1, 148) = 60.29$, $p < .001$, $\eta_p^2 = .29$. Furthermore, lesbian/gay attitudes were significantly related to evaluations, $F(1, 148) = 52.65$, $p < .001$, $\eta_p^2 = .26$, such that more positive lesbian/gay attitudes were related to more positive evaluations ($b = 0.39$, 95% CI [0.28, 0.49]). No significant main effect of the importance of religion was observed.

To interpret the Sexual orientation x Religiosity interaction, we inspected the simple main effects. This demonstrated that when participants read about the prospective adoption by a heterosexual couple, their evaluations were more positive when the couple was religious ($M = 6.17$, $SD = 0.28$, 95% CI [5.99, 6.35]) rather than non-religious ($M = 5.30$, $SD = 0.30$, 95% CI [5.12, 5.47]), $F(1, 148) = 46.38$, $p < .001$, $\eta_p^2 = .24$. The reversed pattern was observed when participants read about the prospective adoption by a lesbian couple: Now their evaluations were *less* positive when the couple was religious ($M = 2.14$, $SD = 1.09$, 95% CI [1.96, 2.32]) rather than non-religious ($M = 2.64$, $SD = 0.61$, 95% CI [2.47, 2.81]), $F(1, 148) = 15.08$, $p < .001$, $\eta_p^2 = .09$. Participants thus had the most negative evaluations of the lesbian couple that identified as Christian.

Simple main effects demonstrated that when people read about the adoption by a Christian couple, they had more positive evaluations of the couple when the couple was heterosexual ($M = 6.17$, $SD = 0.28$, 95% CI [5.99, 6.35]) rather than lesbian ($M = 2.14$, $SD = 1.09$, 95% CI [1.96, 2.32]), $F(1, 148) = 984.08$, $p < .001$, $\eta_p^2 = .87$. When people read about the adoption by a non-Christian couple, the pattern was similar (but less strong), with participants having more positive evaluations of the couple when the couple was heterosexual ($M = 5.30$, $SD = 0.30$, 95% CI [5.12, 5.47]) rather than lesbian ($M = 2.64$, $SD = 0.61$, 95% CI [2.47, 2.82]), $F(1, 148) = 445.46$, $p < .001$, $\eta_p^2 = .75$. These findings demonstrate that deviants (i.e., the lesbian couple) were evaluated more harshly when being in-group members (i.e., Christians) rather than out-group members (i.e., non-Christians).

Subscales. Although we combined all dependent variables into one scale, face value may suggest different theoretical subconstructs. More specifically, there are 12 items that seem related to participants' evaluations of the prospective adoption (items "*heartwarming*" to "*moral*" in Table 4.1, Cronbach's $\alpha = .99$), 5 items that seem related to participant's emotional reactions (items "*happy*" to "*irritated*" in Table 4.1, Cronbach's $\alpha = .99$), 4 items that seem related to perceived similarities between participants and the adoptive parents (items "*your moral values are similar to [Parent 1] and [Parent 2]*" to "*you could be friends with [Parent 1] and [Parent 2]*" in Table 4.1, Cronbach's $\alpha = .98$), and 4 items that seem to be related to evaluations of the adoptive parents (items "*[Parent 1] and [Parent 2] are good people*" to "*[Parent 1] and [Parent 2] adopt a child out of goodness of their heart*" in Table 4.1, Cronbach's $\alpha = .97$).

Table 4.1.
Means and Standard Deviations for Adoptive Parents Evaluation Scale and Separate Items

	Heterosexual couple		Lesbian couple	
	religious couple	non-religious couple	religious couple	non-religious couple
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Adoptive parents evaluation scale	6.22 (0.28)	5.43 (0.30)	1.99 (1.09)	2.62 (0.61)
Separate items				
<i>To what extent do you think the adoption is..</i>				
heartwarming	5.70 (0.74)	4.95 (0.82)	2.10 (1.06)	2.84 (0.97)
acceptable	6.14 (0.42)	5.53 (0.55)	1.93 (1.25)	2.42 (0.95)
repulsive (r)	1.78 (0.42)	2.08 (0.62)	5.53 (1.20)	5.08 (0.71)
good	6.19 (0.46)	5.38 (0.54)	1.90 (1.30)	2.34 (0.71)
objectionable (r)	1.16 (0.37)	1.78 (0.53)	5.78 (1.33)	5.29 (0.84)
admirable	6.73 (0.45)	5.98 (0.70)	2.00 (1.54)	2.84 (1.08)
naturally	5.78 (0.53)	5.05 (0.64)	1.43 (1.03)	1.66 (0.67)
unacceptable (r)	1.30 (0.46)	1.98 (0.48)	5.83 (1.38)	5.47 (0.76)
unusual (r)	1.59 (0.55)	2.48 (0.55)	6.15 (1.03)	5.76 (0.82)
disgusting (r)	1.00 (0.00)	1.15 (0.36)	5.65 (1.48)	4.87 (0.93)
pleasant	5.78 (0.58)	4.98 (0.58)	1.75 (1.15)	2.39 (0.64)
obnoxious (r)	1.35 (0.48)	2.23 (0.58)	5.88 (1.28)	5.24 (0.85)
moral	6.49 (0.51)	5.55 (0.55)	1.40 (1.03)	1.95 (0.90)
<i>To what extent does the adoption makes you..</i>				
happy	5.76 (0.60)	4.93 (0.62)	1.88 (1.04)	2.32 (0.66)
restless (r)	1.38 (0.49)	2.00 (0.45)	5.68 (1.29)	5.16 (0.72)
angry (r)	1.41 (0.50)	1.90 (0.50)	5.45 (1.26)	4.89 (0.95)
enthusiastic	5.51 (0.56)	4.93 (0.66)	1.88 (1.07)	2.50 (0.69)
irritated (r)	1.46 (0.51)	2.15 (0.53)	5.70 (1.36)	5.11 (0.86)
<i>To what extent do you think...</i>				
you agree with the adoption	6.16 (0.50)	5.58 (0.50)	1.88 (1.38)	2.29 (0.80)
your moral values are similar to Simon(e) and Mariëlle	5.84 (0.44)	5.10 (0.59)	1.80 (1.07)	2.63 (0.82)
you have a lot in common with Simon(e) and Mariëlle	5.70 (0.66)	4.80 (0.76)	1.70 (0.88)	2.42 (0.68)
you have the same ideas and values as Simon(e) and Mariëlle	5.95 (0.40)	5.15 (0.53)	1.75 (0.98)	2.55 (0.86)
you could be friends with Simon(e) and Mariëlle	5.19 (1.20)	4.00 (1.20)	1.60 (0.71)	2.55 (0.65)

1

2

3

4

5

6

&

Table 4.1.
Means and Standard Deviations for Adoptive Parents Evaluation Scale and Separate Items (continued)

	Heterosexual couple		Lesbian couple	
	religious couple	non-religious couple	religious couple	non-religious couple
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Simon(e) and Mariëlle are good people	6.43 (0.50)	5.58 (0.55)	2.35 (1.25)	3.47 (0.89)
Simon(e) and Mariëlle adopt a child out of selfish reasons	1.46 (0.56)	2.35 (0.70)	6.13 (0.85)	5.37 (0.59)
Simon(e) and Mariëlle are religious	6.08 (0.60)	3.90 (0.90)	1.70 (0.99)	2.08 (0.63)
Simon(e) and Mariëlle adopt a child out of goodness of their heart	6.35 (0.63)	5.23 (0.73)	2.28 (1.30)	3.61 (0.97)

Note. Items with (r) were reverse coded before they were included in the adoption evaluation scale, such that higher scores on the adoption evaluation scale indicate more positive ratings.

All subscales strongly correlated with each other, with Pearson's correlations ranging between $r = .957$ and $r = .993$ (all p 's < .001). This means that between 91% and 98% of the variation in one of the subscales may be accounted for by the other subscales (r^2 's .91 - .98). Therefore, we considered it more appropriate to combine items into one general scale, rather than dividing them into separate subscales.

However, to further explore the underlying structure of the items in more detail, we also performed a principal components analysis (PCA). Results demonstrated that a one-component solution provided the best fit. More specifically, 92% of the variance was explained by the first component, which had an eigenvalue of 24.86. All subsequent components had eigenvalues below 0.47 (and therefore below the critical cut-off point of eigenvalue > 1). The scree plot also displayed a sharp drop after the first component. Furthermore, all items loaded strongly ($>|.856|$) on this first component. All in all, this supports the notion that all items reflect the same underlying construct. Therefore, we think our data are best interpreted by performing analyses on the general scale, rather than separate subscales.

Discussion

Our study demonstrates an interesting phenomenon: Behavior that is generally seen as admirable and praiseworthy (i.e., adoption of an orphan) becomes negatively evaluated if displayed by in-group deviants. More specifically, Christians in the Dutch Bible Belt evaluated a scenario of a couple that wanted to adopt a child that was orphaned during

a natural disaster and that would otherwise die. As expected, people had more positive evaluations of the heterosexual couple than of the lesbian couple. Of most interest to our current line of reasoning, and in line with subjective group dynamics, the lesbian Christian couple was evaluated most negatively (and thus even more negatively than the non-Christian lesbian couple).

These findings build on work on subjective group dynamics (SGD), which explains when people will derogate deviant in-group members (e.g., Abrams et al., 2003; Marques et al., 1998; 2001). We demonstrate that heterosexual Christians derogate a lesbian couple that wants to adopt an orphan child. In line with SGD, derogation is strongest when the lesbian couple explicitly identifies as Christian. Previous work demonstrated that people derogate in-group deviants especially when deviants show immoral behavior, and thereby undermine moral norms. For example, Abrams and colleagues (2008) demonstrated that children have more positive evaluations of children that adhere to moral norms such as fairness (e.g., turn taking and sharing with others), and derogate others that undermine these moral norms (e.g., by being selfish). We extend this work by demonstrating that people not only derogate in-group deviants when they display immoral behavior, but also when they display behavior that would otherwise be considered as highly normative and moral.

Our findings may be relevant for understanding and reducing prejudice within (religious) communities. One might be inclined to perceive people who respond negatively to homosexuals and lesbians as bigoted and hateful. However, we suggest a more nuanced story: These negative reactions may be caused by a perceived threat to the in-group identity. Heterosexual Christians have the most negative evaluations of a couple that is not only lesbian, but also Christian. This could mean that this couple is perceived as deviating from the in-group (by being Christian but also lesbian), which is more threatening than deviating from the out-group (by being non-Christian and also lesbian), and therefore leads to harsher evaluations (e.g., Marques et al., 1988; Pinto et al., 2010). This interpretation can have important implications for the reduction of negative reactions. For example, buffering the (personal and social) identities of religious people may make them more resistant to threats, which may lead to less prejudicial responses towards gay and lesbian Christians. It may be interesting for future research to investigate whether conservative Christians become more tolerant of deviating members of their congregation (e.g., lesbians or homosexuals) after engaging in group- or self-affirmation (e.g., Cohen & Sherman, 2014). If so, this may offer religious leaders tools to increase tolerance within their community, for instance by including affirmational elements in their sermons.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

One may wonder whether sexual orientation or religiosity was the most salient social identity. Was the lesbian Christian couple considered deviant for being lesbians rather than heterosexuals, or because they were Christians rather than non-Christians? Our study was not designed to investigate which of these two social identities was the most important to our participants, but rather to test our assumption that conservative Christians would have the most negative evaluations of lesbian couples that also identified as Christian. To make matters more complicated, having a heterosexual identity is a keystone in the Christian tradition (remember the slogan about Adam and Steve?), which makes it difficult to disentangle these two identities. Nonetheless, we found negative evaluations of in-group deviants regardless of whether we considered sexual orientation or religiosity to form the most prominent in-group.

A somewhat related point concerns the salience of the manipulated couple's religiosity. In the Christian condition, the couple's religious affiliation is very salient (*"We are both Christians, and we, as devout Christians, feel it is our duty to help one of those children"*). In the non-Christian condition, religious affiliation was not made salient (*"We feel it is our duty to help one of those children"*). So our experimental manipulation may have contrasted salient group membership to the Christian and adherence to Christian norms, with non-salient group membership. This increased salience of Christianity may have triggered more negative reactions towards those who undermine important Christian norms (i.e., the lesbian Christian couple who wanted to adopt a child). In fact, this understanding of our manipulations fits perfectly with a subjective group dynamics account. According to SGD, the evaluation of deviant in-group members should be most negative when *"attention is focused on in-group norms"* (Marques et al., 1998, p. 977), and salience of in-group norms should thus lead to increased intragroup differentiation.

The dependent variable in this study (the adoptive parents evaluation scale) was formed out of a large number of different items, measuring different aspects of evaluation. The adoption itself was evaluated, as were emotions about the adoption, the couple's character (and people's similarities with the couple. See Table 4.1 for a complete overview of the items. The univariate interaction effects between the manipulated orientation and religiosity of the couple were significant for all items separate (all p 's < .005), meaning that we found evidence of the derogation of the lesbian Christian couple on all 27 items. This is interesting because it seems to contrast earlier findings that Christians reject homosexual acts but not homosexual people (e.g., Ford et al., 2009). How to explain this difference? One possible reason that we found derogation on every item may be that it was impossible for participants to see the couple separate from their behavior. The women in the couple

were having a romantic relationship with each other, and wanted to raise a child within this relationship. This particular scenario may have undermined group norms so strongly, that it was no longer possible for our participants to view the women separate from their (intended) behavior. We suggest that the extremely negative evaluations were a way in which participants coped with this aversive scenario and protected their own identities from being tainted by the 'sinning' couple.

We occasionally used the term 'threat' to describe the aversive state people experience when being confronted with deviant in-group members. Important to note is that we have not measured (the perception of) threat in this study. Therefore, it would be an interesting and relevant direction for future research to investigate experienced threat more directly.

Finally, basic social psychological phenomena are often investigated in lab-settings with student samples. Although these studies provide relevant insight into fundamental social psychological processes, they also raise criticisms regarding generalizability to other samples and settings (e.g., Henrich, Heine, & Norenzayan, 2010). The current study used a non-student sample that was collected in the streets of two Dutch cities, and that may better reflect the Dutch population than the average student sample. For instance, only about 8 percent of the current sample completed post-secondary education, which is a relatively good reflection of higher education levels in the Dutch population (CBS, 2012). Importantly, our results offer strong support for the process of subjective group dynamics. For example, the mean difference between evaluations of the heterosexual couple and the lesbian couple are approximately 3.5 points on a 7 point scale, with an effect size of .88 for this main effect, which is extremely large. This increases confidence that important social psychological phenomena are relevant for understanding of our daily lives. The current findings suggest that derogating benevolent behavior of deviant in-group members by heterosexual Christians may be among these phenomena.

1

2

3

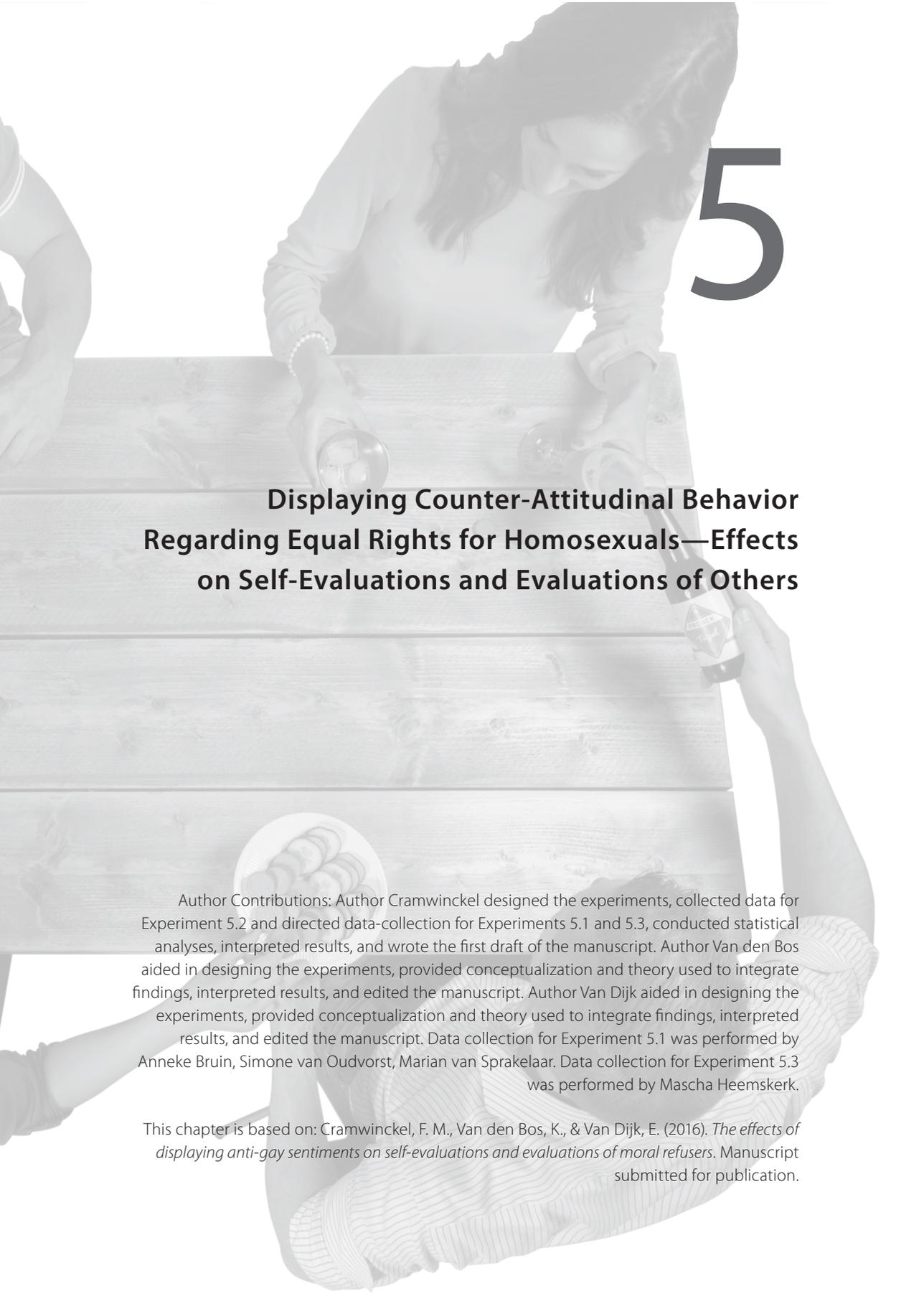
4

5

6

&





5

Displaying Counter-Attitudinal Behavior Regarding Equal Rights for Homosexuals—Effects on Self-Evaluations and Evaluations of Others

Author Contributions: Author Cramwinckel designed the experiments, collected data for Experiment 5.2 and directed data-collection for Experiments 5.1 and 5.3, conducted statistical analyses, interpreted results, and wrote the first draft of the manuscript. Author Van den Bos aided in designing the experiments, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Author Van Dijk aided in designing the experiments, provided conceptualization and theory used to integrate findings, interpreted results, and edited the manuscript. Data collection for Experiment 5.1 was performed by Anneke Bruin, Simone van Oudvorst, Marian van Sprakelaar. Data collection for Experiment 5.3 was performed by Mascha Heemskerk.

This chapter is based on: Cramwinckel, F. M., Van den Bos, K., & Van Dijk, E. (2016). *The effects of displaying anti-gay sentiments on self-evaluations and evaluations of moral refusers*. Manuscript submitted for publication.

Abstract

We investigate the consequences of engaging in counter-attitudinal behavior regarding equal rights for same-sex couples among participants from a pro-gay rights society. Experiment 5.1 demonstrates that these participants have more negative evaluations of themselves when they write an anti-gay essay (vs. a pro-gay essay). Experiment 5.2 shows that negative self-evaluations only occur when participants feel their behavior is visible to others and it is difficult to view their own behavior as moral. Experiment 5.3 reveals that participants only have more negative self-evaluations when they actively engage in counter-attitudinal behavior. In all experiments, participants have positive evaluations of others who refuse to engage in counter-attitudinal behavior out of moral concern. Implications are discussed.

