Moral Biases in Intergroup Context

Submitted by Esra Dasci to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Psychology

May 2019

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……..

Acknowledgements

I would like to start by thanking Turkish Ministry of Education for their generosity and support over the years. Without their support, this project would not exist at the first place.

To Manuela and Joe: Thank you for your time, energy, support and feedback over years. You both are very good scientists in your own ways, and I have learnt so much from you. I am very proud to be your student. I would like to give special thanks to Manuela, I truly appreciate your support and guidance throughout my PhD. There were times that I felt completely lost, but then your enthusiasm and encouragement motivated me to power through this long journey. Thank you for your time, availability, encouragement and support.

To Manuela's lab group, thank you for all the support and feedback you gave me. You are all great, inspiring researchers. I would like to give special thanks to Jolien who has been the best academic buddy to me. Thank you for your support, and for your help whenever I needed. I deeply appreciate all your contributions and inspiring ideas that made my work richer.

Merve, you are an amazing friend, I am very grateful that our ways are intersected at some point. Thank you for being there whenever I needed, for being my family here, for all the laughs and your sympathetic ear. I know we always have each other. To Melika, thank you for being a great, caring and supportive friend. I feel very lucky to have you here. To Afet, thank you for all those beautiful travels, long library nights but most importantly for your smiling face. To Esma, you are the best officemate and an amazing friend one can ever have. Thank you for all those laughs and beautiful memories.

Slieman, I knew that we'd be the best friends as soon as I met you. Thank you for all those beautiful memories and your beautiful, heart-warming smile. To Laura, my choir buddy, thank you for your kindness, inspiring chats and coffee mornings.

Tom, I didn't know that what we had started would grow into this *stupendous* experience. Thank you for being so caring and loving and supporting me in every possible way. You have been both the best friend and a great partner to me from the beginning. I am lucky to have you in my life.

To my family, you gave me the courage to start this journey. Thank you for your support and encouragement for all those years. I would never be able to do this PhD without your love, care and understanding. I will forever be grateful for you.

Abstract

Even though people perceive their moral judgements as objective and unbiased, research suggests that they in fact make biased moral decisions that are then *rationalized* as objective. This project aimed to build on this prior work to further investigate the effect of intergroup biases on judgments of others' morality. In the first line of this project, I aimed to investigate whether or not moral judgments are influenced by group membership where there is no prior information, beliefs, or emotions associated with the target group (i.e., in a minimal group setting). Participants read two versions of trolley dilemmas and judge either ingroup or outgroup targets. Across four experimental studies with minimal groups, I found evidence for group biases, even though the nature of patterns obtained varied. Overall, the findings suggest that group membership has a substantial influence on moral judgments even in minimal group settings, but given the inconsistency between the patterns obtained these warrant further investigation. Then, I aimed to investigate moral biases in pre-existing groups (specifically, men and women) and the role of relevant ideologies (sexism). I specifically focused on the role of a type of sexism that directly refers to women's morality, i.e., benevolent sexism. To further specify the role of sexism in moral judgements, I also varied the moral extremity of the action. The results indicated that benevolent sexism led to *more lenient* judgements of women who displayed *clearly immoral* actions, especially among male participants (Study 5). However, the interplay between benevolent sexism and moral extremity was not moderated by target gender, indicating that this effect is not limited to judgements of women's actions (Study 6). In the final study, I aimed to investigate the role of hostile sexism in counter-stereotypical female actions. The results revealed that hostile sexism negatively affected moral judgements of women who behaved counter-stereotypically, but not of men who displayed the same (stereotypically male) behaviour, but this was only the case for female participants.

Table of Contents

Index of Tables

Index of Figures

1 General Introduction

We come across moral judgements every day, in the media, in everyday conversation, among others. These judgements are sometimes implicit, a gut feeling about what is right and wrong, and often feel personal, but nevertheless fairly objective and shared with others. These seeming contradictions are at the core of what morality is: A deeply held set of rules often so self-evident that they seem objective, but are nevertheless intrinsically linked to identities that are shared with some, but not others. My focus in this thesis to examine group biases in moral judgements.

Research suggests that people automatically evaluate everything they perceive, including individuals and groups (Zajonc, 1980). Person and group perception have a significant impact on future interactions with those individuals and groups that are perceived (Boehm, 2008). For example, the evaluation of someone as untrustworthy reduces willingness to interact with that person (Bambilla, Sacchi, Pagliaro, & Ellemers, 2013). Person and group perceptions are however not necessarily factual or objective. Instead, they often involve affective input and motivations that can lead to biased perceptions. For example, a classic study found that people make more favourable judgments of others whom they expect to interact with than of people they do not expect to meet (Darley & Berscheid, 1967).