1

2

3

4

5

6

&

Gay rights are the focus of debate in many countries. Especially same-sex marriage has attracted attention. In 2001, the Netherlands legalized same-sex marriage. Since then, same-sex marriage has been legalized in twenty other countries, ten of which in the last three years (www.freedomtomarry.org). In 2015, Ireland legalized same-sex marriage by popular vote (www.marriagequality.ie). Later that year, the US supreme court ruled same-sex marriage as a right protected by the US constitution in all 50 states (www.gaymarriage.procon.org); marking the historic significance of this day for the gay rights movement.

In short, more and more people support marriage equality. For example, in the Netherlands, over eighty percent of people supports same-sex marriage (The Netherlands Institute for Social Research [SCP], 2013). What do you imagine would happen when people who endorse equal rights are asked to write an essay that explicitly condemns equal rights for homosexuals? Would they display this behavior that clearly opposed their values? How would this impact their self-concepts if they would? How would they feel about others who refuse to comply?

From classic social psychological research, we know that people often engage in behaviors that do not fit with their attitudes or moral beliefs. For example, people can be persuaded to lie about how much they like a boring task (Festinger & Carlsmith, 1959), conform to the opinion of the majority even knowing the majority is wrong (Asch, 1951), and administer electric shocks to others when someone else instructs them to do so (Milgram, 1963). It may therefore come as no surprise that almost all participants in our current experiments complied with the assignment and wrote anti-gay essays when instructed to do so. But how does engaging in counter-attitudinal behavior affects people's self-concepts?

Consequences for the Self

The theory of moral self-regulation proposes that people's moral self-concepts are based on their own previous behavior (e.g., Zhong, Liljenquist, & Cain, 2009). People compare their behavior with their moral ideals. When people do not live up to their moral ideals, their moral self-concept is threatened and needs to be restored. Based on this, we expect that people have more negative evaluations of themselves after showing behavior that opposes their moral values (counter-attitudinal behavior) than after showing behavior that aligns with their moral values (pro-attitudinal behavior). In our first experiment, this means that when people endorse equal rights for homosexuals, they should have more negative self-evaluations after writing an essay against gay rights (vs. in favor of gay rights).

1

2

3

4

5

6

&

1

2

3

4

5

6

&

This effect may be stronger in some situations than in others. Self-concept maintenance theory states that while people want to maintain a positive view of themselves as moral individuals, they can be tempted to display moral transgressions (Mazar, Amir, & Ariely, 2008). Whether people stick to their moral values or succumb to temptation depends on how well they are able to maintain a positive self-concept while transgressing. This is—among other things—influenced by how easy it is for people to see their behavior as moral (Mazar et al., 2008) and the salience of their own behavior and moral standards (Diener & Wallbom, 1976); aspects we manipulate in Experiment 5.2.

Consequences for the Other

Self-concepts are not the only under-investigated aspect of counter-attitudinal behavior. Counter-attitudinal behavior often takes place in a social setting. Therefore, in the current studies, we expose people to the behavior of another participant. More specifically, participants are first invited to engage in counter-attitudinal behavior themselves, and afterwards read the refusal of someone else to comply. We investigate a social situation where people are exposed to the motivated refusal of someone else, which allows us to assess people's evaluations of these others.

Would people appreciate or derogate others who refuse to engage in counter-attitudinal behavior? We can formulate competing predictions. On the one hand, one might expect people to react negatively. If participants complied to the request while not fully acknowledging their behavior was questionable, the refusal of the other person is essentially new information that imposes an (unanticipated) challenge to one's morality. This may lead participants in anticipated reproach (Chapter 2; Cramwinckel, Van den Bos, & van Dijk, 2015; Minson & Monin, 2012). On the other hand, it could be that participants have positive evaluations of a refuser of an anti-gay essay. This might happen if participants already realized that it had been wrong¹ for them to comply to the task. In that case, internal alarm bells (Zhong et al., 2009) have already gone off, and the other's refusal is not necessarily interpreted as new and unwelcomed attack. In this case, participants may not feel the need to "shoot the messenger", because the (internal) message was already well received. This may lead them to appreciate the other's moral character. An important goal of our first experiment is to examine which of the two competing predictions receives the strongest support.

¹ We occasionally refer to writing an anti-gay essay as 'wrong' or 'immoral'. These terms reflect the mindset of our participants who endorse equal rights, and do not necessarily reflect objective moral truths.

Experiment 5.1

In this experiment, participants indicated their own attitudes towards homosexuality and were asked to write an essay that either opposed equal rights for homosexuals (the anti-gay essay condition) or endorsed those rights (the pro-gay essay condition). We verified that most participants had relatively positive attitudes about homosexuals and disagreed with the statement that homosexuals should *not* have the same rights as heterosexuals. After having written an essay themselves, participants read the bogus reaction of a person who refused to write an anti-gay essay because being gay was normal and acceptable (i.e., the anti-essay refuser) or had refused to write a pro-gay essay because being gay was abnormal and unacceptable (i.e., the pro-essay refuser). Afterwards, we measured participants' evaluations of the refuser and of themselves. As outlined above, we expected that participants would have more negative evaluations of themselves after writing an anti-gay essay, rather than a pro-gay essay (Hypothesis 1). Furthermore, we investigated whether participants would negatively (Hypothesis 2a) or positively (Hypothesis 2b) evaluate the refuser of the anti-gay essay.

Although we were primarily interested in participants' reactions towards the refuser of the anti-gay essay, we included the refuser of the pro-gay essay condition. By doing so, we could rule out the possibility that reactions would be driven by the fact that the refuser was assigned a different task than the participant. However, the pro-essay refuser makes a moral claim that is counter-normative in Dutch society. We discuss the possible implications of this condition in the Discussion section.

Method

Participants and Design. Utrecht University students ($N = 207$) participated in this experiment in exchange for a monetary reward or course credits ($M_{\text{age}} = 21.22$ years, $SD_{\text{age}} = 3.74$; 130 women, 76 men, 1 participant failed to provide demographic information). Participants were randomly assigned to one of the conditions of our 2 (Own essay: pro-gay essay vs. anti-gay essay) \times 2 (Refuser: anti-essay refuser vs. pro-essay refuser) experimental design. The number of participants within each condition varied between 50 and 53.²

² In all experiments, we aimed to collect a minimum of 50 participants per cell, and stopped data collection at the end of the week in which we achieved this aim. We report all data exclusions, all manipulations, and all measures. In Experiment 5.1, 29 additional participants were not included in the analyses: 18 participants participated more than once; we only included data from their first participation. Furthermore, we had two a priori criteria to exclude participants from the analyses: (1) refusing to write the essay (7 participants), and (2) writing an essay of one sentence or less

1

2

3

4

5

6

&

1

Procedure. Participants were seated in a cubicle behind a computer. We assessed participants' attitude regarding homosexuality by asking them to indicate the extent to which they were negative or positive towards homosexuality on a continuous slider ranging from 0 to 100 (Haddock, 1993).

2

3

Hereafter, the essay task started. Participants were asked to write an essay from the viewpoint of a passionate adversary of gay rights (the anti-gay essay condition) or from the viewpoint of a passionate advocate of gay rights (the pro-gay essay condition). In the anti-gay essay condition, participants were asked to defend the viewpoint that homosexuality is abnormal and unnatural and that gays should not have the same rights as straight people. In the pro-gay essay condition, participants were asked to defend the viewpoint that homosexuality is normal and natural, and that gays should have the same rights as straight people. They were asked to imagine as lively as possible that they agreed with the provided position. In both conditions, participants learned that their essay could be evaluated by another participant. In reality, the essays were not seen by other participants.

4

5

6

&

Given the fact that 87% of the Dutch population supports equal rights for lesbians and gay men (The Netherlands Institute for Social Research [SCP], 2013), we anticipated that most participants in our sample would be relatively positive towards homosexuality. We verified this assumption by letting participants, before they wrote their essay, indicate to what extent they agreed with their assigned statement on a 100-point scale ranging from 0 (*disagree completely*) to 100 (*agree completely*). As expected, participants in the pro-gay essay condition agreed more with their assigned position ($M = 90.35$, $SD = 20.73$) than participants in the anti-gay essay condition ($M = 12.75$, $SD = 23.07$), $F(1, 203) = 647.82$, $p < .001$, $\eta^2 = .76$. Furthermore, participants' attitudes towards homosexuality were on average positive ($M = 71.61$, $SD = 25.05$).

Afterwards, participants were exposed to the reaction of someone who had refused to write the essay out of moral concern. In the *anti-essay refuser* condition, the other participant refused to write the anti-gay essay because he/she thought homosexuality was normal and moral, and thus considered it to be morally wrong to write an anti-gay essay. In the *pro-essay refuser* condition, the other participant refused to write the pro-gay essay because

(4 participants). These participants were excluded from the research, resulting in a final N of 207. Besides the measures reported in Experiment 5.1, we also measured participants' moral identity (11 items, $\alpha = .73$, Aquino & Reed, 2002) and social comparison orientation (11 items, $\alpha = .83$, Gibbons & Buunk, 1999). Stimulus materials and data are available on request.

he/she thought homosexuality was abnormal and immoral, and thus considered it to be morally wrong to write a pro-gay essay. In both conditions, participants were instructed to read the reaction of the other participant carefully, and try to form an impression of this other person's character.

Subsequently, the dependent variables were collected. *Refuser evaluation* was measured with 45 items, assessing the extent to which the refuser was perceived as being nice and honest, amongst other things. Table 5.1 shows all items. *Self-evaluation* was measured with 23 items, such as the extent to which participants felt satisfied with themselves. Table 5.2 shows all items. Answers were given on 7-point Likert scales, ranging from 1 (*not at all*) to 7 (*very much so*).

Hereafter, we asked participants what they thought the research was about, and if they had noticed something unusual during the experiment. Last, demographic variables such as age and gender were asked. Hereafter, participants were thanked, debriefed, and rewarded for their participation.

1
2
3
4
5
6
&

Table 5.1
Varimax Rotated Component Loadings on Refuser Evaluation Items in Experiment 5.1

Items	Likability of refuser $\alpha = .98$	Agency of refuser $\alpha = .91$
<i>The other participant seems to be...</i>		
unjust (reverse coded)	-.85	
unkind (reverse coded)	-.85	
kind	.85	
benevolent	.84	
unpleasant (reverse coded)	-.83	
good	.83	
bad (reverse coded)	-.82	
obnoxious (reverse coded)	-.82	
warm	.82	
unreasonable (reverse coded)	-.81	
pleasant	.81	
cold (reverse coded)	-.80	
a good friend	.80	
just	.80	
reasonable	.79	

Table 5.1
Varimax Rotated Component Loadings on Refuser Evaluation Items in Experiment 5.1 (continued)

2	Items	Likability of refuser $\alpha = .98$	Agency of refuser $\alpha = .91$
3	cruel (reverse coded)	-.79	
4	annoying (reverse coded)	-.77	
5	pleasant to work with	.77	
6	unethical (reverse coded)	-.75	
&	moral	.74	
	ethical	.73	
	immoral (reverse coded)	-.73	
	generous	.69	
	dishonest (reverse coded)	-.62	
	intelligent	.57	.55
	stingy (reverse coded)	-.54	
	someone with low self-esteem (reverse coded)		-.84
	someone with high self-esteem		.83
	confident		.82
	insecure (reverse coded)		-.77
	weak (reverse coded)		-.73
	strong		.67
	independent		.65
	dependent (reverse coded)		-.58
	mature	.50	.57
	immature (reverse coded)	-.46	-.54
	stupid (reverse coded)	-.41	-.51
	passive (reverse coded)		-.47
	honest		.42
	active		
	<i>to what extent do you ... the refuser?</i>		
	reject (reverse coded)	-.68	
	trust	.68	
	admire	.68	
	respect	.63	.46
	despise (reverse coded)	-.57	

Note. Component loadings > |.55| are printed in bold. Component loadings < |.30| are not shown (Field, 2009).

Results

Scale construction. The first step in the data analysis was to create scales to measure evaluation of the refuser and the self. We performed Principal Component Analyses (PCAs) with orthogonal (varimax) rotation to create scales, and tested our hypotheses by conducting analyses of variance (ANOVAs) on these scales.

Self-evaluation. We performed an exploratory PCA to investigate the underlying structure of the 23 items that measured self-evaluation. Four components had eigenvalues > 1 . Eigenvalues did not decrease much after the second component. The scree plot showed a large indent after the first component, indicating the underlying structure of a single component. The first component explained 43.66% of the variance.

Hereafter, we performed PCAs, fixed with one, two, or three components. The one-component solution had the most items with good factor loadings ($> |.55|$) on this factor (Tabachnick & Fidell, 2007). Therefore, we opted for the one-component solution. This component was interpreted as *Self-evaluation* ($\alpha = .95$). Only items that loaded strongly ($> |.55|$) were included, which led to the exclusion of five items. Table 5.2 shows which items were included in the self-evaluation scale and how strongly they loaded on this scale.

Refuser evaluation. We performed an exploratory PCA to investigate the underlying structure of the 45 items that measured refuser evaluation. Five components had eigenvalues > 1 . Eigenvalues did not decrease much after the second component. The first two components explained 62.49% of the variance. The scree plot also suggested a two-component solution.

Hereafter, we performed PCAs fixed with one, two, or three components. The two-component solution had the most items with good factor loadings ($> |.55|$) on one of the two components (Tabachnick & Fidell, 2007). Therefore, we opted for the two-component solution. These two components were interpreted as *Likability of the refuser* (e.g., the extent to which the refuser is perceived as kind and good. $\alpha = .98$) and *Agency of the refuser* (e.g., the extent to which the refuser is perceived as confident and independent. $\alpha = .91$). Only items with good factor loadings ($> |.55|$) and without cross loadings (i.e., factor loadings $> |.55|$ on two factors) were included, which led to the exclusion of seven items. Table 5.1 shows which items belonged to which scale, as well as their loadings on each scale.

Main analyses. To test our predictions, we performed ANOVAs with the Own Essay and Refuser manipulations as the independent variables, and our outcome measures as the dependent variables.

1

2

3

4

5

6

&

1

Self-evaluations. Results only showed a significant effect of own essay, ($F[1, 203] = 15.40$, $p < .001$, $\eta_p^2 = .07$), indicating that participants had more positive evaluations of themselves after writing an pro-gay essay ($M = 5.48$, $SD = 0.79$, 95% CI [5.30, 5.66]), than after writing an anti-gay essay ($M = 4.97$, $SD = 1.07$, 95% CI [4.79, 5.15]), which supported Hypothesis 1.

2

3

Refuser likability. Results only showed a significant effect of the Refuser manipulation, ($F[1, 203] = 212.45$, $p < .001$, $\eta_p^2 = .51$), indicating that participants had higher likability ratings of the refuser of the anti-gay essay ($M = 4.96$, $SD = 0.98$, 95% CI [4.76, 5.16]), than of the refuser of the pro-gay essay ($M = 2.90$, $SD = 1.04$, 95% CI [2.70, 3.10]), which supported Hypothesis 2b.

4

5

Refuser agency. Results only showed a significant effect of the Refuser manipulation, $F(1, 203) = 40.50$, $p < .001$, $\eta_p^2 = .17$, indicating that participants ascribed higher agency ratings to the refuser of the anti-gay essay ($M = 5.65$, $SD = 1.06$, 95% CI [5.43, 5.87]), than to the refuser of the pro-gay essay ($M = 4.64$, $SD = 1.21$, 95% CI [4.42, 4.86]), which supported Hypothesis 2b. We did not observe other significant effects (all $ps > .16$).

6

&

Additional analyses. We based our predictions on the assumption that participants have pro-gay attitudes. This may imply that the observed effects are stronger for those who have stronger pro-gay attitudes. To explore the possible role of attitudes towards homosexuality as a moderator, we performed regression analyses. Following West and colleagues (West, Aiken, & Krull, 1996), the Own Essay and Refuser manipulations were effect coded and attitude was mean-centered. The three main effects were included in Step 1. All two-way interactions were added in Step 2. The three-way interaction was added in Step 3. Table 5.3 shows all effects obtained in these analyses.

Self-evaluations. Results showed a significant model in Block 1, $F(3, 203) = 5.17$, $p = .002$, $\Delta R^2 = .07$. We observed only a significant main effect of own essay, ($b = .26$, 95% CI [0.13, 0.39], $t[200] = 3.88$, $p < .001$), already described above.

Refuser likability. Results showed a significant model in Step 2, $F(6, 200) = 56.64$, $p < .001$, $\Delta R^2 = .11$. A significant effect of the refuser manipulation was observed, ($b = -1.03$, 95% CI [-1.15, -0.90], $t[200] = -16.53$, $p < .001$), as well as a significant main effect of attitude, $b = -.005$, 95% CI (-0.01, 0.0004), $t(200) = -2.12$, $p = .035$. These effects were qualified by a significant Refuser x Attitude interaction, $b = -.019$, 95% CI (-0.02, -0.01), $t(200) = -7.54$, $p < .001$. No other significant effects were observed. As described above, participants who read the response of the refuser of the anti-gay essay had more positive evaluations of the refuser than participants who read the response of the refuser of the pro-gay essay. Simple slope analyses showed that this pattern was stronger for participants with very positive

Table 5.2
Unrotated Component Loadings on Self-Evaluation Items in Experiment 5.1

Items	Self-Evaluation $\alpha = .95$
<i>To what extent do you feel ... at this moment?</i>	
good	.66
happy with yourself	.87
determined	.66
happy	.70
comfortable	.60
peaceful	.43
excited	
satisfied with yourself	.83
energetic	.59
optimistic	.73
self-confident	.81
friendly	.59
worried (reverse coded)	
awkward (reverse coded)	-.53
uncomfortable (reverse coded)	-.57
fatigued (reverse coded)	
dissatisfied with yourself (reverse coded)	-.82
irritated by yourself (reverse coded)	-.78
angry with yourself (reverse coded)	-.78
disappointed in yourself (reverse coded)	-.79
self-critical (reverse coded)	-.65
guilty (reverse coded)	-.76
disgusted by yourself (reverse coded)	-.72

Note. Component loadings > |.55| are printed in bold. Component loadings < |.30| are not shown (Field, 2009).

attitudes towards homosexuality, $b = -1.50$, $t(200) = -17.04$, $p < .001$, than for participants with relatively neutral attitudes towards homosexuality, $b = -0.55$, $t(200) = -6.24$, $p < .001$.

Refuser agency. Results showed a significant model in Step 2, $F(6, 200) = 10.95$, $p < .001$, $\Delta R^2 = .08$. We observed a significant main effect of the refuser manipulation, $b = -0.50$, 95% CI (-0.65, -0.35), $t(200) = -6.58$, $p < .001$. This effect was qualified by a significant Refuser x Attitude interaction, $b = -.01$, 95% CI (-0.02, -0.01), $t(200) = -4.22$, $p < .001$. No other

1

2

3

4

5

6

&

significant effects were observed. Simple slopes analyses demonstrated that participants with very positive attitudes towards homosexuality ascribed higher agency ratings to the refuser when they read the response of the anti-essay refuser, rather than when they read the response of the pro-essay refuser, $b = -0.83$, $t(200) = -7.48$, $p < .001$. Participants with relatively neutral attitudes towards homosexuality ascribed similar agency ratings to the refuser regardless of whether they read the response of the anti-essay refuser or the pro-essay refuser, $b = -0.18$, $t(200) = -1.60$, $p = .11$.

Discussion

In line with our first hypothesis, we found that people's evaluation of themselves depend mostly on their previous behavior: Participants had more positive self-concepts after writing an explicit pro-gay than after writing an anti-gay, essay. This is in line with the theory of moral self-regulation, which posits that people have an internal feedback system that regulates their moral behavior (e.g., Zhong et al., 2009). In this system, internal alarm bells go off whenever people engage in wrong behavior, which leads them to feel bad about themselves and motivated to do better in the future.

The consequences for one's self-concept may not, however, solely rely on an internal feedback system. For example, recent research by Greene and Low (2014) demonstrates that although people were more likely to engage in a moral transgression after remembering a moral deed (i.e., moral licensing) rather than an immoral deed, this only occurred when the behavior was private and not when it was public. This fits with literature on public self-awareness which demonstrates that situations where one's "attention is directed toward the implications of one's behavior for other's evaluations of the self" (Hirt, McCrea, & Kimble, 2000, p. 1133) increase awareness of one's own behavior and self-image. Thus, the salience of one's (counter-attitudinal) behavior to others also influences the consequences thereof for one's self-concept. Therefore, in our second experiment, we investigate whether people's self-concepts suffer the most when people feel their behavior is visible to others and it is difficult for participants to view their own counter-attitudinal behavior as moral.

In support of Hypothesis 2b, we demonstrated that people had more positive evaluations of the refuser of the anti-gay essay (vs. the refuser of the pro-gay essay). Additionally, we observed that this pattern was stronger when participants had more positive attitudes towards homosexuality. This means that participants appreciated the moral character of the person who refused to write an anti-gay essay.

Our reasoning hinges on the positive evaluation of refusers of the anti-gay essay. Note, however, that our effect may have been partially driven by negative evaluations of the refuser of the pro-gay essay. The pro-gay essay refuser clearly deviated from important norms in the Dutch society where homosexuality is seen as acceptable. Given that deviating from societal norms is a strong source of dislike and antipathy (Marques, & Yzerbyt, 1988; Skitka, 2010), participants may have disliked the refuser of the pro-gay essay because he/she supported a deviant position in society. To rule out this possibility, we include a different control condition in Experiment 5.2.

Experiment 5.2

In our second experiment, we investigate whether people's moral feedback system depends on whether or not people feel that their counter-attitudinal behavior is visible to others. There are several areas of research that indicate that the visibility of one's behavior to others is important, especially when engaging in wrongful or counter-attitudinal behavior. People have an strong motive to present themselves favorably to others (for a review, see e.g., Geen, 1991) and therefore try to engage in praiseworthy behavior and refrain from undesirable behavior when their behavior is visible to others. For example, people only engage in moral transgressions in private settings and not in public settings (Greene & Lowe, 2014); cues of being watched lead to more pro-social behavior in real-life settings (Bateson, Nettle, & Roberts, 2006) and people are especially likely to buy "green" products in public rather than private settings (i.e., *conspicuous consumption*, Griskevicius, Tybur, & Van den Bergh, 2010). In Experiment 5.1, people's own immoral behavior was salient to others because participants were informed that their essay could be read by the other participant. This increased public self-awareness may have made it more difficult for them to maintain a positive self-concept (e.g., Mazar et al., 2008), because of increased concern on how their behavior would be evaluated by others (e.g., Hirt et al., 2000). We therefore argue that engaging in counter-attitudinal behavior within a context that is highly evaluative leads to negative self-evaluations, while these negative self-evaluations may be attenuated when the evaluative context is less salient. To test this prediction, we manipulated whether people were informed that their anti-gay essays would be read by the experimenter (the strong evaluative context condition) or not (the mild evaluative context condition).

Furthermore, we manipulated how hard it was for participants to see their counter-attitudinal behavior as moral. Building on the work from Mazar and colleagues (2008), we expected that people's self-concepts would be most negatively impacted when it would

1

2

3

4

5

6

&

Table 5.3
Summary of Regression Analyses in Experiment 5.1: Predicting Self-Evaluations, Likability of Refuser, and Agency of Refuser from Own Essay, Refuser, Attitude, and the Interactions Between These Predictor Variables

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
<i>Self-Evaluations</i>				
Step 1				
Own Essay	.256	.066	.264	< .001
Refuser	-.035	.066	-.036	.596
Attitude	.000	.005	.005	.938
Step 2				
Own Essay	.255	.066	.263	< .001
Refuser	-.032	.066	-.033	.626
Attitude	.000	.003	-.004	.952
Own Essay x Refuser	-.107	.066	-.110	.107
Own Essay x Attitude	-.002	.003	-.043	.524
Refuser x Attitude	.002	.003	.053	.439
Step 3				
Own Essay	.254	.066	.262	< .001
Refuser	-.035	.066	-.036	.594
Attitude	.000	.003	-.002	.982
Own Essay x Refuser	-.107	.066	-.110	.108
Own Essay x Attitude	-.002	.003	-.045	.511
Refuser x Attitude	.002	.003	.052	.447
Own Essay x Refuser x Attitude	.001	.003	.027	.694
<i>Likability of Refuser</i>				
Step 1				
Own Essay	-.001	.070	-.001	.985
Refuser	-1.027	.070	-.713	< .001
Attitude	-.006	.003	-.110	.025
Step 2				
Own Essay	.010	.063	.007	.879
Refuser	-1.026	.062	-.712	< .001
Attitude	-.005	.003	-.092	.035
Own Essay x Refuser	.076	.063	.052	.229
Own Essay x Attitude	-.003	.003	-.058	.178
Refuser x Attitude	-.019	.003	-.329	< .001

Table 5.3

Summary of Regression Analyses in Experiment 5.1: Predicting Self-Evaluations, Likability of Refuser, and Agency of Refuser from Own Essay, Refuser, Attitude, and the Interactions Between These Predictor Variables (continued)

	Variable	b	SE b	β	p-value
<i>Step 3</i>					
	Own Essay	.011	.063	.008	.859
	Refuser	-1.020	.063	-.708	< .001
	Attitude	-.006	.003	-.095	.030
	Own Essay x Refuser	.075	.063	.052	.230
	Own Essay x Attitude	-.003	.003	-.056	.195
	Refuser x Attitude	-.019	.003	-.328	< .001
	Own Essay x Refuser x Attitude	-.002	.003	-.035	.419
<i>Agency of Refuser</i>					
<i>Step 1</i>					
	Own Essay	-.114	.080	-.092	.155
	Refuser	-.506	.080	-.406	< .001
	Attitude	.001	.003	.015	.815
<i>Step 2</i>					
	Own Essay	-.107	.077	-.086	.167
	Refuser	-.503	.076	-.404	< .001
	Attitude	.002	.003	.033	.599
	Own Essay x Refuser	-.065	.077	-.053	.396
	Own Essay x Attitude	-.003	.003	-.053	.387
	Refuser x Attitude	-.013	.003	-.262	< .001
<i>Step 3</i>					
	Own Essay	-.105	.077	-.084	.176
	Refuser	-.494	.077	-.397	< .001
	Attitude	.001	.003	.027	.661
	Own Essay x Refuser	-.066	.077	-.053	.395
	Own Essay x Attitude	-.003	.003	-.050	.418
	Refuser x Attitude	-.013	.003	-.261	< .001
	Own Essay x Refuser x Attitude	-.003	.003	-.057	.358

Note. The Own Essay and Refuser conditions were effect coded; the anti-gay essay and anti-essay refuser conditions were coded -1, the pro-gay essay and pro-essay refuser conditions were coded +1.

1

be relatively difficult to view one's own behavior as moral, while their self-concepts would remain relatively unaffected when it was easier to view one's own behavior as moral.

2

Participants therefore read the bogus reaction of another participant who refused to write the anti-gay essay out of moral concern or out of non-moral concern (i.e., because he/she

3

had an injury and therefore could not type an essay). It should be harder for participants to view their own behavior as moral when someone else claimed that writing an anti-gay

4

essay was immoral. We expected people's self-evaluations to be lower when they were confronted with a refuser who makes a moral claim (vs. a non-moral claim), but only when

5

they were in a strong evaluative context (Hypothesis 3).

6

Moreover, including the non-moral condition enables us to rule out the alternative explanation from Experiment 5.1, that the positive evaluation of the refuser of the

&

anti-gay essay was due to the deviant societal position of the refuser. More specifically, in Experiment 5.2, the reactions are equally normative: In both cases the other participant

refuses to engage in counter-normative behavior (i.e., writing the anti-gay essay), albeit for different reasons.

Method

Participants and Design. Utrecht University students ($N=204$) participated in this experiment in exchange for a monetary reward or course credits ($M_{\text{age}} = 21.12$ years; $SD_{\text{age}} = 2.76$ years; 120 women). Participants were randomly assigned to one of the four conditions of our 2 (Refusal Reason: moral reason vs. non-moral reason) x 2 (Evaluative Context: strong vs. mild) experimental design. The number of participants per cell varied between 47 and 54.³

Procedure. The procedure of this experiment was similar to Experiment 5.1; we therefore only elaborate on the differences. All participants were asked to write an essay from the viewpoint of a passionate adversary of gay rights (i.e., the anti-gay essay condition of Experiment 5.1). As expected, agreement with the assigned anti-gay essay was low ($M = 11.80$, $SD = 21.71$), which indicated that the assignment was counter-attitudinal. Agreement did not differ significantly between the different conditions of our experiment, $F(3, 200) = 0.58$, $p = .628$, $\eta_p^2 = .01$. Participants' attitudes towards homosexuality were

³ Twenty-one additional participants completed this experiment but were not included in the analyses. Five participants participated more than once in Experiment 5.2, and we therefore only included their data from the first participation. Furthermore, as in Experiment 5.1, we excluded participants who refused to write the essay (15 participants) and/or wrote an essay of one sentence or less (one participant). This resulted in a final N of 204.

on average positive ($M = 68.86$, $SD = 24.95$), and did not differ significantly between the different conditions of our experiment, $F(3, 200) = 0.72$, $p = .544$, $\eta_p^2 = .01$.

Participants were asked to write a good essay. Following research by Van den Bos, Bruins, Wilke, and Dronkert (1999), half of the participants were informed that their essay would be evaluated by the experimenter and this message was repeated three times throughout the experiment (*strong evaluative context condition*). The other half of the participants was not given the information that the experimenter would evaluate their essay (*mild evaluative context condition*; Van den Bos et al., 1999).

Afterwards, participants were confronted with the reaction of someone who had refused to write the anti-gay essay. In the *non-moral reason* condition, the other participant refused to write the pro-gay essay because he/she had injured his/her hand and could therefore not type an essay. The *moral reason* condition was exactly the same as the anti-essay refuser condition in Experiment 5.1.

Refuser likability (29 items, $\alpha = .98$), *Refuser agency* (9 items, $\alpha = .93$) and *Self-evaluations* (18 items, $\alpha = .95$) were measured similarly as in Experiment 5.1. We measured participants' fear of negative evaluation with a state version of the brief Fear of Negative Evaluation scale (12 items, $\alpha = .91$; Leary, 1983), consisting of items such as "Right now, I am afraid others will not approve of me". Participants' attitudes towards lesbians, gay men, and bisexuals (LGBs) were measured with an adapted version of the attitudes towards homosexuals scale (21 items, $\alpha = .91$; Kite & Deaux, 1986), which included items such as "I would not mind having a lesbian/gay/bisexual friend". We also assessed participants' sexual orientation.

Results

Check on attitudes. There was a significant and strong correlation between the one-item thermometer question that assessed attitudes towards homosexuality, and the 21-item extended questionnaire that assessed attitudes towards LGBs ($r = .67$, $p < .001$), thereby validating our use of the one-item thermometer question as an assessment of participants' attitudes towards homosexuality.

Check on evaluative context. An ANOVA indicated that the evaluative context manipulation had some influence on participants' measured state FNE (Leary, 1983). Participants in the strong evaluative context condition scored somewhat higher on this scale ($M = 3.80$, $SD = 1.16$) than participants in the mild evaluative context condition ($M = 3.53$, $SD = 1.16$), although this effect was marginally significant only, $F(1, 202) = 2.76$, $p = .099$, $\eta^2 = .01$.

1

2

3

4

5

6

&

1

Main analyses. We performed ANOVAs with the refusal reason and evaluative context manipulations as the independent variables.

2

3

4

5

6

&

Self-evaluations. Results showed a significant main effect of evaluative context, $F(1, 200) = 6.02, p = .015, \eta_p^2 = .03$, indicating that after writing an anti-gay essay, participants had more positive evaluations of themselves in a mild evaluative context ($M = 5.30, SD = 0.95, 95\% \text{ CI } [5.10, 5.49]$), rather than a strong evaluative context ($M = 4.95, SD = 1.14, 95\% \text{ CI } [4.73, 5.14]$). Furthermore, we observed a significant Refusal Reason \times Evaluative Context interaction, $F(1, 200) = 4.84, p = .029, \eta_p^2 = .02$. Simple main effects demonstrated that participants in the non-moral reason conditions had similar self-evaluations, regardless of whether they were in a mild evaluative context ($M = 5.18, SD = 0.98$) or strong evaluative context ($M = 5.14, SD = 0.93$), $F(1, 200) = 0.03, p = .857, \eta_p^2 < .001$; see Figure 5.1. Importantly, and in support of Hypothesis 3, participants in the moral reason condition had more negative self-evaluations in the strong evaluative context ($M = 4.74, SD = 1.31$) than in the mild evaluative context ($M = 5.41, SD = 0.90$), $F(1, 200) = 10.70, p = .001, \eta_p^2 = .05$.

Refuser likability. Results only showed a significant effect of the refusal reason manipulation, $F(1, 200) = 170.34, p < .001, \eta_p^2 = .46$, indicating that participants in the moral reason condition evaluated the refuser as more likable ($M = 5.20, SD = 1.00, 95\% \text{ CI } [5.01, 5.38]$), than participants in the non-moral reason condition ($M = 3.48, SD = 0.86, 95\% \text{ CI } [3.30, 3.67]$). This supported Hypothesis 2b.

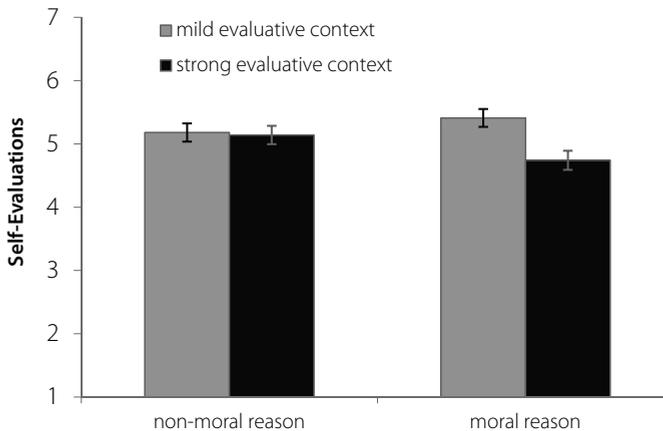


Figure 5.1. Self-Evaluations in Experiment 5.2, as a Function of Refusal Reason and Evaluative Context Manipulations. Higher numbers represent more positive self-evaluations. Error bars are plotted one standard error above and below the means.

Refuser agency. Results only showed a significant effect of the refusal reason manipulation, $F(1, 200) = 67.14, p < .001, \eta_p^2 = .25$, indicating that participants in the moral reason condition ascribed higher agency ratings to the refuser ($M = 5.68, SD = 1.20, 95\% \text{ CI } [5.44, 5.90]$), than did participants in the non-moral reason condition ($M = 4.32, SD = 1.15, 95\% \text{ CI } [4.09, 4.55]$). This supported Hypothesis 2b.

Additional analyses. To explore the possible role of attitudes towards homosexuality as a moderator, we performed a set of regression analyses. The refusal reason and evaluative context conditions were effect coded, attitude was mean-centered. The three main effects were included in Step 1. All two-way interactions were added in Step 2. The three-way interaction effect was added in Step 3. Table 5.4 shows all effects obtained in these analyses.

Self-evaluations. Results showed a significant model in Step 3, $F(7, 196) = 3.11, p = .004, \Delta R^2 = .01$. We observed a significant main effect of evaluative context, ($b = 0.17, 95\% \text{ CI } [0.02, 0.31], t[196] = 2.31, p = .022$), and a marginally significant main effect of attitude, $b = -0.005, 95\% \text{ CI } (-0.011, 0.001), t(196) = -1.69, p = .093$. These effects were qualified by a significant Refusal Reason x Evaluative Context interaction, ($b = -0.15, 95\% \text{ CI } [-0.29, -0.01], t[196] = -2.05, p = .042$), and a significant Evaluative Context x Attitude interaction, $b = 0.006, 95\% \text{ CI } (0.001, 0.012), t(196) = 2.17, p = .032$. Furthermore, a marginally significant Refusal Reason x Evaluative Context x Attitude interaction was observed, $b = -0.005, 95\% \text{ CI } (-0.011, -0.001), t(196) = -1.70, p = .092$. No other significant effects were observed. To interpret the marginally significant Refusal Reason x Evaluative Context x Attitude interaction, we inspected the simple slopes with GLM analyses and plotted the regression lines, the results can be seen in Figures 5.2a and 5.2b. When participants had relatively neutral attitudes towards homosexuality (-1 SD), their self-evaluations were similarly high regardless of whether they read the moral reason or non-moral reason in the strong, ($F[1,196] = 0.07, p = .788, \eta^2 < .001$), or mild evaluative context conditions, ($F[1,196] = 0.43, p = .512, \eta^2 = .002$), as can be seen in Figure 5.2a. When participants had very positive attitudes towards homosexuality ($+1 \text{ SD}$) on the other hand, participants in the strong evaluative context condition had lower self-evaluations when they were in the moral reason condition than in the non-moral reason condition, $F(1,196) = 7.47, p = .007, \eta^2 = .04$. When participants were in the mild evaluative context condition, they had similar levels of self-evaluation, regardless of whether they were in the moral reason condition, or the non-moral reason condition, ($F[1,196] = 0.99, p = .320, \eta^2 = .005$), as can be seen in Figure 5.2b.

1

2

3

4

5

6

&

Table 5.4
Summary of Regression Analyses in Experiment 5.2: Predicting Self-Evaluation, Likability of Refuser, and Agency of Refuser from Evaluative Context, Refusal Reason, Attitude and the Interactions Between These Variables

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
<i>Self-Evaluations</i>				
Step 1				
Evaluative Context	.172	.073	.163	.020
Refusal Reason	.027	.074	.026	.713
Attitude	-.004	.003	-.085	.223
Step 2				
Evaluative Context	.176	.072	.167	.015
Refusal Reason	.038	.072	.036	.596
Attitude	-.004	.003	-.105	.130
Evaluative Context x Refusal Reason	-.147	.072	-.140	.043
Evaluative Context x Attitude	.007	.003	.154	.028
Refusal Reason x Attitude	.003	.003	.078	.256
Step 3				
Evaluative Context	.166	.072	.157	.022
Refusal Reason	.031	.072	.030	.665
Attitude	-.005	.003	-.117	.093
Evaluative Context x Refusal Reason	-.147	.072	-.140	.042
Evaluative Context x Attitude	.006	.003	.149	.032
Refusal Reason x Attitude	.004	.003	.091	.189
Evaluative Context x Refusal Reason x Attitude	-.005	.003	-.117	.092
<i>Likability of Refuser</i>				
Step 1				
Evaluative Context	-.024	.065	-.019	.710
Refusal Reason	-.844	.065	-.668	<.001
Attitude	.006	.003	.128	.014
Step 2				
Evaluative Context	-.032	.063	-.025	.613
Refusal Reason	-.843	.063	-.668	<.001
Attitude	.006	.003	.120	.019
Evaluative Context x Refusal Reason	-.092	.064	-.073	.148
Evaluative Context x Attitude	-.002	.003	-.041	.417
Refusal Reason x Attitude	-.008	.003	-.166	.001

Table 5.4

Summary of Regression Analyses in Experiment 5.2: Predicting Self-Evaluation, Likability of Refuser, and Agency of Refuser from Evaluative Context, Refusal Reason, Attitude and the Interactions Between These Variables (continued)

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
Step 3				
Evaluative Context	-.035	.064	-.028	.585
Refusal Reason	-.845	.064	-.669	< .001
Attitude	.006	.003	.118	.022
Evaluative Context x Refusal Reason	-.092	.064	-.073	.148
Evaluative Context x Attitude	-.002	.003	-.042	.408
Refusal Reason x Attitude	-.008	.003	-.163	.002
Evaluative Context x Refusal Reason x Attitude	-.001	.003	-.026	.611
<i>Agency of Refuser</i>				
Step 1				
Evaluative Context	-.068	.082	-.050	.408
Refusal Reason	-.662	.082	-.489	< .001
Attitude	.008	.003	.155	.011
Step 2				
Evaluative Context	-.074	.081	-.054	.363
Refusal Reason	-.660	.081	-.487	< .001
Attitude	.008	.003	.148	.015
Evaluative Context x Refusal Reason	-.142	.081	-.105	.080
Evaluative Context x Attitude	-.002	.003	-.043	.482
Refusal Reason x Attitude	-.007	.003	-.131	.031
Step 3				
Evaluative Context	-.071	.081	-.052	.386
Refusal Reason	-.658	.081	-.486	< .001
Attitude	.008	.003	.150	.014
Evaluative Context x Refusal Reason	-.142	.081	-.105	.081
Evaluative Context x Attitude	-.002	.003	-.042	.494
Refusal Reason x Attitude	-.007	.003	-.133	.029
Evaluative Context x Refusal Reason x Attitude	.001	.003	.027	.658

Note. The evaluative context and refusal reason conditions were effect coded with the strong evaluative context and moral reason conditions coded as -1 and the mild evaluative context and non-moral reason conditions coded as +1.