Just as person perception can be biased by preconceived beliefs and motivations, group perceptions are guided by pre-existing stereotypes. For example, the Stereotype Content Model (Fiske, Cuddy, Glick, & Xu, 2002) proposes that different outgroups are associated with specific stereotypes, which in turn drive emotions and influence behaviour, including moral judgments. For example, research has shown that the acceptability of sacrificing lives to save other lives depended on whether those sacrificed were outgroup members and those saved were ingroup members (Cikara, Farnsworth, Harris & Fiske, 2010).

In fact, a very recent study examining changes in moral values between 1900 and 2007 has shown a steady increase in ingroup-based morality with time (Wheeler, McGrath, & Haslam, 2019), underlining the timely need to improve our understanding of group biases in moral judgements.

In this thesis, I take on board existing knowledge about the flexibility of moral judgments, leading to biases that can have severe consequences for certain groups in society. My overall aim is to advance understanding of how moral judgements might be influenced by group biases and associated factors. To do so, I bring together insights from moral psychology with findings from social psychology. I started by examining group biases in moral judgments in a minimal group paradigm, to investigate group biases in groups that are not associated with specific meanings, histories, or ideologies. Then, I proceeded by examining group biases in real social groups (i.e., men and women) and whether associated ideologies (i.e., sexism) interfere with intergroup biases.

In Chapter 2, I review the literature on which my research was based. I start by defining morality and continue by discussing existing theory and evidence bearing on the aspects of morality I address in this thesis.

In Chapter 3, I report four studies examining the effect of group membership on moral judgments in a minimal group setting. I also examine the role of group identification and type of action (i.e., moral dilemma).

In Chapter 4, I report three studies examining whether sexism affects moral judgments of women as a function of the moral extremity of the action (i.e., how immoral the action is). I was specifically interested in investigating whether the role of benevolent sexist attitudes toward women can be triggered in actions that do not directly bear on the gender hierarchy.

Chapter 5 provides a summary of my findings in relation to the overarching research questions of this research. I conclude by discussing the limitations of this research, and directions for the future.

2 Literature Review

Research has shown that, though we evaluate others' actions in different domains such competence and warmth (Fiske et al., 2002), evaluations of people's moral virtue seem to be primary (Abele & Wojiszke, 2007; Brambilla & Leach, 2014; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Van Lange & De Bruin, 2000). Specifically, moral traits such as trust score highest among the characteristics people seek in others (Cottrell, Neuberg, & Li, 2007) and those perceived as moral are also perceived as most deserving (Brambilla & Riva, 2016). Of course all domains of impression formation play a role in our evaluations of a person and can be primary in particular situations, but individuals seem to seek and value most information about another person's morality (Brambilla & Leach, 2014). For example, being deficient in sociability might not lead to ostracism or punishment in a society whereas lacking in morality undoubtedly comes with consequences, ranging from ostracism to prison sentences (Opotow, 1990; Van der Lee, Ellemers, Scheepers, & Rutjens, 2017). In sum, morality is a critical domain of social judgement and one that can be very consequential. If so, it seems crucial to understand how morality functions and whether moral judgements can be biased, potentially leading to the biased and undeserved social spread of sanctions. But what exactly is morality?

2.1. What is Morality?

In everyday language, the word 'morality' broadly refers to a socially shared system of values and principles that distinguishes between what is right and wrong. As a system, morality can be defined as sets of rules, virtues, norms, practices, identities, institutions, and psychological mechanisms that work together to supress selfishness and maintain a cooperative and coherent social life (Haidt, 2008; Haidt & Kesebir, 2010). This set of moral norms is unique to each society and each member of a specific society is expected to live by these norms. Abidance to a society's moral system is monitored through regular evaluations

of individual actions, individuals, or groups—moral judgements (Haidt, 2001). When individuals or groups fail to meet these moral rules, they are condemned, criticized, or even excluded from the society. Moral guidelines impose a social contract that does not need to be formalized or written, but is often reflected and reinforced by formal procedures and institutions (legal, religious or political). Irrespective of whether or not they are formalized, the main function of these guidelines is the same: To supress selfishness and promote prosocial behaviour.

Selfishness and pro-sociality are understood by reference to how the self is defined (Tajfel, 1978; Tajfel, Billig, & Bundy, 1971). Within groups and societies, pro-sociality emerges from coordination within the group and selfishness is about the failure to do that. When more groups are taken into account, pro-sociality can be bounded by group lines, in that to favour a competitor is not likely to be perceived as moral (Levine, Prosser, Evans, & Reicher, 2005; Levine & Thomson, 2004)—though to help a more disadvantaged group might be (Smith, Aquino, Koleva, & Graham, 2014). In this way, moral guidelines serve to regulate the behaviour of individuals within their own group, but also towards other groups, in a complex and often nuanced system that prescribes what is right or wrong.