- 1
- 2
- 3
- 4
- 5
- 6
- &



Figure 5.2a. Self-Evaluations for Participants With a Moderate Attitudes Towards Homosexuality in Experiment 5.2, as a Function of Refusal Reason and Evaluative Context Manipulations. Higher numbers indicate more positive self-evaluations. Slopes are plotted one SD below the mean of attitude. Error bars are plotted one standard error above and below the mean.

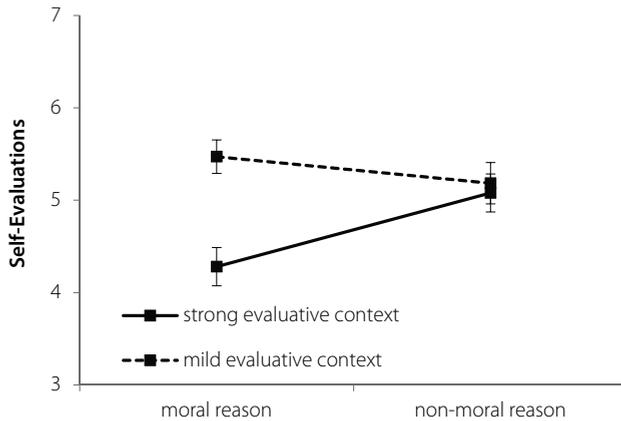


Figure 5.2b. Self-Evaluations for Participants With Very Positive Attitudes Towards Homosexuality in Experiment 5.2, as a Function of Refusal Reason and Evaluative Context Manipulations. Higher numbers indicate more positive self-evaluations. Slopes are plotted one SD above the mean of attitude. Error bars are plotted one standard error above and below the mean.

Refuser likability. Results showed a significant model in Step 2, $F(6, 197) = 33.87, p < .001$, $\Delta R^2 = .03$. A significant main effect of refusal reason was observed, ($b = -0.84$, 95% CI [-0.97, -0.72], $t[197] = -13.29, p < .001$), as well as a significant main effect of attitude, $b = .006$, 95% CI (0.001, 0.011), $t(197) = 2.37, p = .019$. These effects were qualified by a significant Refusal Reason x Attitude interaction, $b = -0.008$, 95% CI (-.013, -0.003), $t(197) = -3.28, p = .001$. No other significant effects were observed.

As described above, participants in the moral reason condition had more favorable evaluations of the refuser than participants in the non-moral reason condition. Simple slope analyses showed that this pattern was stronger for participants who had very positive attitudes towards homosexuality, ($b = -1.04$, $t[197] = -11.24, p < .001$), than for participants who had relatively neutral attitudes towards homosexuality, $b = -0.64$, $t(197) = -7.15, p < .001$.

Refuser agency. Results showed a significant model in Step 2, $F(6, 197) = 14.32, p < .001$, $\Delta R^2 = .03$. A significant main effect of refusal reason was observed, ($b = -0.66$, 95% CI [-0.82, -0.50], $t[197] = -8.16, p < .001$), as well as a significant main effect of attitude, $b = .008$, 95% CI (0.002, 0.015), $t(197) = 2.45, p = .015$. These effects were qualified by a significant Refusal Reason x Attitude interaction, $b = -0.007$, 95% CI (-0.014, -0.001), $t(197) = -2.18, p = .031$. We observed a marginally significant Refusal Reason x Evaluative Context interaction, $b = -0.14$, 95% CI (-0.30, 0.02), $t(197) = -1.76, p = .080$. However, inspection of the simple slopes indicated no statistically significant slopes so we do not interpret this interaction further. No other significant effects were observed.

As described above, participants ascribed more agency to the refuser in the moral reason condition than in the non-moral reason condition. Simple slope analyses showed that this pattern was stronger for participants with very positive attitudes towards homosexuality, ($b = -0.84$, $t[197] = -7.11, p < .001$), than for participants with relatively neutral attitudes towards homosexuality, $b = -0.49$, $t(197) = -4.26, p < .001$.

Discussion

Our findings extend self-concept maintenance theory (Mazar et al., 2008). This theory states that people are less likely to engage in unethical behavior when it is hard to categorize their behavior as moral and people are self-aware, because this makes it harder to maintain a positive self-concept. We demonstrate that this also works the other way around: After engaging in counter-attitudinal behavior, people's self-concepts were only lowered when they believed their essays would be read by the experiment leader and they read a moral reason of someone else to engage in the same task. Arguably, the knowledge that their

1

2

3

4

5

6

&

1

essays would be read by another person made them more self-conscious and thereby increased awareness of their own behavior and moral standards. Furthermore, the explicit claim by the other person that writing an anti-gay essay is wrong made it difficult to see their own behavior as moral. Combined, these findings show that engaging in counter-attitudinal behavior that does not fit with one's moral values has negative consequences for one's self-concept, but only under specific conditions.

2

3

4

5

6

&

Furthermore, we again observed that people had positive evaluations of others who refused to write an anti-gay essay out of moral concern—now compared to someone who refused to write an anti-gay essay out of non-moral concern. Apparently, this positive evaluation is not only observed when contrasted with an extremely deviant position in society (as in Experiment 5.1) but also when contrasted with a non-deviant position in society. This is important because it strengthens our belief that we investigate a situation where people appreciate the character of those who—out of moral reasons—refuse to engage in counter-attitudinal behavior. In two experiments, we now observed support that people have positive evaluations of others who refuse to write an anti-gay essay out of moral reasons.

Our findings may be relevant for research on “do-gooder derogation”; the phenomenon that people derogate others who refuse to engage in immoral behavior out of moral reasons (e.g., Chapter 2; Minson & Monin, 2012; Monin, 2007). For example, Monin and colleagues (2008) showed that people who picked a black man as the most likely subject of a crime, derogated a bogus participant who ostensibly had refused to partake in this task because it was racist. However, because of differences in the experimental set-up, it is difficult to directly relate the results of our first two experiments to these other lines of research. Therefore, in our third experiment, we examine the possible linkage between our work and earlier studies.

Experiment 5.3

In this experiment, we examine the possible relevance of our current work for earlier studies, most notably the work by Monin and colleagues (2008). More specifically, we investigate the critical role of involvement by manipulating whether participants themselves engage in the wrongful behavior (the actor conditions) or whether they merely observe the reaction of the moral refuser (the observer conditions; similar to Monin et al., 2008).

According to self-concept maintenance theory (Mazar et al., 2008), people's self-concepts should remain unharmed when it is easy to view their own behavior as moral. We predict

that it is easier for people to see their own behavior as moral when they have not engaged in wrongful behavior, rather than when they have engaged in wrongful behavior. Thus, we expect that participants' self-evaluations are lower when they are actors rather than observers (Hypothesis 4). Another aspect that may make it easier for participants to maintain a positive self-concept is when moral standards are less salient (Mazar et al., 2008). Therefore, we manipulate the task the other participant ostensibly received. We confront participants with the reaction of someone else who refused to write an anti-gay essay (the same topic conditions; similar to Experiments 5.1 and 5.2) or with the reaction of someone else who refused an anti-abortion essay (the different topic conditions). We expect that reading a refusal to engage in the same task would make moral standards more salient than refusal to engage in a different task. Therefore, we expect that self-evaluations would be lower for participants in the same task conditions. However, this should only be the case for participants who engaged in the counter-attitudinal behavior, because observers did not receive a task themselves (Hypothesis 5).

1

2

3

4

5

6

&

Method

Participants and design. Utrecht University students ($N = 201$) participated in this experiment, in exchange for a monetary reward or course credits ($M_{\text{age}} = 22.17$ years; $SD_{\text{age}} = 3.90$ years; 143 women). Participants were randomly assigned to one of the four conditions of our 2 (Role: observer vs. actor) \times (Task: same topic vs different topic) experimental design. Cell sizes varied between 47 and 54 participants.⁴

Procedure. The procedure of this experiment was similar to Experiments 5.1 and 5.2, we therefore only elaborate on the differences. Participants were either asked to write an essay from the viewpoint of a passionate adversary of gay rights (i.e., the actor condition, which was exactly the same as in Experiments 5.1 and 5.2), or participants were not asked to write an essay (i.e., the observer condition; similar to the observer condition in Monin et al., 2008).

As expected, participants' attitudes towards homosexuality were on average positive ($M = 75.31$, $SD = 25.33$), and did not differ significantly between the different conditions of our experiment, $F(3, 197) = 0.13$, $p = .941$, $\eta_p^2 = .002$.

⁴ Eight additional participants completed this experiment but were not included in the analyses. Three participants already participated in Experiment 5.1 or 5.2. Furthermore, as in Experiments 5.1 and 5.2, we excluded participants who refused to write the essay (five participants) and/or wrote an essay of one sentence or less (none of the participants). This resulted in a final N of 201.

1

Furthermore, general agreement with the anti-gay essay was low ($M = 18.24$, $SD = 31.06$), which indicated that this task was counter-attitudinal. However, agreement differed significantly between the different conditions of our experiment.⁵

2

3

4

5

6

&

After writing their essay (i.e., actor conditions) or not (i.e., the observer conditions), participants were confronted with the reaction of someone who had refused to write an anti-gay essay or an anti-abortion essay, out of moral concern. The *same task* condition was exactly the same as in Experiments 5.1 and 5.2. The *different task* condition showed the response of another person who ostensibly indicated to refuse to write an anti-abortion essay because “*abortion is normal and moral, and women should have the right to have an abortion*”.

Refuser likability (29 items, $\alpha = .97$), *Refuser agency* (9 items, $\alpha = .92$) and *Self-evaluations* (18 items, $\alpha = .95$) were assessed similarly as in Experiments 5.1 and 5.2.

Results

Main analyses. We performed ANOVAs with the role and task manipulations as the independent variables.

Self-evaluations. In partial support of Hypothesis 4, results only showed a marginally significant effect of role, indicating that participants had lower self-evaluations when they were actors ($M = 5.15$, $SD = 1.06$) rather than observers ($M = 5.44$, $SD = 0.96$), $F(1, 197) = 3.87$, $p = .051$, $\eta_p^2 = .02$. We found no support for our hypothesis that participants would have lower self-evaluations when they read the response of a same task refuser rather than a different task refuser, but only when they were actors themselves (Hypothesis 5).

Refuser likability. Results only showed a significant effect of task, indicating that participants ascribed higher likability to the refuser of the anti-gay essay ($M = 5.18$, $SD = 0.99$) rather

⁵ We observed significant main effects of Role and Task, which were qualified by a significant Role x Task interaction, $F(1, 197) = 6.64$, $p = .011$, $\eta_p^2 = .033$. Simple main effects illustrated that observers agreed less with the anti-gay statement when they were confronted with the same task refuser ($M = 14.89$, $SD = 30.41$) rather than the different task refuser ($M = 36.66$, $SD = 36.65$), $F(1, 197) = 14.28$, $p < .001$, $\eta_p^2 = .068$. Actors' agreement with the anti-gay statement was equally low regardless of whether they were confronted with the refuser of the same task ($M = 10.42$, $SD = 23.62$) or with the refuser of the different task ($M = 10.83$, $SD = 24.48$), $F(1, 197) = 0.01$, $p = .945$, $\eta_p^2 < .001$.

than the refuser of the anti-abortion essay ($M = 4.63$, $SD = 1.06$), $F(1, 197) = 13.86$, $p < .001$, $\eta_p^2 = .07$. This finding supported Hypothesis 2b, that participants have positive evaluations of others who refuse to write an anti-gay essay out of moral concern.

Refuser agency. Results showed no significant main or interaction effects ($M = 4.99$, $SD = 1.36$; all $ps > .11$).

Additional analyses. To explore the possible role of attitudes towards homosexuality as a moderator, we performed a set of regression analyses. The manipulations were effect coded, attitude was mean-centered. The three main effects were included in Step 1. All two-way interaction effects were added in Step 2. The three-way interaction effect was added in Step 3. Table 5.5 shows all effects obtained in these analyses.

Self-evaluations. Results showed a non-significant model in Step 2, $F(6, 194) = 1.52$, $p = .172$, $\Delta R^2 = .02$. However, a significant main effect of role was observed, ($b = -0.14$, 95% CI [-0.29, -0.003], $t[194] = -2.02$, $p = .045$), indicating that people had lower self-evaluations when they actors rather than observers. Furthermore, a significant Role x Attitude interaction was observed, $b = -0.006$, 95% CI (-0.011, -0.0002), $t(194) = -2.04$, $p = .043$. Simple slopes analyses demonstrated that participants with moderate attitudes towards homosexuality had similar self-evaluations, regardless of their role, $b = 0.008$, $t(194) = 0.08$, $p = .937$. Participants with very positive attitudes towards homosexuality had lower self-evaluations when they were actors rather than observers, $b = -0.296$, $t(194) = -2.94$, $p = .004$.

Refuser likability. Results showed a significant model in Step 3, $F(7, 193) = 5.62$, $p < .001$, $\Delta R^2 = .04$. A significant main effect of task was observed, ($b = -0.26$, 95% CI [-0.39, -0.12], $t[193] = -3.71$, $p < .001$), as well as a significant main effect of attitude, $b = .011$, 95% CI (0.005, 0.016), $t(193) = 3.78$, $p < .001$. These effects were qualified by a significant Task x Role x Attitude interaction, $b = 0.008$, 95% CI (0.003, 0.014), $t(193) = 2.85$, $p = .005$. No other significant effects were observed.

Figures 5.3a and 5.3b illustrate the Task x Role x Attitude interaction effect. As can be seen in Figure 5.3a, for *actors*, likability ratings of the refuser depended on their attitudes towards homosexuality ($t_{\text{slope difference}} = 2.87$, $p = .005$). Simple slope analyses demonstrated that participants with moderate attitudes towards homosexuality had higher likability ratings of refusers of the same task than of refusers of the different task, $b = -0.56$, $t(193) = -3.90$, $p < .001$, whereas participants with very positive attitudes towards homosexuality had similar likability ratings of refusers of the same and different task, $b = 0.049$, $t(193) = 0.327$, $p = .744$.

1

2

3

4

5

6

&

Table 5.5
Summary of Regression Analyses in Experiment 5.3: Predicting Self-Evaluation, Refuser Likability, and Refuser Agency from Role, Task, Attitude and the Interactions Between These Predictor Variables

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
<i>Self-Evaluations</i>				
Step 1				
Role	-.142	.072	-.140	.049
Task	.048	.072	.048	.499
Attitude	-.001	.003	-.019	.789
Step 2				
Role	-.144	.071	-.142	.045
Task	.044	.071	.043	.538
Attitude	-.002	.003	-.044	.548
Role x Task	.011	.071	.011	.880
Role x Attitude	-.006	.003	-.144	.043
Task x Attitude	-.002	.003	-.044	.544
Step 3				
Role	-.143	.071	-.140	.047
Task	.048	.071	.047	.505
Attitude	-.002	.003	-.045	.535
Role x Task	.011	.071	.011	.874
Role x Attitude	-.005	.003	-.122	.094
Task x Attitude	-.001	.003	-.037	.613
Role x Task x Attitude	.004	.003	.089	.220
<i>Refuser Likability</i>				
Step 1				
Role	-.058	.071	-.054	.416
Task	-.268	.070	-.253	<.001
Attitude	.010	.003	.234	.001
Step 2				
Role	-.054	.071	-.051	.447
Task	-.265	.071	-.251	<.001
Attitude	.011	.003	.260	<.001
Role x Task	<.001	.071	.000	.996
Role x Attitude	.002	.003	.059	.382
Task x Attitude	.004	.003	.085	.223

Table 5.5
Summary of Regression Analyses in Experiment 5.3: Predicting Self-Evaluation, Refuser Likability, and Refuser Agency from Role, Task, Attitude and the Interactions Between These Predictor Variables (continued)

Variable	<i>b</i>	SE <i>b</i>	β	<i>p</i> -value
Step 3				
Role	-.051	.069	-.048	.463
Task	-.257	.069	-.244	<.001
Attitude	.011	.003	.257	<.001
Role x Task	.002	.069	.002	.981
Role x Attitude	.004	.003	.106	.121
Task x Attitude	.004	.003	.100	.142
Role x Task x Attitude	.008	.003	.194	.005
<i>Refuser Agency</i>				
Step 1				
Role	-.109	.073	-.103	.137
Task	-.012	.073	-.012	.868
Attitude	.009	.003	.216	.002
Step 2				
Role	-.107	.073	-.102	.142
Task	-.013	.073	-.013	.855
Attitude	.010	.003	.242	.001
Role x Task	-.076	.073	-.072	.299
Role x Attitude	.001	.003	.035	.611
Task x Attitude	.004	.003	.102	.155
Step 3				
Role	-.107	.073	-.102	.145
Task	-.012	.073	-.012	.868
Attitude	.010	.003	.242	.001
Role x Task	-.076	.073	-.072	.301
Role x Attitude	.002	.003	.042	.561
Task x Attitude	.004	.003	.104	.149
Role x Task x Attitude	.001	.003	.026	.713

Note. The Role and Task conditions were effect coded with the observer and same task refuser conditions coded as -1 and the actor and different task refuser conditions coded as +1.

- 1
- 2
- 3
- 4
- 5
- 6
- &

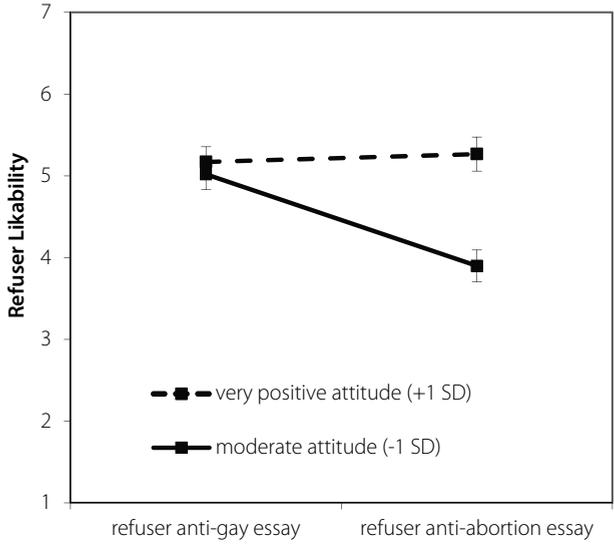


Figure 5.3a. Actors Likability Ratings in Experiment 5.3, as a Function of Refuser Task and Attitudes Towards Homosexuality. Slopes are plotted one SD above the mean of attitude. Error bars are plotted one standard error above and below the mean.

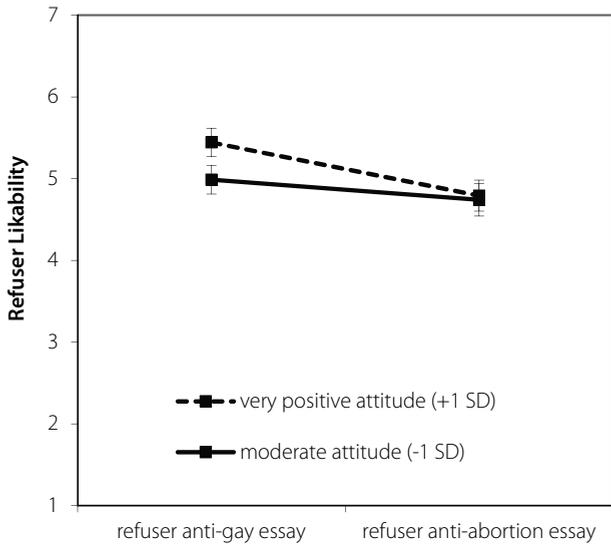


Figure 5.3b. Observers Likability Ratings in Experiment 5.3, as a Function of Refuser Task and Attitudes Towards Homosexuality. Slopes are plotted one SD above the mean of attitude. Error bars are plotted one standard error above and below the mean.