Some argue that there are universal moral principles that apply to all cultures, people and time periods. In this vein, cross-cultural differences in what is deemed moral are considered superficial, or not core to morality. These universalist approach seeks commonalities across cultures, which, in turn, suggests a relatively systematic and non-discriminative set of moral rules focused on judgements of harm and justice (Turiel, 1983). Specifically, those who hold a universalist approach propose that there is only one true moral guideline, which applies to everyone and to which everyone ought to adhere. Early moral development theories supported this universalist standpoint by suggesting that moral development is realised in pre-determined ordered stages that are equal for everyone

(Kohlberg, 1963). Although this approach acknowledged differences across cultures and individuals, it ranks those differences as developmental stages, adding up to only one superior sense of morality that applies to all people.

By contrast, relativists focus more on the ways in which moral principles vary across groups or societies. One of the first and loudest voices in opposition to universalist approaches was Carol Gilligan (1983), who suggested that following Kohlberg's theory women would be judged as less moral than men. In her book, "In a different voice", Gillian claims that Kohlberg's approach is 'male-centric' and fails to take into account women's values and priorities. Although Gillian's views have been criticized, and subsequent research could not find gender differences in moral reasoning, the work of scholars such as Gilligan have highlighted that moral judgements can be based on more than one type of consideration and can therefore vary across individuals and groups in meaningful ways that do not merely reflect developmental stages. Moral relativism (also called 'moral pluralism') therefore pointed to the importance of examining differences in what is deemed as moral, paving the way for more recent interest in the examination of moral reasoning in intergroup contexts (Graham & Haidt, 2010).

When individuals deviate from their society's mores, they are penalized harshly to maintain societal order and cohesion (Balliet, Mulder, Van Lange, 2011; Fehr & Gachter, 2002; Nowak & Sigmund, 2005; Wu, Balliet & Van Lange, 2016). However, there are times when members of the same society cannot agree on what qualifies as moral, or even on how to apply certain moral principles (Monin & Jordan, 2008). For example, members of religious sub-groups in a society might have different dress codes or dietary requirements that their followers ought to comply with. Although transgression of these rules might be deemed morally wrong by one sub-group, this might not be shared by other sub-groups in the same society. That is, moral guidelines serve their regulatory function by driving evaluations of

actions, individuals, and groups as moral or immoral, which in turn leads to sanctions for behaviour considered immoral (Ellemers & van den Bos, 2012; Sleebos, Ellemers, & de Glider, 2006). Deviance is thus defined by comparing individuals' behaviour against the group's moral guidelines (Pagliaro, Ellemers, & Barreto, 2011). It stands to reason that when individuals or groups hold different ideas about what qualifies as moral, then they also hold different beliefs about what qualifies as deviant.

How people make moral judgements and how these often converge but sometimes diverge has been the focus of a growing body of work. Though many psychological theories include aspects related to morality (e.g., social psychological theories linking norms to attitudes and behaviour), only a smaller set of theories is explicitly concerned with moral judgements. Theories have differed in their focus, a broad category of theories focusing more on the content or motives of moral judgements, while others focus on the routes through which people arrive at those judgements, and yet others at how these judgements pan out in and are affected by social contexts. In what follows, I briefly review some of the most important theories proposed to examine these issues and some key the evidence they have generated.

2.2. Moral Motives: Beyond Harm and Justice

As argued, early moral psychology research tended to treat morality as a one universal set of rules which can be applied to any moral situation in question. Universalist researchers developed their theories around the idea that morality concerns beliefs about harm and justice. However, it has been refuted by pluralistic and relativist researchers, such as Gillian, who indicated that there is more to morality than just these two concerns. Here I review some of the most important theories which extend morality beyond harm and justice.

One of the theories that aims to define the content or motives underlying moral judgements was developed by Shweder (1990). This author proposed that there are three

different ways of talking about morality that shape moral decision making, which are called as *the big three of ethical discourse*: 1) The ethics of autonomy, which taps into the concepts of harm, fairness, and justice and aims to protect the autonomy of individuals during their interpersonal relations; 2) The ethics of community, which is related to the concepts of duty, respect, and loyalty and aims to regulate the social order within a community, therefore enabling the groups to maintain their cohesion and coordination over time; and 3) The ethics of divinity is related to sanctity and sin and aims to preserve the purity of each individual. Shweder (1991) suggests that cultural norms and culturally shaped thoughts and emotions affect the moral decision making process. For example a study with Brazilian and US samples found that people from low vs. high socio-economic status used these three moral discourses differently when judging harmless but offensive actions (Haidt, Koller, & Dias, 1993). Similarly, another study indicated that college students relied more on the ethics of autonomy while older people were more willing to talk about issues related to the ethics of community and divinity (Arnett & Balle-Jensen, 1993). In this sense, what this theory suggests is that moral judgments are sensitive to cultural norms and cannot be evaluated without taking the cultural context into account.

A similar approach was developed by Rai and Fiske (2011) who conceptualized four moral motives regulating interpersonal relations. This theory argues that different kinds of social relationships entail different moral obligations and prohibitions that should be examined in order to understand moral judgments and moral disagreements. A. Fiske suggested that moral judgements are driven by different moral motives that are embedded in social-relational schemas, the activation of which can vary depending on the situation and the individual: 1) Unity, which is related to the issues aroused by the fact that people live in social groups and need to maintain cohesion within the group; 2) hierarchy, which is based on respect to authority and loyalty to higher status group members; 3) equality matching, which

promotes fairness and justice between individuals and between groups; and 4) market pricing, which is related to rewards and punishments based on a utilitarian cost and benefit calculation. In support of this theory, a series of experiments has shown that priming the social-relational schemas led to activation of certain moral motives. For example, priming group relationships as communal and sharing led to activation of unity related motives, while priming a relational frame based on authority led to activation of hierarchy motives (Rai, 2012). The theory argues that moral judgments are based on evaluations of others as prospects for social relations and tries to specifically explain the cases where an agent's intentions are not taken into account by an observer who evaluating their action. In this way, according to the theory, it could be possible that an action involving violence or unequal treatment might be deemed as morally right depending on which relational motives are activated and what is the relevant relational structure. Overall, although these motives are proposed to be universal, Fiske claims that there is a difference in when and how these motives are activated across cultures, religious groups, and ideologies.

Haidt and Joseph's (2004) moral foundations theory aims to synthesize and integrate different moral theories in a broader theory of morality. Moral foundations theory addresses cultural differences in morality by reference to the idea that morality is not just about judgements of harm and fairness, but there are various other foundations that morality is built on: 1) The first foundation is related to the avoidance of harm and promotes care and warmth towards others; 2) The second foundation is fairness, which refers to concerns about cheating and unfairness among people; 3) The third foundation is loyalty and focuses on concerns about on threads and challenges one's ingroup disrespectful behaviours and betrayal of authorities; 4) The fourth foundation is authority, which refers to concerns about hierarchy and respect for authorities 5) The fifth is purity and aims to prevent degradation and promote wholesomeness and sanctity of individuals in a society. Finally, 6) the sixth foundation is

liberty and refers to concerns about oppression and corruption related to abuse of power. Studies testing this model have shown that different groups (e.g., religious groups, national groups, and people with different political ideologies) utilize or prioritise different moral foundations when making their moral judgements. For example, a study comparing conservatives and liberals in a US sample, found that liberals prioritised harm and fairness foundations when making moral judgements about a range of political issues, whereas conservatives were more likely to use all five foundations when making these judgements (Graham, Haidt, & Nosek, 2009). Haidt and colleagues have suggested that moral foundation theory can explain how ideologies that are seen by some people as problematic can be perceived as perfectly moral by other people: It is not that some people are happy to be immoral, but that ideologies can shape what foundations are prioritised when making moral judgements. Interestingly, foundations such as loyalty can both be used as a basis to promote cohesion and a basis for racism, nationalism, and prejudice towards outgroups, and authority can justify respectful treatment, but also oppression (Graham & Haidt, 2012; Haidt & Graham, 2007).

2.3. Routes to Moral Judgements

A second category of theoretical approaches focuses on the routes through which moral judgements are made. The earliest attempts to explain moral judgments came from philosophers (like Hutcheson, Hume, Smith) who suggested that moral judgements are led by our inner moral sense and immediate emotional reactions to the eliciting situations. According to emotion philosophers, reason is the ability to calculate the odds and the best means to a preferred outcome. Therefore, it is an instrument to understand that certain actions might lead to certain moral or immoral results. However, moral judgments are themselves based on our own sentimental interests and passions.

Later, Hume, and other scholars who supported this idea, were refuted by rationalist philosophers (like Kant, Leibniz, Descartes) who suggested that although emotions aroused can guide our moral reasoning, in the end we draw moral judgements from rational evidence we gather from the environment. In this rational context, it is conceptualized that human morality is shaped by certain abstract rational moral principles that people ought to comply with. This approach to morality gained extensive support from the early psychology research (specifically cognitive developmental researchers like Kohlberg, Piaget, Turiel), where moral judgments are led by a rational process as consisting of deliberate reasoning leading to moral behaviours.