As shown in Figure 5.3b, for *observers*, likability ratings depended on their attitudes towards homosexuality, but the pattern was reversed (and the difference between the slope was not significant; $t_{\text{slope difference}} = -1.10, p = .273$). Simple slope analyses demonstrated that participants with moderate attitudes towards homosexuality had similar likability ratings of refusers of the same and different task, ($b = -0.158, t[193] = -1.121, p = .264$), whereas participants with very positive attitudes towards homosexuality had higher likability ratings of refusers of the same task than of refusers of the different task, $b = -0.360, t(193) = -2.659, p = .009$.

Refuser agency. Results showed a significant model in Step 3, $F(7, 193) = 2.29, p = .029, \Delta R^2 = .001$. Only a significant main effect of attitude was observed, $b = 0.01, 95\% \text{ CI } (0.004, 0.016), t(193) = 3.37, p = .001$; the more positive participants' attitude towards homosexuality were, the more agentic they considered the refuser to be.

Discussion

In line with Hypothesis 4, participants had more favorable self-concepts when they were only observers. This effect was stronger for people who had very strong positive attitudes about homosexuality. However, in contrast to expectations, people's self-concepts did not depend on the exact task that was refused by the other participant, and we thus observed no support for Hypothesis 5. Furthermore, we again demonstrated that participants had positive evaluations of others who refused to write this anti-gay essay out of moral concern, thereby replicating the results from Experiments 5.1 and 5.2.

One may wonder why we did not observe support for our hypothesis that people's self-concept would depend on the specific task the other person refused. We reasoned that people's self-concepts should remain unaffected when moral standards are less salient (Mazar et al., 2008) and that moral standards about anti-gay behavior would be more salient when someone else refused an anti-gay task, rather than an anti-abortion task. However, in hindsight, both tasks may have actually increased the salience of moral norms. Research demonstrates that the people's attitudes towards abortion are among the best predictors of people's support for gay-rights (Besen & Zicklin, 2007; Hicks & Lee, 2006), meaning that people's moral values concerning homosexuality are strongly intertwined with their moral values concerning abortion. This implies that increasing the salience of moral norms about abortion may have heightened the salience of moral norms about homosexuality. This could explain why self-evaluations were unaffected by the content of the task the other person refused. If true, this may have some interesting implications. For example, although people can be flexible in their attempts to maintain a positive self-concept in the face of

1

2

3

4

5

6

&

1

moral transgressions (e.g., Mazar et al., 2008), being confronted with the moral behavior of others can prevent this. This behavior of the other does not even have to be exactly in the same moral domain in order to have this effect.

2

3

4

5

6

&

At first sight, the finding that people have positive evaluations of the other participant who refused to write an anti-gay essay may seem incompatible with findings from earlier work (e.g., Chapter 2; Cramwinckel et al., 2015; Monin et al., 2008). In this previous research, people had more *negative* evaluations of moral refusers who refused to go along with 'immoral' behavior, while in the current research we find that people have more *positive* evaluations of moral refusers. As noted earlier, we investigate a situation where people engage in behavior that is blatantly counter-attitudinal and is thus likely to set off internal alarm-bells. This contrasts earlier work where people engaged in behavior that seemed innocent, such as tasting a slice of sausage (Chapter 2). Thus, the confrontation with the moral refuser in the current experiments does not seem to signal participants' immoral behavior, and therefore does not evoke defensive responses from participants. Our findings hereby expand the literature on moral refusers and do-gooder derogation (e.g., Chapter 2; Minson & Monin, 2012; Monin et al., 2008) because we demonstrate that in situations where the moral quality of the refused behavior is clear, social support for this refuser can be strong.

General Discussion

It is not uncommon for people to display counter-attitudinal behavior. For example, people go along with the majority opinion even when they know the majority is wrong (Asch, 1951), deliver shocks to others when prompted by an experiment leader (Milgram, 1963), and write anti-gay essays even though they explicitly disagree with the assignment (the current experiments). But how does engaging in counter-attitudinal behaviors such as these influence people's image of themselves? We systematically investigated the consequences of displaying counter-attitudinal behavior. Our paper highlights how people's self-concepts and their evaluations of others are influenced after displaying behavior that goes against one's moral beliefs.

In Experiment 5.1, we demonstrated that participants had more positive self-evaluations when they wrote a pro-gay, rather than an anti-gay essay. Furthermore, participants had more positive evaluations of the other person when they read a refusal to write an anti-gay essay, rather than an refusal to write a pro-gay essay. In Experiment 5.2, all participants wrote an anti-gay essay. We showed that participants self-evaluations were lower when they read a moral reason rather than a non-moral reason for refusing to write an anti-gay

essay, but only when they believed that their own anti-gay essay would be read by the experimenter. As in Experiment 5.1, people had positive evaluations of the other participant who refused to write an anti-gay essay out of moral concern. In Experiment 5.3, participants' self-evaluations were lower when they wrote an anti-gay essay, rather than were mere observers. Again, participants had positive evaluations of others who, out of moral concern, refused to write an anti-gay essay.

Our findings align with the theory of moral self-regulation (e.g., Zhong et al., 2009). More specifically, people's self-concepts were more negative after writing an anti-gay essay rather than a pro-gay essay (Experiment 5.1) or doing nothing (Experiment 5.3). Interestingly, we also present evidence that people's self-concepts are flexible and only affected under specific conditions. In an extension of self-concept maintenance theory (Mazar et al., 2008), we demonstrate that people's self-concepts are only negatively impacted when their own behavior is salient (Experiment 5.2) and when it is hard to view their behavior as moral (Experiments 5.2 and 5.3). This implies that people can often engage in counter-attitudinal behavior without negative repercussions. Apparently, failing to adhere to one's moral standards only affects people's psychological well-being in some of the cases. These insights may help explain why counter-attitudinal behavior happens so often and possibly serve as a first step in preventing counter-attitudinal behavior.

Another interesting future step may be to investigate how engaging in counter-attitudinal behavior influences people's future behavior. It may be that people would feel the need to make up for going against their moral beliefs by displaying extra moral behavior, as the theory of moral self-regulation would suggest (e.g., Zhong et al., 2009). This would imply, for instance, that although people could be easily enticed to write an anti-gay essay, they might afterwards donate more time or money to a pro-gay charity in order to restore their moral self-worth.

Somewhat related to this point is our observation that being confronted with someone who refuses to engage in counter-attitudinal behavior may sometimes teach people how to stick to their convictions. Maybe witnessing a refusal to engage in counter-attitudinal behavior can liberate people to act in line with their own moral beliefs as well. This relates to literature on *moral elevation* that demonstrates that witnessing moral behavior by others can induce moral or pro-social behavior (e.g., Aquino, McFerran, & Laven, 2011; Schnall & Roper, 2012; Schnall, Roper, & Fessler, 2010). Systematically investigating under which conditions people learn from the exemplary behavior of the refuser therefore seems an interesting direction for future research.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

Our findings may be linked to literature on retaliation against whistleblowers. Although it would be in the interest of companies and society at large when people would blow the whistle on organizational misconduct, this often does not happen. And when it does, those who dare to blow the whistle are often targets of negative reactions and retaliation (e.g., Mesmer-Magnus & Viswesvaran, 2005; Yeargain & Kessler, 2009). Our current work can offer important insights for this research area. We show that people can have positive evaluations of others who refuse to engage in behavior they think is wrong, even when most others agreed to go along. This may be relevant because it demonstrates that people not necessarily always feel the need to retaliate against those call out wrongful behavior.

To conclude, our experiments indicate that behaving in accordance with one's moral beliefs remains a challenge. Most people can be easily persuaded to engage in counter-attitudinal behavior, in some cases even without any repercussions to their personal self-concepts. However, our results also paint a more optimistic picture. In three experiments, we show that people have positive views of others who refuse to compromise their moral values and uphold these values in the face of difficulty. This suggests that even though people themselves may sometimes fail to behave in accordance with their moral beliefs, they will not fail to appreciate defenders of the moral cause they believe in.





6

General Discussion

Author contributions: Solo-authored with helpful comments from Kees van den Bos and Eric van Dijk.

This chapter is partly based on Cramwinckel, F. M., Van den Bos, K., & Van Dijk, E. (2015). Reactions to morally motivated deviance. *Current Opinion in Psychology*, 6, 150-156. doi: 10.1016/j.copsy.2015.08.007

You have been reading four empirical chapters about how people react to social situations involving mundane morality. What do we learn from this research, how can this knowledge be used, and where should we go from here? In this last chapter, I will reflect on these issues. In order to do so, let's start with an overview of the different findings in the different empirical chapters.

In Chapter 2, I have investigated a situation where people first tasted a sausage themselves and were afterwards confronted with the reaction of a peer who had refused to do the same. This other person had either moral or non-moral reasons to refuse tasting the sausage. I therefore refer to them as a moral refuser or a non-moral refuser.

In Experiment 2.1, participants' physiological responses were measured while delivering a speech on camera—ostensibly to be sent to the refusing participant. The key findings were that participants who were exposed to the reaction of a moral refuser displayed more physiological signs of threat and evaluated the refuser as less likeable than participants who were exposed to the reaction of a non-moral refuser. This supported my reasoning that being confronted with a moral refuser is threatening, and leads to negative reactions.

In Experiment 2.2, I investigated whether this negative reaction could be prevented by buffering the moral self-concept. According to earlier work and theorizing (e.g., Zhong & Liljenquist, 2006), people's sense of morality is connected to their sense of physical cleanliness, which means that people should feel more moral after cleansing their body. Therefore, in Experiment 2.2, half of the participants engaged in physical cleansing (i.e., they cleansed their hands with antibacterial soap) before being exposed to the reaction of the refuser. Importantly, participants who had cleansed their hands before they were confronted with a moral refuser did not show negative reactions on refuser evaluations, and evaluations of the self. These findings further supported my reasoning that a threatened self-concept underlies negative reactions to a moral refuser.

In Chapter 3, I investigated a similar process, but within a different domain. Here, participants were invited to drink and evaluate beer. After drinking beer themselves, participants were exposed to the reaction of a peer who had ostensibly refused to taste the beer out of moral reasons (because he/she thought drinking alcohol was wrong), or out of non-moral reasons (because he/she did not like the taste of alcohol). In reality, the beer that was used in the experiments was non-alcoholic, but participants were unaware of this fact.

In Chapter 3, I was mostly interested in whether the negative reactions of people towards moral refusers would also translate into actual other-directed and self-directed behavior.

1

2

3

4

5

6

&

1

2

3

4

5

6

&

In Experiment 3.1, the most important outcome variable was how much beer participants allocated to the non-drinking peer. In a different room, and outside participant's awareness, the amount of beer participants allocated to the refuser was measured. What I found is that participants allocated more beer to a moral refuser than to a non-moral refuser, unless they had engaged in a self-affirmation manipulation. The effectiveness of the self-affirmation manipulation provided further support that a threatened self-concept underlies negative reactions to exposure with a moral refuser.

To zoom in on the impact of exposure to a moral refuser on the self-concept, I used a self-directed behavioral outcome measure in Experiment 3.2—the amount of time participants kept their hands submersed in ice-cold water. This cold-presser test induces pain and the longer one keeps ones hand submersed, the more pain one inflicts on oneself (Bastian, Jetten, & Fasoli, 2011; Bastian, Jetten, & Stewart, 2013; Chen, Poon, & DeWall, 2015; Jackson, Iezzi, Nagasaka, Fritch, & Gunderson, 2002). Therefore, the cold-presser test can be used as a measure of self-directed punishment. Furthermore, self-punishment (such as inflicting pain to oneself) can also be used to restore one's moral adequacy (Bastian et al., 2013; Cohen & Sherman, 2014). The most important finding from Experiment 3.2 was that people who were exposed to a moral refuser, rather than a non-moral refuser, kept their hands submersed in ice-cold water for a longer period of time. Again, this supported our reasoning that exposure to a moral refuser threatens one's self-concept and invokes the need to cope with this threat, for example by inflicting pain to oneself. In Experiment 3.2, I also assessed how much beer participants allocated to the refuser but this effect failed to reach statistical significance (although the pattern was in the expected direction).

In Chapters 2 and 3, I showed that participants felt threatened by the morally motivated behavior of another person who refused to display behavior that participants had displayed themselves. Moral threat, however, may also arise in situations that are not directly related to one's own previous behavior. This is what I investigated in Chapter 4.

In Chapter 4, conservative Christians read a vignette about a couple that wanted to adopt a child that was orphaned in a natural disaster, and would die if it was not adopted. I manipulated whether the adoptive parents-to-be were described as a heterosexual or lesbian couple, and whether they explicitly identified as Christians or not. As expected, participants had more positive evaluations of a heterosexual couple than of a lesbian couple. Importantly, I demonstrated that while participants had the most positive evaluations of a heterosexual couple who identified as being Christian, they had the most negative evaluations of a lesbian couple that identified as being Christian. Apparently, in-group

deviants (i.e., the Christian lesbian couple) were evaluated extremely negative. These negative evaluations were found on a wide array of variables, ranging from characteristics of the couple to the adoption itself and to perceived similarities between participants and the couple. These findings demonstrate that behavior that would otherwise be considered as highly moral is derogated when it is performed by in-group deviants, arguably because in-group deviants threaten the positive identity of the in-group.

The main message in Chapters 2, 3 and 4 seems somewhat negative: People in various situations tend to react negatively to moral behavior displayed by others. However, other research shows that people may sometimes appreciate the moral behavior of others (see, e.g., Aquino, McFerran, & Laven, 2011; Silvers & Haidt, 2008). Therefore, in Chapter 5, I investigated a situation where people were able to appreciate the moral character of a morally motivated deviant.

More specifically, in Chapter 5, I implemented an experimental setting where people were aware of the moral implications of their behavior while they were performing this behavior. The experiments had the following general set-up: Participants were instructed to write an essay that directly opposed their values. More specifically, they were instructed to write an essay denouncing equal rights for homosexuals. Although participants disagreed strongly with the content of this essay, almost all of them complied with the task and wrote such a counter-attitudinal essay. Notice that this forms a contrast with the experiments in Chapters 2 and 3, where people only became aware of the possible moral implications of their behavior (i.e., tasting sausage or drinking beer) when they were exposed to the reaction of the moral refuser. After writing their own essay, participants were exposed to the reaction of either a moral refuser who had refused to write this anti-gay essay out of moral reasons (i.e., because homosexuality is normal and moral and writing an anti-gay essay would be against one's values) or a control condition. In all experiments, I assessed participants evaluations of themselves and the refuser.

In Experiment 5.1, I manipulated whether participants wrote an anti-gay essay or a pro-gay essay and whether they read the response of a peer who had moral reasons to refuse to write the anti-gay essay or to refuse to write the pro-gay essay. Participants who wrote an anti-gay essay had more negative evaluations of themselves than participants who wrote a pro-gay essay, arguably because writing an anti-gay essay opposed participants' moral values. Furthermore, participants had more positive evaluations of the refuser who refused to write an anti-gay essay, than of the refuser who refused to write a pro-gay essay. These findings demonstrate that although participants' self-evaluations were lowered after

1

2

3

4

5

6

&

1

engaging in behavior that went against their moral beliefs, they appreciated the moral character of another person who refused to do so.

2

3

4

5

6

&

In Experiment 5.2, all participants wrote an anti-gay essay. Half of the participants was exposed to the reaction of a peer who refused to write the anti-gay essay out of moral reasons, while the other half was exposed to the reaction of a peer who refused to write the anti-gay essay out of non-moral reasons (i.e., because of an injured hand). Furthermore, half of the participants learned that their essay would be read by the experimenter, while the other half did not. Here, I found that participants' self-evaluations were only lowered when they read the response of the moral refuser and believed their essay would be read by the experimenter. These findings show that engaging in counter-attitudinal behavior that does not fit with one's moral values has negative consequences for one's self-concept, but only under specific conditions. Furthermore, participants had more positive evaluations of the moral refuser than of the non-moral refuser.

In Experiment 5.3, I investigated the critical role of involvement by manipulating whether participants themselves engaged in the wrongful behavior (the actor conditions) or whether they merely observed the reaction of the moral refuser (the observer conditions). Furthermore, I also manipulated whether the refuser allegedly had moral reasons to refuse the anti-gay essay (i.e., the same task as participants) or an anti-abortion essay (i.e., a different task as participants). Demonstrating the critical role of involvement for one's self-concept, I found that participants had lower self-evaluations when they were actors rather than observers. Furthermore, participants had more positive evaluations of refusers of the anti-gay essay than of those who refused to write the anti-abortion essay.

Taken together, the findings in Chapter 5 demonstrate that people can be easily enticed to engage in behavior they do not agree with, such as writing an anti-gay essay when they endorse pro-gay values. In some cases, this influences their positive self-concepts negatively, for example, when they engage actively in this behavior, the moral dimension of the behavior is made clear by someone else, or when they feel their behavior is visible to others. On a more positive note, I also show converging evidence that people can have positive evaluations of others who refuse to compromise their values.

Reactions to Morally Motivated Deviance

First and foremost, this dissertation provides insight into how people respond to the morally motivated behavior of others—in other words, how people respond to *morally motivated deviance* (Chapter 1). The research described in this dissertation builds on, and

extends, literature on the derogation of moral behavior (e.g., Minson & Monin, 2012; Monin, Sawyer, & Marquez, 2008; Parks & Stone, 2010). I will address the most important theoretical implications of my work below.

Threat

An important contribution to existing literature is my focus on the threatened moral self-concept after exposure to morally motivated deviance. Importantly, I assess self-threat with relatively direct measures, something that has not yet received systematic research attention so far. More specifically, in Chapter 2, I demonstrate that when people are confronted with a moral refuser (i.e., someone who refused to eat meat out of moral reasons), they experienced threat on a physiological level. This was indicated by physiological indices of threat, operationalized by the biopsychosocial model of challenge and threat (e.g., Blascovich & Mendes, 2010). In Chapters 2 and 5, I assessed self-threat in a different way and demonstrated that participant's explicit self-concepts were lowered after exposure to morally motivated deviants. Taken together, these findings show that exposure to morally motivated deviance affects responses outside people's awareness (i.e., their bodily reactions) as well as their explicit cognitions about themselves and their behavior.

Furthermore, in earlier work (Minson & Monin, 2012; Monin et al., 2008, Monin, 2007), it was reasoned that self-threat underlies negative reactions to morally motivated deviants. If so, buffering or restoring the self-concept in some way should prevent negative reactions (e.g., Monin et al., 2008). I provide support for this line of reasoning in Chapters 2 and 3. That is, in Chapter 2, I implemented physical cleansing (e.g., Zhong & Liljenquist, 2006) as a way to buffer participants' self-concepts. And indeed, I showed that engaging in physical cleansing, before exposure to the morally motivated deviant, prevented negative reactions. In Chapter 3, I showed that buffering one's self-concept, by engaging in self-affirmation (e.g., Cohen & Sherman, 2014), prevented negative behavioral reactions to a moral refuser who refused to drink alcohol out of moral reasons. Together, these findings provide support for the notion that a self-threat underlies negative reactions to morally motivated deviance.

The Moral Label

Another important contribution of my work is that it disentangles the effects of mere deviance from the motivation to display deviance. In earlier work (e.g., Monin et al., 2008), reactions to a morally motivated deviant were often compared with reactions to a compliant other. For example, Monin and colleagues (2008) investigated people who engaged in a "police task" where they had to pick the most likely perpetrator out of several

1

2

3

4

5

6

&

1

2

3

4

5

6

&

suspects. Afterwards, these people were confronted with either the manipulated reaction of an earlier participant who went along with the task and picked one of the suspects, or with the reaction of an earlier participant who ostensibly refused to participated out of moral reasons (i.e., because the task was racist). Real participants had more negative reactions to the morally motivated deviant than to the compliant other. It was argued that this derogation of the *moral rebel* was due to the moral nature of his/her rebellion (e.g., Monin et al., 2008; Monin, 2007).

However, these negative reactions may also have been caused by the fact that this rebel was deviant and not by his/her moral motivation to deviate. In fact, there is an abundance of research showing that people in general have negative evaluations of those who deviate from norms (Abrams, Marques, Bown, & Henson, 2000; Chekroun & Nugier, 2011; Hichy, Mari, & Capozza, 2008; Mendoza, Lane, & Amodio, 2014; Parks-Stamm, 2013; Parks & Stone, 2010; Pinto, Marques, Levine, & Abrams, 2010; Travaglino, Abrams, De Moura, Marques, & Pinto, 2014). In this dissertation, I tried to rule out this alternative explanation that negative reactions are cause by mere deviance, rather than morally motivated deviance. More specifically, in Chapters 2, 3 and 5, I investigated how compliant participants react to deviants who have either moral or non-moral reasons for their refusal to engage in the task. I demonstrate that negative reactions occur as a response to morally motivated deviance, not deviance alone. Thereby, I provide direct support for the reasoning that negative reactions are evoked by morally motivated deviance, rather than deviance alone.