The most influential rationalist theory was developed by Kohlberg, who argued that human morality develops through stages that require certain cognitive abilities to be developed in order to go to the next stage (1969). This idea that cognitive and moral development progress in tandem suggests that human morality is led by deliberate and controlled processes. Therefore, moral judgments can be made in a similar way as we make judgements about the physical world, where things are either right or wrong. In Kohlberg's moral development theory, the last stage is justice, which is viewed as the ultimate virtue. Therefore, one cannot be truly moral unless they are proved to be just in terms of Kohlberg's moral theory.

Kohlberg's theory has been criticized due to its limited conceptualization of human morality, specifically disregarding the voice of female morality (Gilligan, 1983). Blasi (1983) also criticized Kohlberg's theory because of its neglect of the effect of the moral self and social dispositions when making moral judgments. What Kohlberg suggest is that there is a responsible self who is aware of moral principles and, therefore, acts upon them. However, Blasi proposes the Self Model of moral action which stresses the relevance of the moral self in moral judgements and actions. Blasi (1980) argued that human morality is more than

deontic morals (judgements and actions), but rather it is an interplay between controlled conscious and emotional unconscious psychological mechanisms shaped by experiences (Blasi, 1983). Blasi's self-model of moral action broadens the understanding of moral judgments and opens the way to exploring the alternative routes to moral judgments (Narvaez & Lapsley, 2009).

Following Blasi's idea, Haidt (2001) proposed the Social Intuitionist Model where moral judgments are product of both moral intuitions and moral reasoning. It has been argued that people always make their moral judgments based on their moral intuitions, which is a quick reaction to a morally eliciting situation. This can then be followed by the slower and more deliberate moral reasoning that rationalizes the moral judgement that was initially regulated by moral intuition. Haidt supported his model with empirical evidence showing that affective responses were a better predictor of moral judgments than reasoning. For example, a study examining the reactions to sexual moral issues which did not involve harm (homosexuality, incest and different form of masturbation) among liberals and conservatives found that initial affective reactions predicted the moral judgments better than the perceptions of harmfulness (Haidt & Hersh, 2001). Later, in another study (Björklund, Haidt, & Murphy, 2000) participants were asked to provide reasoning to their initial moral judgments of harmless but disgusting moral dilemmas. However, those who judged the scenarios as immoral were not able to explain their reasoning (moral dumbfounding) but they nevertheless did not change their initial moral judgments. Another study found that there was a disassociation between people's moral judgements and their justification of the judgment, indicating that people failed to justify their reasoning (Hauser, Cushman, Young, Jin, & Mikhail, 2007). This research indicates that moral judgements can be arrived at through 'gut feelings', that is, immediate and intuitive reactions to (im)moral situations, and that, at least sometimes, subsequent moral reasoning might fail to happen.

Additional evidence that moral judgements are not necessarily arrived at purely by reference to rational standards comes from research by Greene and colleagues (2001). Using MRI to track brain activity, these researchers found that certain moral dilemmas trigger activity in brain areas that are associated with social emotional processing while other types of moral dilemmas (and non-moral) trigger activity in brain areas that are associated with working memory controlled decision making. In addition to demonstrating that people can arrive at moral decisions through more emotional or through more rational routes, these findings suggest that one important factor that affects which route leads to moral decision making is the characteristic of the action being judged. That is, they suggested that characteristics of the action being judged can determine whether the moral judgement is arrived at through emotional or purely rational processes.

Green and colleagues further differentiated between personal and impersonal actions. An action is seen as personal if it meets these three criteria: 1) It is likely to cause serious bodily harm, 2) to a particular person, 3) in such a way that the harm does not come from the deflection of an existing threat onto a different party. If an action does not meet these criteria, it is seen as impersonal. In a series of follow up studies, Greene and colleagues (2001) manipulated the characteristics of the action to clarify how these affect moral decision making processes. Their research suggested that a personal-moral action is consistently perceived as less moral than an impersonal-moral (and non-moral). Studies have subsequently found that whether an action is personal or impersonal can trigger activity in different areas of the brain, suggesting that different processes are involved in moral decision making referring to these actions—personal actions being associated with emotional processes and impersonal actions being associated with more controlled and effortful processes.

On the basis of these findings, Greene (2009) developed the dual process model of moral judgements. This model describes how particular moral dilemmas involve different types of moral decision making processes. For instance, Greene (2009) describes the different processes involved in the moral judgements people make when considering different train track dilemmas, where the agent in the scenario has to decide on a trade-off between one person and five people on the train tracks. Moral psychology widely uses these trolley dilemmas that are easily modifiable to manipulate specific factors (e.g., factors regarding the action, or of the agent and or the victim) to test specific research questions (Greene, 2008; Mikhail, 2011). In one version of this scenario (footbridge), the agent tries to stop the train by shoving a fat man off the footbridge. Since this involves a direct intervention, it qualifies as a personal action and it has been shown to trigger judgements driven by emotional processes in the brain: Greene (2009) designates this judgement as *intuitive emotional processing*. However, in another version of this scenario (switch), the agent tries to stop the train by switching the train towards a track where there is only one person. Participants perceive this action as impersonal and therefore rate it as more moral than the footbridge scenario. This is designated as *controlled cognitive process.*

Despite the claims and evidence put forth by Greene and colleagues, Hauser (2006) argues that in any moral decision cognitive process take the lead, feeding the intuition and finally triggering emotions. Therefore, the model that Hauser (2006) suggests is not that different moral decisions are related to different (cognitive or emotional) moral processes, but that cognitive and emotional processes are inter-related and both contribute to moral decision making

2.4. The Social Context of Moral Judgements

The theories reviewed above focus on what types of beliefs drive moral judgements and the routes through which these operate. Research has also examined how morality works

in social context and how factors that characterise social contexts might affect moral judgements. For example, while many researchers have suggested that the most important function of morality is to promote prosocial behaviour, it is clear that this is bounded by social categorizations as people are more likely to show altruism towards ingroup members than towards outgroup members (e.g., Batson et al., 1995; Yamagishi & Mifune, 2008). In addition, what is considered pro-social is likely to differ across groups, or societies, because they are likely to prioritise different moral norms (Haidt, 2008; Sachdeva, Singh, & Medin, 2011).

Groups are regulated by a variety of norms, not all of which are deep moral convictions (Skitka, 2010) and the extent to which norms influence group members' behaviour relies on the extent to which they are seen as moral. For example, groups can have norms that regulate task performance, as well as norms that regulate cheating—and the latter is likely to be considered more of a moral issue. Research has shown that norms framed in moral terms are more powerful regulators of ingroup members' behaviour than norms framed in competence terms. For example, studies have shown that group members are more likely to follow ingroup norms framed in moral terms than in competence terms (Pagliaro, Ellemers, & Barreto, 2011) and that they are more protective of the ingroup's reputation when an ingroup member displays immoral than an incompetent behaviour (Marques, Yzerbyt & Leyens, 1988).

The reason why the moral framing of group norms is so potent is that people are motivated to see themselves and their group as highly moral (and more so than highly competent; Leach, Ellemers, & Barreto, 2008), both because morality is central to the self-concept and because they are likely to anticipate harsher sanctions when they behave immorally than when they behave incompetently (Pagliaro, Ellemers, & Barreto, 2011). In addition, morality is expected to be a more fixed or immutable characteristic than other traits

such as competence (Pagliaro, Ellemers, Barreto, & Di Cesare, 2016), motivating people to avoid being seen as immoral—since 'once immoral, always immoral'. Perhaps because of this, but also because another person's morality indicates whether they can be expected to be benevolent or threatening, research has shown that people prefer others who are seen as moral and avoid those who are perceived to lack moral qualities to a greater extent than those who lack sociability or competence (Brambilla & Leach, 2014; Brambilla, Sacchi, Menegatti, & Moscatelli, 2016). In addition, groups only regulate the behaviour of ingroup (vs. outgroup) members (Gollwitzer & Keller, 2010)—and indeed, people care more about what their ingroup deems moral than about what other groups deem moral (Ellemers & van den Bos, 2012).

Importantly, studies have also shown that, contrary to what has been argued by Greene (2008), but consistent with Haidt's ideas (Haidt, 2001; Haidt & Joseph, 2004), the very same action can be judged very differently depending on local norms (or cultures), who engages in that action, or who its target is. The next section expands on these biases in moral judgements.

2.5. Biases in Moral Judgements

All in all, the work reviewed in the previous section suggests that moral judgements are flexible and responsive to contextual features, rather that driven purely by rigid principles or particular characteristics of the action. That also suggests that, if this is so, then moral judgements are not in themselves necessarily moral—if we consider fairness central to our conceptualization of morality—but very open to group and ideological biases.

There is ample evidence showing how different types of intergroup biases can lead to unequal treatment based on social group membership (Batson et al., 1995; Blair, Judd, & Cheapleau, 2004; Correll, Park, Judd & Wittenbrink, 2007; Eberhardt, Davies, Purdie-Vaughns & Johnson, 2006; Mealy, Stephan, & Urrutia, 2007; Mifune, Hashimoto &