These findings may have important implications for theory and practice. It may prove to be relevant for literature on deviance (e.g., Abrams et al., 2000; Pinto et al., 2010; Marques, Yzerbyt, & Leyens, 1988) because it demonstrates that it is possible for people to deviate from the norms without social backlash from their peers. This can, for example, also be relevant in whistle-blower situations. Whistle-blowers are often somewhat stronger on moral reasoning and feel it is their moral obligation to blow the whistle (e.g., Mesmer-Magnus & Viswesvaran, 2005). This may indicate that whistleblowers often have moral motivations for their decision to report organizational misconduct. Unfortunately, retaliation by the management or co-workers often occurs. This retaliation can take the form of coercion, exclusion, and sometimes even violence (Mesmer-Magnus & Viswesvaran, 2005). The research reported in this dissertation suggests that having non-moral reasons for engaging in deviant behavior may prevent negative responses. Therefore, it may be possible that when whistleblowers claim to have non-moral reasons for their decision to blow the whistle (e.g., it is in the company's best interest to report wrongdoing rather than they feel it is their moral duty to report wrongdoing), this may prevent retaliation against them. This would be an interesting future direction of research.

Threatening Moral Messages

Another possible future direction would be to investigate how senders can best formulate a moral message to ensure it will be well-received by the receiver. Morally motivated deviants often want to achieve social change that is beneficial for all of us, such as increased equality (e.g., Bashir, Lockwood, Chasteen, Nadolny, & Noyes, 2013), or the abolition and prevention of discrimination of minorities (e.g., Chapter 5; Monin et al., 2008). Therefore, enabling them to successfully convey these messages to others may be very relevant for society as a whole.

This insight may also be relevant for the field of social psychology and the methodologies this field employs. Recently, a lot of attention has been paid to ‘questionable research practices’ (e.g., John, Loewenstein, & Prelec, 2011; Steneck, 2006). However, labeling certain behaviors as ‘questionable’ may be threatening for those who have engaged in that behavior. As I have shown in this dissertation, it can lead people to react defensively, lash out to the messengers, and to question themselves. This is not constructive, I argue, because this may obscure the content of the message and prevents it from being heard and processed by receivers. I propose this is unfortunate because moving from questionable research practices to responsible code of conduct would increase confidence in obtained findings (Steneck, 2006). This would be beneficial for science as well as society and is something that most researchers care strongly about.

Based on my research, it seems plausible that moral messages can be conveyed in ways that are less threatening for perceivers. There are already examples of people who seem to have understood this. For example, Simmons and colleagues (2012) gift-wrapped an important message about improving reporting practices in a non-moral package.

The idea that messages can be threatening, and lead to negative reactions when they are framed in moral terms rather than non-moral terms, also fits with research on value conflicts (e.g., Harinck, De Dreu, & Van Vianen, 2000; Harinck, & Ellemers, 2014; Rexwinkel, Ellemers, & Harinck, 2010) and moral convictions (e.g., Skitka, Washburn, & Carsel, 2015; Tagar, Morgan, Halperin, & Skitka, 2014). For example, Rexwinkel and colleagues (2010) investigated how framing of the exact same conflict influenced conflict resolution. They found that when the conflict was framed in moral terms rather than non-moral terms, people were less likely to adhere to adaptive forms of conflict resolution (i.e., searching for a win-win solution) and more likely to adhere to less constructive forms of conflict resolution (i.e., sticking to one’s own views). Furthermore, Tagar and colleagues (2014) investigated moral convictions. This means that “a given position is based on one’s core moral beliefs

1

2

3

4

5

6

&

1

2

3

4

5

6

&

and convictions” (Skitka et al., 2015, p. 41). Tagar and colleagues (2014) demonstrated that Jewish-Israelis whose political views about the Israeli-Palestinian conflict were based on moral convictions were more likely to endorse extreme political actions (e.g., had higher tolerance for collateral damage) than Jewish-Israelis who whose political views were not based on moral convictions. This implies that when people view conflicts in a moral light, they are less willing to resolve these conflicts.

In short, labeling the exact same behavior or issue as non-moral rather than moral, may prove to be a valuable tool. Especially when people want others to change their behavior, it may be worthwhile to frame a message in a way that enables receivers to hear it and does not threaten receivers’ sense of moral adequacy. Investigating whether changing the framing of a message from moral to non-moral does indeed change their behavior for the better may be a fruitful area of future research.

Methodological and Statistical Reflections

This General Discussion offers me an opportunity to critically reflect on my own research practices. As you may have noticed while reading Chapters 2-5, I have used a wide array of innovative methodologies and different experimental set-ups in this dissertation. Using such a diverse methodological toolbox has several advantages, as well as some disadvantages.

In this dissertation, I have chosen to use immersive behavioral paradigms, where people were actively engaged in the experimental situation. For example, in my experiments, participants were tasting sausage, giving a speech, cleansing their hands (in Chapter 2), drinking beer, writing about important values (Chapter 3) or counter-attitudinal topics (Chapter 5). Afterwards, I measured their responses to these kind of engaging situations with several different measures. For instance, I assessed self-reported self- and other-evaluations (Chapters 2, 3, 4 and 5). I also used implicit outcome measures such as physiological responses during a speech task (Chapter 2) and even actual negative behaviors—beer allocation to an explicit non-drinker and self-inflicted pain by submersing one’s hand into ice-cold water (Chapter 3).

Limitations

Perhaps as a consequence of using innovative measures and paradigms, not all effects are equally strong, and some are even non-significant. For example, in Chapter 3, I did not observe a significant effect of refuser type (moral vs. non-moral) on the allocation

of beer in the second experiment. Thus, participants in Experiment 3.2 did not allocate significantly more beer to the moral refuser than to the non-moral refuser. This makes it hard to draw firm conclusions about the robustness of the finding that people allocate more beer to a moral refuser than a non-moral refuser, as observed in Experiment 3.1. However, the means in Experiment 3.2 were in the hypothesized direction. The reason why I think the difference between conditions was not significant, is that the standard deviations were quite large (see Figure 3.2b). Because of these large standard deviations, I might have needed a larger sample to have enough power to detect the difference between the conditions. Furthermore, there were other differences between the two experiments reported in Chapter 3, such as the absence of a self-affirmation manipulation and an extra dependent variable (the cold presser test). These differences may have had unexpected effects on the beer allocation. One way of resolving this would be to run a (close) replication of Experiment 3.1 with a larger sample. I am currently running such a close replication of Experiment 3.1.

Another limitation of my research is that I conducted the research in Chapter 2 before the importance of adequate power—and thus large samples—was as commonly known and accepted as it is now (see, e.g., Simmons, Nelson, Simonsohn, 2011, 2012, 2013). As a consequence, in hindsight, I think power could have been improved in Chapter 2. For example, in Experiment 2.1, cell sizes varied between 22 and 26 participants. Although a minimum of 20 participants per cell was deemed acceptable a few years ago, it is considered low now for most experimental designs (e.g., Simmons et al., 2013). Therefore, I have increased the sample sizes of my experiments and/or conducted power analyses in Chapters 3-5. Moreover, in line with recent guidelines, in Chapters 3, 4 and 5, I explicitly mention how I have determined sample size (Simmons et al., 2012). More specifically, in Chapter 3, I report power analyses; in Chapters 4 and 5, I aimed to collect a minimum of 50 participants per cell. Because of these measures, the cell sizes in the research reported in this dissertation have increased from a minimum of 22 in Chapter 2, to a minimum of 47 in Chapter 5.

Furthermore, another issue has to do with the clear distinction between main analyses and exploratory or additional analyses. In Chapter 2, I did not make such a distinction which may have inadvertently increased the chance of false positive findings (e.g., Simonsohn, Nelson, & Simmons, 2014). For instance, in retrospect, I think it would have been more accurate to position the moderation with moral identity in Chapter 2 under the heading of additional analyses. Because of these potential negative consequences of not specifying the nature of the analyses, I made a distinction between main analyses and additional

1

2

3

4

5

6

&

1

2

3

4

5

6

&

analyses in Chapters 3, 4 and 5. More specifically, in Chapters 3, 4 and 5, the explicitly stated hypotheses were tested with the analyses specified under the “main analyses” headings, while potential alternative explanations and/or not explicitly hypothesized findings were reported under the “additional analyses” headings. Although this is certainly a step in the right direction, it may be even more informative to follow up interesting research with pre-registered (replication) experiments (e.g., Van ’t Veer & Giner-Sorolla, 2016).¹ Concluding, although there may have been elements in the research that could have been improved, I am confident that the findings reported in this dissertation reflect valuable insights about how people respond to the moral behavior of others.

The Benefits of Immersive Behavioral Paradigms

Implementing immersive behavioral paradigms such as the ones I have used is challenging. Almost every aspect of doing research with behavioral elements—most notably designing a study, designing the materials, getting IRB approval, preparing all the materials, collecting the data, and coding and analyzing the data—takes more effort and is therefore more expensive in terms of energy, time and money than running a study without behavioral elements. To give some examples, coding and analyzing the physiological data and video data in Chapter 2 lasted a few weeks and collecting data for the two studies reported in Chapter 3 took about three months of fulltime data collection. This forms a sharp contrast to studies collected on online platforms such as mTurk, where one can collect data of several hundred participants within 24 hours and analyze the data within a few hours. Furthermore, publishing papers with only behavioral data is difficult (e.g., Baumeister, Vohs, & Funder, 2007), so rewards (in the form of publications) can sometimes be difficult to obtain. In short, using behavioral interaction paradigms in your research can be a frustrating and lengthy process that is not easily rewarded. So why did I—or should anyone, for that matter—bother?

First, and most importantly, some research domains may be best investigated when people are immersed into an actual setting where they experience and undergo the phenomenon you are interested in as a researcher. As I also discussed in the General Introduction, morality is a tricky subject to investigate. Although people do not agree on the content of morality, they often do care deeply about being moral (e.g., Schaumberg & Wiltermuth, 2014; Schwartz, 1992). People are also extremely vigilant to threats to their moral self-concepts

¹ As a matter of fact, I am currently in the process of executing a pre-registered (close) replication of Experiment 3.1, in order to investigate the robustness of the effect reported here.

(Cohen & Sherman, 2014). One of the consequences hereof is that people are not very willing or able to think about, or report, their own moral shortcomings or failures to do the right thing. This is neatly demonstrated by the observation that people tend to forget their own immoral behavior or remember immoral acts less clearly and vividly than other types of behavior they performed (Kouchaki & Gino, 2016). It also becomes evident in the large discrepancy between how people indicate they would react in hypothetical moral situations and how they react in actual moral situations (see, e.g., Baumert, Halmburger, & Schmitt, 2013; Monin et al., 2008). People's self-concepts can be strongly affected after engaging in transgressions. For example, in Chapter 5, I demonstrated that people who wrote an anti-gay essay themselves had lower self-evaluations than people who did not write an anti-gay essay. Because being a good and moral person is so important to people (e.g., Aquino & Reed, 2002; Cohen & Sherman, 2014; Monin, 2007; Sherman & Cohen, 2002, 2006), especially situations that you expect to threaten or undermine people's moral self-concepts are, in my view, best investigated when people actually experience this situation.

Another good reason to use immersive behavioral paradigms is that it is incredibly fascinating. How riveting is it to learn, for example, that participants picked up a can of beer, and poured a real glass of actual beer for another person? And not just any person, but someone who explicitly indicated not to drink alcohol out of moral reasons (Chapter 3). Or that people who were exposed to the reaction of a moral non-drinker kept their hands submerged in ice-cold water for about 10 seconds longer than people who were exposed to the reaction of a non-moral non-drinker? 10 seconds is a long time when you are in pain, try it yourself.

I remember learning about social psychological experiments (e.g., Asch, 1956; Milgram, 1963; Zimbardo, 1973) for the first time in an introductory course. It has led me to fall deeply and madly in love with social psychology and has inspired me to become a social psychological researcher. However, as inevitably happens when you fall in love is that you learn things about the object of your affection that you do not particularly like. For me, this was the realization that the role of actual behavior and behavioral paradigms in current social psychological studies has declined tremendously in the last decades (Baumeister et al., 2007). This has even led researchers to argue that (personality and) social psychology is no longer the science of behavior, but rather "the science of self-reports and finger movements" (Baumeister et al., 2007, p. 396)—arguably because most studies use self-reported ratings, key-strokes and reaction times as outcome variables. Thus, using immersive behavioral paradigms is a way to contribute to the social psychological tradition that leads to strong and important insights about human behavior in situations that closely resemble reality.

1

2

3

4

5

6

&

1

Recommendations and Future Directions

2

Positive Reactions

3

An important future direction may be to explore reactions to morally motivated deviants in a longitudinal setting. It may be the case that although primary reactions to morally motivated deviants are negative (as we have seen in the work reported in this dissertation), people may learn from this confrontation to change their own behavior in future situations. I have some anecdotal evidence to support this suggestion. For instance, In Chapter 5, I reported three experiments that were somewhat similar, and it incidentally occurred that participants participated in more than one of these experiments. Sometimes these non-naïve participants became moral refusers themselves, the second time they participated. For instance, one participant in Experiment 5.3 wrote that *"I was supposed to write an essay about how abnormal and unnatural homosexuality is, and that homosexuals should not have the same rights as heterosexuals. However, this is directly opposed to my personal convictions so I cannot and will not put myself in this position. Everyone should have the same rights"* (Participant #47). As becomes clear from this reaction, it is almost a direct quote of the moral refuser manipulation I used in Chapter 5. This seems to suggest that witnessing a refusal to engage in counter-attitudinal behavior can liberate people to act in line with their own moral beliefs in future situations. Systematically investigating under which conditions people learn from the exemplary behavior of morally motivated deviants therefore seems like an interesting direction for future research.

4

5

6

&

Another interesting direction may be to focus on what makes people appreciate rather than derogate morally motivated deviants. There may be contexts where they are not threatening, but accepted or even admired. It may be insightful to link the current work to research on moral elevation (e.g., Aquino, McFerran, & Laven, 2011) and heroism (e.g., Franco, Blau, & Zimbardo, 2011). Although I did not focus on the deviants themselves, but rather on reactions to them, it may also be interesting to investigate what makes people decide to become deviants. As you may recall, in Chapters 2, 3 and 5, there were always some participants who refused to engage in the task themselves. It may be interesting to investigate which personal and situational factors increase or decrease the likelihood that people will stick to their moral beliefs and behave as moral refusers themselves.

Group Settings

Another possibility may be to investigate reactions to morally motivated deviants in group settings, rather than interpersonal settings. Morally motivated deviance may be a form of deviance that is beneficial for other group members, and evoke positive reactions as

a consequence. One of the ways in which morally motivated deviance can be beneficial, is because it may increase the level of morality of the whole group. People consider their group's morality to be the most important characteristic of their group, want to belong to moral groups (Brambilla, Sacchi, Pagliaro, & Ellemers, 2013; Ellemers, Pagliaro, Barreto, & Leach, 2008; Ellemers, Pagliaro, & Barreto, 2013; Ellemers & Van den Bos, 2012; Leach, Ellemers, Barreto, 2007) and are vigilant to moral threats to their group (Täuber & Van Zomeren, 2013).

For example, Parks-Stamm (2013) demonstrated that people excluded high-performing group members from their group when they were focused on themselves, but included these members when they were focused on the group as a whole. Apparently, focusing on oneself leads to a threatening comparison between one's own performance and the superior performance of the deviant. However, focusing on your group as a whole leads to the positive reflection of the deviant's excellent performance on the whole team. In a parallel fashion, focusing on what morally motivated deviants mean for you personally can be threatening, but focusing on what morally motivated deviants mean for your group's moral standing may lead to positive reactions.

Mundane Morality in the Social World

Finally, future research into moral reactions may want to focus on mundane morality. The moral behaviors people experience and perform in their regular day-to-day live are very far away from artificial settings such as moral dilemmas and thought experiments (Hofmann, Wisneski, Brandt, & Skitka, 2014). In fact, as I have argued in the General Introduction, morality may center around mundane events, such as eating, drinking, love, and raising children, to name just a few. This was also demonstrated by Hofmann and colleagues (2014) who assessed daily morality in 1252 people over the course of three days. Participants were signaled five times a day on their mobile phones and answered whether they had committed, were a target of, witnessed or learned about, immoral or moral behavior in the past hour. In this research, something morally relevant had occurred in the past hour for about 30 percent of the reactions. Morality is thus a very pervasive and important aspect of people's live, that has a large influence on their happiness and sense of meaning (Hofmann et al., 2014). Perhaps even more relevant for the research discussed in this dissertation is that most morally relevant behaviors occur in social settings (i.e., 64.3%. Hofmann et al., 2014).

These insights strengthen my belief that the settings I have used in this dissertation are well-suited for investigating morality. Topics such as eating meat, drinking alcohol, and gay rights are mundane moral topics that people have ample experience with in their daily

1

2

3

4

5

6

&

1

lives. The study of morality in social settings where people are exposed to the morally motivated behavior of others invites researchers to directly address and incorporate the fact that our morality is shaped in relation to others and does not exist in a social vacuum.

2

3

Concluding Words

4

I have investigated what happens to people when they are confronted with morally motivated behavior of other people. I have tried to answer fundamental question about how morality functions and how this influences people in their daily lives. More specifically, in this dissertation, I have used eating meat, drinking beer, supporting (or opposing) gay rights and adoption as behaviors of interest. These behaviors cover a wide range of content domains and some of them may be considered as more or less moral than others. Important to note is that I do not argue that these behaviors *are* moral behaviors, but I have investigated what happens when these behaviors are labeled or perceived as such. What we have seen is that moral labels—regardless of the exact content of the behaviors that are labeled—have a strong influence on people’s self-concepts, their evaluations of others, and their behaviors. Especially important in this process is people’s aversive reaction to engage in behaviors that are labeled as immoral by others. Thus, the construct of morality regulates people’s attitudes and behaviors in the social world.

5

6

&





&

Addendum

References

Dutch Summary | Nederlandse Samenvatting

Acknowledgements | Dankwoord

KLI dissertation series

References

- Abrams, D., Marques, J. M., Bown, N. J., & Henson, M. (2000). Pro-norm and anti-norm deviance. *Journal of Personality and Social Psychology, 78*(5), 906-912.
- Abrams, D., Rutland, A., Cameron, L., & Marques, J. M. (2003). The development of subjective group dynamics: When in-group bias gets specific. *British Journal of Developmental Psychology, 21*(2), 155-176. doi: 10.1348/026151003765264020
- Abrams, D., Rutland, A., Ferrell, J. M., & Pelletier, J. (2008). Children's judgments of disloyal and immoral peer behavior: Subjective group dynamics in minimal intergroup contexts. *Child Development, 79*, 444-461.
- Ali, M. M., & Dwyer, D. S. (2010). Social network effects in alcohol consumption among adolescents. *Addictive Behaviors, 35*(4), 337-342. doi:10.1016/j.addbeh.2009.12.002
- Anderson, C., & Bushman, B. (2002). Human aggression. *Annual Review of Psychology, 53*, 27-51. doi:10.1146/annurev.psych.53.100901.135231
- Aquino, K., Freeman, D., Reed, A., II, Lim, V. K. G., & Felps, W. (2009). Testing a social-cognitive model of moral behavior: The interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology, 97*(1), 123-141. doi:10.1037/a0015406
- Aquino, K., McFerran, B., & Laven, M. (2011). Moral identity and the experience of moral elevation in response to acts of uncommon goodness. *Journal of Personality and Social Psychology, 100*(4), 703-718.
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423-1440. doi:10.1037//0022-3514.83.6.1423
- Aquino, K., Reed, A., Thau, S., & Freeman, D. (2007). A grotesque and dark beauty: How moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. *Journal of Experimental Social Psychology, 43*, 385-392. doi:10.1016/j.jesp.2006.05.013
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (Ed.) *Groups, leadership and men*. Pittsburgh, PA: Carnegie Press.
- Asch, S. E. (1956). Studies of independence and conformity. A minority of one against a unanimous majority. *Psychological Monographs, 70*(9), 1-70
- Ayduk, O., Gyurak, A., & Luerssen, A. (2008). Individual differences in the rejection-aggression link in the hot sauce paradigm: The case of rejection sensitivity. *Journal of Experimental Social Psychology, 44*(3), 775-782. doi:10.1016/j.jesp.2007.07.004
- Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology, 107*(3), 393-413. doi: 10.1037/a0037207

1

2

3

4

5

6

&

1

Bashir, N. Y., Lockwood, P., Chasteen, A. L., Nadolny, D., & Noyes, I. (2013). The ironic impact of activists: Negative stereotypes reduce social change influence. *European Journal of Social Psychology, 43*(7), 614-626. doi: 10.1002/ejsp.1983

2

Bastian, B., Jetten, J., & Fasoli, F. (2011). Cleansing the soul by hurting the flesh: The guilt-reducing effect of pain. *Psychological Science, 22*(3), 334-335. doi:10.1177/0956797610397058

3

Bastian, B., Jetten, J., Hornsey, M. J., & Leknes, S. (2014). The positive consequences of pain: A biopsychosocial approach. *Personality and Social Psychology Review, 18*(3), 256-279. doi:10.1177/1088868314527831

4

5

Bastian, B., Jetten, J., & Stewart, E. (2013). Physical pain and guilty pleasures. *Social Psychological and Personality Science, 4*(2), 215-219. doi:10.1177/1948550612451156

6

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters, 2*, 412-414. doi: 10.1098/rsbl.2006.0509

&

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396-403. doi: 10.1111/j.1745-6916.2007.00051.x

Baumert, A., Halmburger, A., & Schmitt, M. (2013). Interventions against norm violations: Dispositional determinants of self-reported and real moral courage. *Personality and Social Psychology Bulletin, 39*(8), 1053-1068. doi: 10.1177/0146167213490032

Beauchamp, T. L. (1991). *Philosophical ethics: An introduction to moral philosophy*. New York, New York: McGraw-Hill.