Yamagishi, 2010; Valdesolo, & DeSteno, 2007). A meta-analysis on intergroup bias by Bettencourt et al. (2001) showed that ingroup bias occurs even when people were merely categorized into different groups. However, the research on intergroup biases has mostly focused on sociability and competence judgements of different groups, or of individuals from different groups. In this body of work, it has been shown that people rate the ingroup as more sociable and competent than the outgroup. However, bias in morality ratings has not been examined, although there are studies that suggest that biases occur in morality domain as well. For example, one study showed that the acceptability of lying for Euro-Americans and Ecuadorians depended on what the intention of the liar was perceived to be (self-promoting vs. other-oriented), who was lying (ingroup vs. outgroup), to whom (ingroup vs. outgroup), and what the cultural orientation of the perceiver was (collectivist vs. individualistic) (Mealy, et al., 2007; see also Ellemers, et al., 2019; Haidt, 2001). Beside cultural context, experimental work by Valdesolo and DeSteno (2007) also showed that even in a minimal group setting the very same behaviour (an unequal distribution of resources) was judged more harshly when it was displayed by an outgroup member than when it was displayed by an ingroup member. Given the fact that moral judgements might have serious consequences for individuals and some groups, it is very important to examine how biases emerge in this domain.

Biases are not just driven by ingroup-outgroup differences, they can also be driven by specific political ideologies. For instance, endorsement of social dominance orientation is found to be related several different types of group based biases (e.g., racism, sexism, ageism. see Sidanus & Pratto, 1999). In a similar vein, when asked about the appropriateness of the killing of civilians during war, American participants' moral judgements depended on the group membership of the victim: Conservatives were more likely to endorse the killing of civilians when these were Iraqi than when they were American, whereas liberals were not

affected by the group membership of the victim (Uhlmann, Pizzaro, Tannenbaum, & Ditto, 2009). Research also suggests that the endorsement of patriarchal attitudes towards women can moderate the link between offender's gender and moral judgements. The next section elaborates on what is already known about this.

2.6. How Sexism Affects Moral Judgments

Undoubtedly, one expects that moral principles applied by a higher legal institution should not be affected by biases. However, there is ample evidence of institutional biases, as well as biases made by individuals representing legal authorities. For example, research has shown that stereotypes of Black targets as criminals drives bias against Black (vs. White) targets in shooting decisions (Correll, Park, Judd & Wittenbrink, 2007). Similarly, the more stereotypically Black a defendant was the more likely to be judged as deserving death penalty when the victim was White than (Eberhardt, Davies, Purdie-Vaughns & Johnson, 2006). Consistent with this evidence, Labour MP David Lammy (2017) recently pointed to the pervasiveness of racial discrimination against Black, Asian, and minority ethnic ('BAME') people in the British criminal justice system, highlighting how important it is to examine the relationship between prejudice and moral judgements in this national context too.

Prejudice and biases in moral judgements are not limited to race. Research also suggests that an offender's gender might affect sentencing judgements, though evidence from this line of research is mixed. The first account, drawn from Chivalry Theory, suggests that women are treated more leniently than men by the justice system (Doerner & Demuth, 2010; Koons-Witt, 2002; Tillyer, Hartley, & Ward, 2015). This theory suggests that women are less likely to be arrested or detained compared to men, for the same crime. By contrast, there are also accounts suggesting that women are treated more harshly than men by the justice system because punishment decisions are made within a male dominated criminal system (Chesney-Lind, 1999, Heidesohn, 1989). There is evidence in support of both theories but that the

specific pattern of bias depends on other factors. Experimental studies indicated that there might be several different factors moderating this link between gender and moral judgements, or punishment (Herzog & Oreg, 2008; Masser, Lee, & McKimmie, 2010). One factor is the gender stereotypicality of the female offender, or her behaviour, or the extent to which she (or her behaviour) conforms to traditional female stereotypes. Indeed, research found that when female offenders portrayed a more stereotypically female image (i.e., being a caregiver) they were given more lenient treatment than when they did not present themselves as feminine (Herzog & Oreg, 2008; Koons-Witt, 2002).

Sexism cam be defined as an ideology that aims to regulate women's actions in order to maintain the patriarchal system were men are granted higher status than women. Ambivalent Sexism Theory (Glick & Fiske, 1996) suggests that sexism can be expressed in different ways, rooted in the same underlying belief that women are inferior to men. Hostile sexism expresses this belief through antagonism towards women who challenge the status quo (Glick, Diebold, Bailey-Werner, & Zhu, 1997). Benevolent sexism, on the other hand, idealizes women as incompetent, but demure, domestic, and moral (Glick & Fiske, 1996, 1997, 2011; Glick, Diebold, Bailey-Werner, & Zhu, 1997; Glick et al., 2000; Jackman, 1999; Krys et al., 2018; Leach, Carraro, Garcia, & Kang, 2015; Shields, 2007). Importantly, in line with the descriptive and prescriptive function of gender stereotypes (Glick & Rudman, 2010), benevolent and hostile sexism work in tandem to regulate women's behaviour and bring it in line with traditional gender roles. Specifically, benevolent sexism idealizes traditional women and rewards women who behave traditionally, whereas hostile sexism punishes women who deviate from this traditional ideal (Glick & Fiske, 1996, 1997, 2011; Glick, Wilkerson & Cuffe, 2015; Wilkerson, 2014).