Begue, L., Subra, B., Arvers, P., Muller, D., Bricout, V., & Zorman, M. (2009). A message in a bottle: Extrapharmacological effects of alcohol on aggression. *Journal of Experimental Social Psychology, 45*(1), 137-142. doi:10.1016/j.jesp.2008.07.018

Bentham, J. (1789). *An introduction to the principles of morals and legislation*. Oxford: Clarendon Press.

Besen, Y., & Zicklin, G. (2007). Young men, religion and attitudes towards homosexuality. *Journal of Men, Masculinities, and Spirituality, 1*(3), 250-266.

Blascovich, J. (2008). Challenge and threat. In A. J. Elliott (Ed.), *Handbook of approach and avoidance motivation* (pp. 431-445). New York, NY: Psychology Press.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology, 80*(2), 253-267. doi: 10.1037//0022-3514.80.2.253

Blascovich, J., Mendes, W. B., & Seery, M. D. (2002). Intergroup encounters and threat: A multi-method approach. In D. M. Mackie, & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 89-109). New York: Psychology Press.

- Blascovich, J., Mendes, W. B., Tomaka, J., Salomon, K., & Seery, M. (2003). The robust nature of the biopsychosocial model of challenge and threat: A reply to Wright and Kirby. *Personality and Social Psychology Review*, *7*(3), 234-243.
- Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 5, pp. 194-33). New York: Wiley.
- Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, R. K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology*, *40*(5), 683-688. doi:10.1016/j.jesp.2003.10.007
- Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 1-51). San Diego, CA: Academic Press.. doi:10.1016/S0065-2601(08)60235-X
- Brambilla, M., Sacchi, S., Pagliaro, S., & Ellemers, N. (2013). Morality and intergroup relations: Threats to safety and group image predict the desire to interact with outgroup and ingroup members. *Journal of Experimental Social Psychology*, *49*(5), 811-821. doi:10.1016/j.jesp.2013.04.005
- Broeders, R., Van den Bos, K., Müller, P. A., & Ham, J. (2011). *Journal of Experimental Social Psychology*, *47*, 923-934. doi: 10.1016/j.jesp.2011.03.018
- Chekroun, P., & Nugier, A. (2011). "I'm ashamed because of you, so please, don't do that!": Reactions to deviance as a protection against a threat to social image. *European Journal of Social Psychology*, *41*(4), 479-488. doi:10.1002/ejsp.809
- Chen, Z., Poon, K., & DeWall, C. N. (2015). When do socially accepted people feel ostracized? Physical pain triggers social pain. *Social Influence*, *10*(1), 68-76. doi:10.1080/15534510.2014.926290
- Choo, H., & Shek, D. (2013). Quality of parent-child relationship, family conflict, peer pressure, and drinking behaviors of adolescents in an asian context: The case of singapore. *Social Indicators Research*, *110*(3), 1141-1157. doi:10.1007/s11205-011-9977-4
- Cikara, M., Farnsworth, R. A., Harris, L. T., & Fiske, S. T. (2010). On the wrong side of the trolley track: Neural correlates of relative social valuation. *Social Cognitive and Affective Neuroscience*, *5*, 404-413. doi:10.1093/scan/nsq011
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis in the behavioural sciences* (Third Edition ed.). Mahwah: New Jersey: Lawrence Erlbaum Associates, Publishers.
- Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*, 333-371. doi:10.1146/annurev-psych-010213-115137

1

2

3

4

5

6

&

1

Cook, R. D. (1977). Detection of influential observation in linear-regression. *Technometrics*, 19(1), 15-18. doi:10.2307/1268249

2

Cramwinckel, F. M., De Cremer, D., & Van Dijke, M. (2013). Dirty hands make dirty leaders?! The effects of touching dirty objects on rewarding unethical subordinates as a function of leader's self-interest. *Journal of Business Ethics*, 115(1), 93-100. doi:10.1007/s10551-012-1385-4

3

Cramwinckel, F.M., Van den Bos, K., & Van Dijk, E. (2015). Reactions to Morally Motivated Deviance. *Current Opinion in Psychology*, 6, 150-156. doi: 10.1016/j.copsyc.2015.08.007

4

Dawson, J. F., & Richter, A. W. (2006). Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology*, 91(4), 917-926. doi:10.1037/0021-9010.91.4.917

5

6

&

DeWall, C. N., Anderson, C. A., & Bushman, B. J. (2011). The general aggression model: Theoretical extensions to violence. *Psychology of Violence*, 1(3), 245-258. doi:10.1037/a0023842

De Wit, F. R. C., Scheepers, D., & Jehn, K. A. (2012). Cardiovascular reactivity and resistance to opposing viewpoints during intragroup conflict. *Psychophysiology*, 49, 1691-1699. doi: 10.1111/j.1469-8986.2012.01456.x

Diener, E., & Wallbom, M. (1976). Effects of self-awareness on antinormative behavior. *Journal of Research in Personality*, 10(1), 107-111.

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(8), 972-985. doi: 10.1177/0146167214533131

Ellemers, N., Pagliaro, S., & Barreto, M. (2013). Morality and behavioural regulation in groups: A social identity approach. *European Review of Social Psychology*, 24(1), 160-193. doi: 10.1080/10463283.2013.841490

Ellemers, N., Pagliaro, S., Barreto, M., & Leach, C. W. (2008). Is it better to be moral than to be smart? The effects of morality and competence norms on the decision to work at group status improvement. *Journal of Personality and Social Psychology*, 95(6), 1397-1410. doi: 10.1037/a0012628

Ellemers, N., & Van den Bos, K. (2012). Morality in groups: On the social-regulatory functions of right and wrong. *Social and Personality Psychology Compass*, 6(12), 878-889. doi: 10.1111/spc3.12001

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. doi:10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. doi:10.3758/BF03193146

- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117-140. doi: 10.1177/001872675400700202
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203-211.
- Field, A. P. (2009). *Discovering Statistics using SPSS: (and sex and drugs and rock 'n' roll)*. (3rd ed.). London: SAGE publications.
- Foot, P. (1978). *Virtues and vices and other essays in moral philosophy*. Berkeley, California: University of California Press.
- Ford, T.E., Brignall, T., Van Valey, T.L., & Macaluso, M.J. (2009). The Unmaking of Prejudice: How Christian Beliefs Relate to Attitudes toward Homosexuals. *Journal for the Scientific Study of Religion*, 48 (1), 146–160. doi: 10.1111/j.1468-5906.2009.01434.x
- Franco, Z. E., Blau, K., & Zimbardo, P. G. (2011). Heroism: A conceptual analysis and differentiation between heroic action and altruism. *Review of General Psychology*, 15(2), 99-113. doi: 10.1037/a0022672
- Frings, D., Hurst, J., Cleveland, C., Blascovich, J., & Abrams, D. (2012). Challenge, threat, and subjective group dynamics: Reactions to normative and deviant group members. *Group Dynamics-Theory Research and Practice*, 16(2), 105-121. doi:10.1037/a0027504
- Fulton, A. S., Gorsuch, R. L., & Maynard, E. A. (1999). Religious orientation, antihomosexual sentiment, and fundamentalism among Christians. *Journal for the Scientific Study of Religion*, 38(1), 14-22. doi: 10.2307/1387580
- Geen, R. G. (1991). Social motivation. *Annual Review of Psychology*, 42, 377-399. doi: 10.1146/annurev.ps.42.020191.002113
- Gibbons, F., & Buunk, B. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, 76(1), 129-142.
- Godbold, L. C., & Pfau, M. (2000). Conferring resistance to peer pressure among adolescents - using inoculation theory to discourage alcohol use. *Communication Research*, 27(4), 411-437. doi:10.1177/009365000027004001
- Goode, C., Balzarini, R. H., & Smith, H. J. (2014). Positive peer pressure: Priming member prototypicality can decrease undergraduate drinking. *Journal of Applied Social Psychology*, 44(8), 567-578. doi:10.1111/jasp.12248
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111, 364-371. doi: 10.1016/j.cognition.2009.02.001
- Greene, M., & Low, K. (2014). Public integrity, private hypocrisy, and the moral licensing effect. *Social Behavior and Personality: An international journal*, 42(3), 391-400. doi: 10.2224/sbp.2014.42.3.391

1

2

3

4

5

6

&

1

Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: Status, reputation, and conspicuous consumption. *Journal of Personality and Social Psychology, 98*(3), 392-404.

2

Haddock, G. (1993). Assessing the structure of prejudicial attitudes: The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology, 65*(6), 1105-1118.

3

Haidt, J., Rosenberg, E., & Hom, H. (2003). Differentiating diversities: Moral diversity is not like other kinds. *Journal of Applied Social Psychology, 33*(1), 1-36. doi:10.1111/j.1559-1816.2003.tb02071.

4

Harinck, F., De Dreu, C. K. W., & Van Vianen, A. E. M. (2000). The impact of conflict issue on fixed-pie perceptions, problem solving, and integrative outcomes in negotiation. *Organizational Behavior and Human Decision Processes, 81*, 329-358. doi: 10.1006/obdh.1999.2873

5

6

&

Harinck, F. & Ellemers, N. (2014). How conflict issues change the nature of the conflict game. In: C.K.W. De Dreu (Ed.). *Social conflict within and between groups*. Current Issues in Social Psychology, pp. 19-36. East Sussex, UK: Psychology Press

Henrich, J., Heine, S.J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2/3), 61-83. doi:10.1017/S0140525X0999152X

Hichy, Z., Mari, S., & Capozza, D. (2008). Pronorm and antinorm deviants: A test of the subjective group dynamics model. *Journal of Social Psychology, 148*(5), 641-645. doi: 10.3200/SOCP.148.5.641-644

Hicks, G. R., & Lee, T. T. (2006). Public attitudes toward gays and lesbians: trends and predictors. *Journal of Homosexuality, 51*(2), 57-77.

Hirt, E. R., McCrea, S. M., & Kimble, C. E. (2000). Public self-focus and sex differences in behavioral self-handicapping: Does increasing self-threat still make it "just a man's game"? *Personality and Social Psychology Bulletin, 26*(9), 1131-1141.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science, 345*(6202), 1340-1343. doi: 10.1126/science.1251560

Inbar, Y., Pizarro, D., Gilovich, T., & Ariely, D. (2013). Moral masochism: On the connection between guilt and self-punishment. *Emotion, 13*(1), 14-8. doi:10.1037/a0029749

Jackson, T., Iezzi, T., Nagasaka, T., Fritch, A., & Gunderson, J. (2002). Does the mere presence of over-the-counter pain medication affect pain perception? Some preliminary findings. *Psychology, Health, & Medicine, 7*(2), 215-7. doi:10.1080/13548500120116102

Jamieson, J. P., Valdesolo, P., & Peters, B. J. (2014). Sympathy for the devil? The physiological and psychological effects of being an agent (and target) of dissent during intragroup conflict. *Journal of Experimental Social Psychology, 55*, 221-227. Doi: 10.1016/j.jesp.2014.07.011

John, L. K., Loewenstein, G., & Prelec, D. (2011). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524-532. doi: 10.1177/0956797611430953

- Kouchaki, M. & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences of the United States of America*, advance online publication. doi: 10.1073/pnas.1523586113
- Kant, I. (1785). *Groundwork for the metaphysics of morals*. (A. W. Wood, Trans.). Binghampton, New York: Vail-Ballou Press.
- Kim, J., & Kim, K. S. (2012). The role of sensation seeking, perceived peer pressure, and harmful alcohol use in riding with an alcohol-impaired driver. *Accident Analysis and Prevention*, 48, 326-334. doi:10.1016/j.aap.2012.01.033
- Kite, M. E., & Deaux, K. (1986). Attitudes toward homosexuality - assessment and behavioral consequences. *Basic and Applied Social Psychology*, 7(2), 137-162. doi:10.1207/s15324834basp0702_4
- Kite, M. E., & Whitley, B. E. (1996). Sex Differences in Attitudes Toward Homosexual Persons, Behaviors, and Civil Rights: A Meta-Analysis. *Personality and Social Psychology Bulletin*, 22(4), 336-353. doi: 10.1177/0146167296224002
- Knee, C. R., & Neighbors, C. (2002). Self-determination, perception of peer pressure, and drinking among college students. *Journal of Applied Social Psychology*, 32(3), 522-543. doi:10.1111/j.1559-1816.2002.tb00228.x
- Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin*, 136(6), 1045-1066. doi:10.1037/a0020970
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93(2), 234-249. doi: 10.1037/0022-3514.93.2.234
- Leary, M. R. (1983). A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin*, 9(3), 371-375.
- Lee, S. W. S., & Schwarz, N. (2010). Washing Away Postdecisional Dissonance. *Science*, 328, 709. doi: 10.1126/science.1186799
- Lee, S. W. S., & Schwarz, N. (2011). Wiping the slate clean: Psychological consequences of physical cleansing. *Current Directions in Psychological Science*, 20(5), 307-311. doi:10.1177/0963721411422694
- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Lieberman, J., Solomon, S., Greenberg, J., & McGregor, H. (1999). A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior*, 25(5), 331-348. doi:10.1002/(SICI)1098-2337(1999)25:5<331::AID-AB2>3.0.CO;2-1
- Liu, T. J., & Steele, C. M. (1986). Attributional analysis as self-affirmation. *Journal of Personality and Social Psychology*, 51(3), 531-540. doi:10.1037/0022-3514.51.3.531

1

2

3

4

5

6

&

1

Marques, J. M., Abrams, D., Paez, D., & Martinez-Taboada, C. (1998). The role of categorization and in-group norms in judgments of groups and their members. *Journal of Personality and Social Psychology, 75*, 976- 988.

2

Marques, J. M., Abrams, D., & Serôdio, R. G. (2001). Being Better by Being Right: Subjective Group Dynamics and Derogation of In-Group Deviants When Generic Norms are Undermined. *Journal of Personality and Social Psychology, 81*(3), 436-447. doi: 10.1037//0022-3514.81.3.436

3

4

Marques, J. M., Yzerbyt, V. Y., & Leyens, J. P. (1988). The black sheep effect - extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology, 18*(1), 1-16. doi:10.1002/ejsp.2420180102

5

6

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633-644.

&

McKay, M. T., & Cole, J. C. (2012). The relationship between alcohol use and peer pressure susceptibility, peer popularity and general conformity in northern irish school children. *Drugs-Education Prevention and Policy, 19*(3), 213-222. doi:10.3109/09687637.2011.641133

Mendoza, S. A., Lane, S. P., & Amodio, D. M. (2014). For members only: Ingroup punishment of fairness norm violations in the ultimatum game. *Social Psychological and Personality Science, 5*(6), 662-670. doi: 10.1177/1948550614527115

Mesmer-Magnus, J. R., & Viswesvaran, C. (2005). Whistleblowing in organizations: an examination of correlates of whistleblowing intentions, actions, and retaliation. *Journal of Business Ethics, 62*, 277-297. doi: 10.1007/s10551-005-0849-1

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*, 371-378.

Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son, and Bourn, West Strand.

Minson, J. A., & Monin, B. (2012). Do-gooder derogation: Disparaging morally motivated minorities to defuse anticipated reproach. *Social Psychological and Personality Science, 3*(2), 200-7. doi:10.1177/1948550611415695

Monin, B. (2007). Holier than me? threatening social comparison in the moral domain. *Revue Internationale De Psychologie Sociale-International Review of Social Psychology, 20*(1), 53-68.

Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*(1), 76-93. doi:10.1037/0022-3514.95.1.76

Nelissen, R. M. A. (2012). Guilt-induced self-punishment as a sign of remorse. *Social Psychological and Personality Science, 3*(2), 139-144. doi:10.1177/1948550611411520

Nelissen, R. M. A., & Zeelenberg, M. (2009). When guilt evokes self-punishment: Evidence for the existence of a dooby effect. *Emotion, 9*(1), 118-122. doi:10.1037/a0014540

- Newman, G. E., & Cain, D. M. (2014). Tainted altruism: When doing some good is evaluated worse than doing no good at all. *Psychological Science, 25*(3), 648-655. doi: 10.1177/0956797613504785
- Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition, 100*, 530-542. doi: 10.1016/j.cognition.2005.07.005
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois.
- Pagliaro, S., Ellemers, N., & Barreto, M. (2011). Sharing moral values: Anticipated ingroup respect as a determinant of adherence to morality-based (but not competence-based) group norms. *Personality and Social Psychology Bulletin, 37*(8), 1117-1129. doi:10.1177/0146167211406906
- Parks-Stamm, E. J. (2013). Exclusion as self-protection: The function of subtypes for ingroup members. *Personality and Social Psychology Bulletin, 39*(5), 651-662. doi: 10.1177/0146167213479978
- Parks, C. D., & Stone, A. B. (2010). The desire to expel unselfish members from the group. *Journal of Personality and Social Psychology, 99*(2), 303-313. doi: 10.1037/a0018403
- Pedersen, W., Vasquez, E., Bartholow, B., Grosvenor, M., & Truong, A. (2014). Are you insulting me? exposure to alcohol primes increases aggression following ambiguous provocation. *Personality & Social Psychology Bulletin, 40*(8), 1037-1049. doi:10.1177/0146167214534993
- Pinto, I. R., Marques, J. M., Levine, J. M., & Abrams, D. (2010). Membership status and subjective group dynamics: Who triggers the black sheep effect? *Journal of Personality and Social Psychology, 99*(1), 107-119. doi:10.1037/a0018187
- Poelen, E. A. P., Engels, R. C. M. E., Van Der Vorst, H., Scholte, R. H. J., & Vermulst, A. A. (2007). Best friends and alcohol consumption in adolescence: A within-family analysis. *Drug and Alcohol Dependence, 88*(2-3), 163-173. doi:10.1016/j.drugalcdep.2006.10.008
- Pronin, E. (2008). How we see ourselves and how we see others. *Science, 320*(5880), 1177-1180. doi:10.1126/science.1154199
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*(1), 57-75. doi: 10.1037/a0021867
- Reed, A., & Aquino, K. F. (2003). Moral identity and the expanding circle of moral regard toward out-groups. *Journal of Personality and Social Psychology, 84*(6), 1270-1286. doi:10.1037/0022-3514.84.6.1270
- Rexwinkel, R. B., Ellemers, N., & Harinck, F. (2010). Kom je aan mijn waarden, dan kom je aan mij! Hoe verwijzingen naar verschillende waarden of informatie bepalen hoe mensen

1

2

3

4

5

6

&

1

een conflict benaderen. In B. Derks, R. J., Renes, K. Ruys, N. Van de Ven, & M. Vliek (Ed.), *Jaarboek Sociale Psychologie 2010* (pp. 323-330). ASPO pers, Groningen: The Netherlands.

2

Rijksinstituut voor Volksgezondheid en Milieu (RIVM, 2009). *Ranking van drugs: Een vergelijking van de schadelijkheid van drugs*. (RIVM publication no. 340001001). Retrieved from <http://www.rivm.nl/bibliotheek/rapporten/340001001.html>

3

Ritter, D., & Eslea, M. (2005). Hot sauce, toy guns, and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior, 31*(5), 407-419. doi:10.1002/ab.20066

4

5

Rutchick, A. M., & Slepian, M. L. (2013). Handling ibuprofen increases pain tolerance and decreases perceived pain intensity in a cold pressor test. *Plos One, 8*(3), e56175. doi:10.1371/journal.pone.0056175

6

&

Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Science, 5*, 279-291. doi:10.1177/1745691610369468

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science, 20*(4), 523-528. doi: 10.1111/j.1467-9280.2009.02326.x

Schaumberg, R. L., & Wiltermuth, S.S. (2014). Desire for a positive moral self-regard exacerbates escalation of commitment to initiatives with prosocial aims. *Organizational Behavior and Human Decision Processes, 123*, 110-113.

Scheepers, D. T. (2009). Turning social identity threat into challenge: Status stability and cardiovascular reactivity during inter-group competition. *Journal of Experimental Social Psychology, 45*(1), 228-233. doi: 10.1016/j.jesp.2008.09.011

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096-1109. doi:10.1177/0146167208317771

Schnall, S., & Roper, J. (2012). Elevation puts moral values into action. *Social Psychological and Personality Science, 3*(3), 373-378. doi:10.1177/1948550611423595

Schnall, S., Roper, J., & Fessler, D. M. T. (2010). Elevation leads to altruistic behavior. *Psychological Science, 21*(3), 315-320.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25*, 1-65.

Seery, M. D., Blascovich, J., Weisbuch, M., & Vick, B. (2004). The relationship between self-esteem level, self-esteem stability, and cardiovascular reactions to performance feedback. *Journal of Personality and Social Psychology, 87*(1), 133-145. doi:10.1037/0022-3514.87.1.133

Sethi, S., & Seligman, M. E. P. (1993). Optimism and fundamentalism. *Psychological Science, 4*(4), 256-259.

- Shao, R., Aquino, K., & Freeman, D. (2008). Beyond moral reasoning: A review of moral identity research and its implications for business ethics. *Business Ethics Quarterly*, *18*(4), 513-540.
- Sherman, D. K., & Cohen, G. L. (2002). Accepting threatening information: Self-affirmation and the reduction of defensive biases. *Current Directions in Psychological Science*, *11*(4), 119-123. doi:10.1111/1467-8721.00182
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. *Advances in Experimental Social Psychology*, *38*(38), 183-242. doi:10.1016/S0065-2601(06)38004-5
- Silvers, J. A., & Haidt, J. (2008). Moral elevation can induce nursing. *Emotion*, *8*(2), 291-295. doi: 10.1037/1528-3542.8.2.291
- Simons-Morton, B., Haynie, D. L., Crump, A. D., Eitel, P., & Saylor, K. E. (2001). Peer and parent influences on smoking and drinking among early adolescents. *Health Education & Behavior*, *28*(1), 95-107.
- Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359-1366. doi: 10.1177/0956797611417632
- Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2012). A 21 word solution. *Dialogue: the Official Newsletter for the Society of Personality and Social Psychology*, *26*(2), 4-7.
- Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2013). Life after p-hacking. *Meeting of the Society for Personality and Social Psychology*, New Orleans, LA, 17-19 January 2013
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, *Advance online publication* doi:10.1177/0956797614567341
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*(2), 534-547. doi: 10.1037/a0033242
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, *4*, 267-281.
- Skitka, L., Bauman, C., & Sargis, E. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*(6), 895-917. doi:10.1037/0022-3514.88.6.895
- Skitka, L. J., Washburn, A. N., & Carsel, T. S. (2015). The psychological foundations and consequences of moral conviction. *Current Opinion in Psychology*, *6*, 41-44. doi: 10.1016/j.copsyc.2015.03.025
- Statistics Netherlands [CBS]. (2012, September 14). Vruchtbaarheid van mannen en vrouwen naar opleidingsniveau. Retrieved from <http://www.cbs.nl/nl-NL/menu/themas/dossiers/vrouwen-en-mannen/publicaties/artikelen/archief/2012/2010-09-btwob-vruchtbaarheid-opleidingsniveau.htm>

1

2

3

4

5

6

&

1

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*(1), 5-19. doi:10.1037/0022-3514.45.1.5

2

Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, *12*(1), 53-74. doi: 10.1007/s11948-006-0006-y

3

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson International Edition.

4

Tagar, M. R., Morgan, G. S., Halperin, E., & Skitka, L. J. (2014). When ideology matters: Moral conviction and the association between ideology and policy preferences in the Israeli-Palestinian conflict. *European Journal of Social Psychology*, *44*(2), 117-125. doi: 10.1002/ejsp.1993

5

Tauber, S. & Van Zomeren, M. (2013). Outrage towards whom? Threats to moral group status impede striving to improve via outgroup-directed outrage. *European Journal of Social Psychology*, *43*(2), 149-159. doi: 10.1002/ejsp.1930

6

&

The Netherlands Institute for Social Research [SCP] (2013). *Acceptance of lesbian, gay, bisexual and transgender individuals in the Netherlands 2013*. Retrieved from http://www.scp.nl/english/Publications/Publications_by_year/Publications_2013/Acceptance_of_lesbian_gay_bisexual_and_transgender_individuals_in_the_Netherlands_2013

The Netherlands Institute for Social Research [SCP] (2016). *LGBT-monitor 2016 Summary*. Retrieved from http://www.scp.nl/english/Publications/Summaries_by_year/Summaries_2016/LGBT_Monitor_2016 on May 27th, 2016

Travaglino, G. A., Abrams, D., Randsley de Moura, G., Marques, J. M., & Pinto, I. R. (2014). How groups react to disloyalty in the context of intergroup competition: Evaluations of group deserters and defectors. *Journal of Experimental Social Psychology*, *54*, 178-187. doi:10.1016/j.jesp.2014.05.006

Van 't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology- A discussion and suggested template. *Journal of Experimental Social Psychology*, advance online publication. doi: 10.1016/j.jesp.2016.03.004

Van den Bos, K., Bruins, J., Wilke, H., & Dronkert, E. (1999). Sometimes unfair procedures have nice aspects: On the psychology of the fair process effect. *Journal of Personality and Social Psychology*, *77*(2), 324-336.

Van den Bos, K., Van Lange, P. A. M., Lind, E. A., Venhoeven, L. A., Beudeker, D. A., Cramwinckel, F. M., Smulders, L., Van der Laan, J. (2011). On the benign qualities of behavioral disinhibition: Because of the prosocial nature of people, behavioral disinhibition can weaken pleasure with getting more than you deserve. *Journal of Personality and Social Psychology*, *101*(4), 791-811. doi:10.1037/a0023556

Van den Bos, K., Van Veldhuizen, T. S., & Au, A. K. C. (2015). Counter cross-cultural priming and relative deprivation: The role of individualism-collectivism. *Social Justice Research*, *28*, 52-23. doi:10.1007/s11211-014-0230-6

- Waldmann, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus throwing a person on a bomb. *Psychological Science*, *18*(3), 247-253. Doi: doi: 10.1111/j.1467-9280.2007.01884.x
- Webber, D., Schimel, J., Martens, A., Hayes, J., & Faucher, E. H. (2013). Using a bug-killing paradigm to understand how social validation and invalidation affect the distress of killing. *Personality and Social Psychology Bulletin*, *39*(4), 470-481. doi: 10.1177/0146167213477891
- Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *132*(2), 249-268. doi: 10.1037/0033-2909.132.2.249
- West, S., Aiken, L., & Krull, J. (1996). Experimental personality designs: Analyzing categorical by continuous variable interactions. *Journal of Personality*, *64*(1), 1-48. doi:10.1111/j.1467-6494.1996.tb00813.x
- Wiesenfeld, B., Brockner, J., & Martin, C. (1999). A self-affirmation analysis of survivors' reactions to unfair organizational downsizings. *Journal of Experimental Social Psychology*, *35*(5), 441-460. doi:10.1006/jesp.1999.1389
- Yeargain, J. W., & Kessler, L. L. (2010). Organizational hostility toward whistleblowers. *Journal of Legal, Ethical and Regulatory Issues*, *13*(1), 87-92.
- Zimbardo, P. G. (1973). On the ethics of intervention in human psychological research: With special reference to the Stanford prison experiment. *Cognition*, *2*(2), 243-256. doi: 10.1016/0010-0277(72)90014-5
- Zhong, C. B., & Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, *314*(5797), 254-254. doi: 10.1126/science.1130726
- Zhong, C. B., Liljenquist, K., & Cain, D. M. (2009). Moral self-regulation, licensing and compensation. In D. De Cremer (Ed.), *Psychological perspectives on ethical behavior and decision making*. (pp. 75-43). Charlotte, NC: Information Age Publishing, Inc.
- Zhong, C., Strejcek, B., & Sivanathan, N. (2010). A clean self can render harsh moral judgment. *Journal of Experimental Social Psychology*, *46*(5), 859-862. doi:10.1016/j.jesp.2010.04.003

1

2

3

4

5

6

&

Dutch Summary | Nederlandse Samenvatting

In dit sociaal-psychologische proefschrift onderzoek ik hoe mensen reageren op het morele gedrag van anderen. Ik focus op de reactie van mensen op moreel gemotiveerde deviantie, wat ik definieer als gedrag dat (1) wordt vertoond vanuit morele overwegingen en dat (2) afwijkt van de standaard. In vier empirische hoofdstukken laat ik zien hoe een confrontatie met een moreel gemotiveerde deviant evaluaties over deze deviant beïnvloedt, evenals gedrag ten opzichte van deze deviant. Daarnaast onderzoek ik hoe deze confrontatie het zelfbeeld en op zichzelf gericht gedrag beïnvloedt. Ten slotte identificeer ik verschillende moderatoren die beïnvloeden hoe mensen reageren op moreel gemotiveerde deviantie.

Mijn onderzoek begint met de observatie dat mensen soms negatief reageren op anderen die moreel gedrag laten zien (zie bijv., Minson & Monin, 2012; Monin, 2007; Parks & Stone, 2010). Zo laat recent onderzoek bijvoorbeeld zien dat mensen een negatief oordeel hebben over anderen die weigeren een racistische taak uit te voeren (Monin, Sawyer, & Marquez, 2008). Daarnaast kunnen mensen ook negatief oordelen over activisten die streven naar gelijkheid (Bashir, Lockwood, Chasteen, Nadolny, & Noyes, 2013) en over mensen die doneren aan goede doelen (Newman & Cain, 2014). Soms straffen we zelfs diegene die pro-sociaal gedrag laten zien (Barasch, Levine, Berman, & Small, 2014). Waarom gebeurt dit?

In Hoofdstuk 2 laat ik zien dat participanten die tijdens het onderzoek worst aten zich bedreigd voelden door een morele vegetariër die dit tijdens het experiment weigerde. Ook bleek dat participanten zichzelf negatiever evalueerden wanneer ze met zo'n morele vegetariër (in plaats van een niet-morele vegetariër) werden geconfronteerd. Het beschermen van het zelfbeeld door middel van lichamelijke reiniging—het reinigen van de handen—zorgde ervoor dat deze negatieve reacties niet optraden.

In Hoofdstuk 3 onderzoek ik een vergelijkbaar proces in een ander domein. Hier laat ik zien dat participanten die tijdens het onderzoek bier dronken, meer bier uitdeelden aan een andere participant die morele redenen had om geen bier te willen drinken dan aan een andere participant die niet-morele redenen had om geen bier te willen drinken. Daarnaast deden participanten zichzelf ook meer pijn—door hun hand langer ondergedompeld te houden in ijskoud water—wanneer ze met een morele weigeraar in plaats van een niet-morele weigeraar werden geconfronteerd. Wanneer participanten hun zelfbeeld hadden beschermd door middel van zelf-affirmatie (e.g., Cohen & Sherman, 2014), traden de negatieve effecten op het uitdelen van bier niet op.

1

2

3

4

5

6

&

1

In Hoofdstuk 4 toon ik aan dat mensen zelfs gedrag dat normaal gesproken als erg bewonderingswaardig wordt gezien kunnen afkeuren, mits dat wordt vertoond door een deviant uit hun eigen groep: Conservatieve Christenen hadden een negatief oordeel over de adoptie van een weeskindje wanneer de toekomstige adoptie-ouders lesbisch in plaats van hetero waren. De meest negatieve beoordeling werd echter gegeven wanneer het stel lesbisch was en zichzelf identificeerde als Christelijk.

2

3

4

5

6

&

In Hoofdstuk 5 onderzoek ik verschillende factoren die beïnvloeden welk effect de confrontatie met een moreel gemotiveerde deviant heeft op het zelfconcept. Deze factoren zijn bijvoorbeeld het actief deelnemen aan gedrag dat tegen iemands eigen waarden ingaat (of niet), de mate waarin iemands eigen gedrag zichtbaar is voor anderen, en de mate waarin de het morele aspect van iemands gedrag opvalt. Daarnaast laat ik ook zien dat mensen waardering kunnen opbrengen voor de moreel gemotiveerde weigering van anderen in situaties waarin de moraliteit van dit gedrag duidelijk is—namelijk wanneer het gedrag dat geweigerd wordt duidelijk tegen de waarden van de participanten ingaat.

Samenvattend, in dit proefschrift heb ik onderzocht wat er gebeurt met mensen op het moment dat zij in aanraking komen met het moreel gemotiveerde gedrag van anderen. Dit moreel gemotiveerde gedrag was bijvoorbeeld de beslissing om geen worst te eten of bier te drinken vanwege morele redenen of geen anti-homo betoog te willen schrijven uit morele overtuiging. Ik heb geprobeerd antwoord te geven op fundamentele vragen over hoe moraliteit functioneert en hoe dit mensen beïnvloedt in hun dagelijks leven. Ik heb daarbij gefocust op verschillende gedragingen in verschillende domeinen, zoals vlees eten, bier drinken, het bevorderen of afwijzen van gelijke rechten voor homoseksuelen en adoptie. Deze verschillende gedragingen beslaan een groot gebied aan verschillende inhoudelijke domeinen. Sommige van deze gedragingen worden wellicht als meer of minder moreel gezien dan anderen. Belangrijk is dat ik in dit proefschrift niet beargumenteer dat deze gedragingen moreel *zijn*. Ik onderzoek juist wat er gebeurt wanneer deze gedragingen worden bestempeld of gezien als moreel. Ik laat hiermee zien dat het label 'moreel' of 'immoreel' een sterke invloed heeft op het zelfconcept van mensen, hun evaluaties van anderen, en hun gedrag—onafhankelijk van de precieze inhoud van de gedragingen. De aversie die mensen ervaren om gedrag te laten zien dat door anderen als immoreel wordt bestempeld speelt hierin een belangrijke rol. Kortom, moraliteit reguleert de attitudes en het gedrag van mensen in de sociale wereld.

KLI dissertation series

The “Kurt Lewin Institute Dissertation Series” started in 1997. Since 2014 the following dissertations have been published in this series:

- | | | |
|----------|---|---|
| 2014-01: | Marijn Stok: Eating by the Norm: The Influence of Social Norms on Young People’s Eating Behavior | 1 |
| 2014-02: | Michèlle Bal: Making Sense of Injustice: Benign and Derogatory Reactions to Innocent Victims | 2 |
| 2014-03: | Nicoletta Dimitrova: Rethinking errors: How error-handling strategy affects our thoughts and others’ thoughts about us | 3 |
| 2014-04: | Namkje Koudenburg: Conversational Flow: The Emergence and Regulation of Solidarity through social interaction | 4 |
| 2014-05: | Thomas Sitser: Predicting sales performance: Strengthening the personality – job performance linkage | 5 |
| 2014-06: | Goda Perlaviciute: Goal-driven evaluations of sustainable products | 6 |
| 2014-07: | Said Shafa: In the eyes of others: The role of honor concerns in explaining and preventing insult-elicited aggression | & |
| 2014-08: | Felice van Nunspeet: Neural correlates of the motivation to be moral | |
| 2014-09: | Anne Fetsje Sluis: Towards a virtuous society: Virtues as potential instruments to enhance | |
| 2014-10: | Gerdien de Vries: Pitfalls in the Communication about CO2 Capture and Storage | |
| 2014-11: | Thecla Brakel: The effects of social comparison information on cancer survivors’ quality of life: A field-experimental intervention approach | |
| 2014-12: | Hans Marien: Understanding and Motivating Human Control: Outcome and Reward Information in Action | |
| 2014-13: | Daniel Alink: Public Trust: Expectancies, Beliefs, and Behavior | |
| 2014-14: | Linda Daphne Muusses: How Internet use may affect our relationships: Characteristics of Internet use and personal and relational wellbeing | |
| 2014-15: | Hillie Aaldering: Parochial and universal cooperation in intergroup conflicts | |
| 2014-16: | Martijn Keizer: Do norms matter? The role of normative considerations as predictors of pro-environmental behavior | |
| 2015-01: | Maartje Elshout: Vengeance | |
| 2015-02: | Seval Gündemir: The Minority Glass Ceiling Hypothesis: Exploring Reasons and Remedies for the Underrepresentation of Racial-ethnic Minorities in Leadership Positions | |

1

2015-03: Dagmar Beudeker: On regulatory focus and performance in organizational environments

2

2015-04: Charlotte Koot: Making up your mind about a complex technology: An investigation into factors that help or hinder the achievement of cognitive closure about CCS

3

2015-05: Marco van Bommel: The Reputable Bystander: The Role of Reputation in Activating or Deactivating Bystanders

4

2015-06: Kira O. McCabe: The Role of Personality in the Pursuit of Context-Specific Goals

5

2015-07: Wiebren Jansen: Social inclusion in diverse work settings

6

2015-08: Xiaoqian Li: As time goes by: Studies on the subjective perception of the speed by which time passes

&

2015-09: Aukje Verhoeven: Facilitating food-related planning. Applying metacognition, cue-monitoring, and implementation intentions

2015-10: Jasper de Groot: Chemosignaling Emotions: What a Smell can Tell

2015-11: Hedy Greijdanus: Intragroup Communication in Intergroup Conflict: Influences on Social Perception and Cognition

2015-12: Bart de Vos: Communicating Anger and Contempt in Intergroup Conflict: Exploring their Relational Functions

2015-13: Gerdientje Danner: Psychological Availability. How work experiences spill over into daily family interactions

2015-14: Hannah Nohlen: Solving ambivalence in context. The experience and resolution of attitudinal ambivalence

2015-15: Stacey Sanders: Unearthing the Moral Emotive Compass: Exploring the Paths to (Un)Ethical Leadership

2015-16: Marc Heerdink: Regulating deviance with emotions: Emotional expressions as signals of acceptance and rejection

2015-17: Danny Taufik: "Can you feel it" The role of feelings in explaining pro-environmental behavior

2015-18: Sarah Elbert: Auditory information and its parameters in health persuasion. The development of a tailored smartphone application to support behavior change

2016-01: Anna van 't Veer: Effortless morality — cognitive and affective processes in deception and its detection

2016-02: Thijs Bouman: Threat by association: How distant events can affect local intergroup relations

2016-03: Tim Theeboom: Workplace coaching: Processes and effects

2016-04:	Sabine Strofer: Deceptive intent: Physiological reactions in different interpersonal contexts	1
2016-05:	Caspar van Lissa: Exercising Empathy: The Role of Adolescents' Developing Empathy in Conflicts with Parents	2
2016-06:	Marlon Mooijman: On the determinants and consequences of punishment goals: The role of power, distrust, and rule compliance	3
2016-07:	Niels van Doesum: Social mindfulness	4
2016-08:	Leonie Venhoeven: A look on the bright side of an environmentally-friendly life: Whether and why acting environmentally-friendly can contribute to well-being	5
2016-09:	Florien Cramwinckel: The social dynamics of morality	6

&