Research on sexism and moral judgments has suggested that endorsement of sexist attitudes predicts harsher judgments of women who do not comply with gender expectations.

For instance, a study by Sakalli-Ugurlu and Glick (2003) found that sexist beliefs (specifically benevolent sexism) predicted a harsher judgment of women involved in pre-marital sex. In a similar vein, a study examining the effect of sexist attitudes on judgements of rape victims found that endorsement of benevolent sexism was associated with victim blame when the victims were counter-stereotypical (Masser, Lee, & McKimmie, 2010). Aside from illustrating the effect of sexism on moral judgements, these findings suggest that benevolent sexism is particularly relevant in this context, which makes sense since it is benevolent sexism that refers to women's superior moral sensibility. That is, benevolent sexism idealizes women as morally pure and superior to men in moral sensibility, while hostile sexism does not speak to women's morality. It is therefore a form of sexism that inherently involves both the expectation that women are moral (descriptive stereotype) and the imperative that they should be so (prescriptive stereotype).

In summary, though research appears to support the idea that sexism moderates moral judgements of female targets, results are somewhat inconsistent and in need of further exploration. Besides, research examining the effect of benevolent sexism on moral judgements has mostly focused on sexual abuse and harassment, or actions clearly violating gender roles or traditional views of women as pure and demure (e.g., murder, pre-marital sex). These behaviours have direct bearings on the power relations between men and women (reflecting it and reinforcing it) and therefore also on the gender hierarchy. Would benevolent sexism have similar effects when the target action does not necessarily have such implications? When actions have a bearing on the gender hierarchy, the link between sexist beliefs and moral judgements seems relatively self-evident. I additionally propose that, since benevolent sexism idealizes women as morally pure and superior, any immoral action by a female might be, by definition, considered as breaching gender roles or traditional views of women. Evidence for this proposal is as yet lacking. In addition, it is as yet unclear whether

benevolent sexism shapes moral judgements by making it less hard to detect immorality in women, or by making immoral behaviour seem more problematic.

2.7. Overview of This Research

In summary, research on biases in moral judgments highlights the significance of the context of the action being judged (e.g., who the agent is, or who the victim is). However, it is also important to see whether these effects would be present in a minimal intergroup context where there is no prior information or no stereotype content present, so as to distinguish mere categorization effects from effects that are driven by historical relationships between groups, or associated ideologies. Therefore, in the first part of this thesis (Studies 1-4 reported in Chapter 3), I focus on whether moral judgments are moderated by group membership. In these studies, I also examine whether or not group identification plays a role in this link, drawing on ample evidence that group identification often moderates intergroup biases (Branscombe, Wann, Noel, & Coleman, 1993; Doosje, Branscombe, Spears, & Manstead, 1998; Doosje, Ellemers, & Spears, 1995; Ellemers, Spears, & Doosje, 1997; Jetten, Spears, & Manstead, 1997; Sidanus, Pratto, & Mitchell, 1994; Terry & Hogg, 1996; Turner, Hogg, Oakes, & Smith, 1984).

Moreover, the present research aims to bring social psychology research and moral psychology approaches together. As argued above, moral psychology research suggests that factors such as intention to harm, or direct (personal) intervention between the person who causes the harm and the victim, influence moral decision making. To test whether intergroup biases in moral judgments depend on characteristics of the action examined within moral psychology, Studies 1-4 use two trolley dilemmas that have been shown to vary in their permissibility: The switch (more permissible) and the footbridge (less permissible) scenarios. Based on ample research with minimal groups that has revealed ingroup biases in other

domains (Bettencourt, Kelly, Nancy, & Hume, 2001). I expect to find group biases in moral judgements in these scenarios, using minimal group categorizations.

Another goal of this research is to examine biases in moral judgements in real groups, where group categorization comes together with specific relationship patterns, meanings, and history. In Studies 5-7 (reported in Chapter 5), I focus on gender categorization (and specifically on men vs. women) and examine the extent to which moral judgments are influenced by target and perceiver gender, and by group-based ideologies, in this case, sexist attitudes towards women. I propose that since benevolent sexism idealizes women as morally pure and superior to men, immoral actions displayed by women would be perceived as violating traditional gender roles and therefore moral judgements of women should be affected by the extent to which perceivers endorse benevolent sexism. Though there is no research I can directly draw on to draw hypotheses, two possibilities seem consistent with existing knowledge. First, since benevolent sexism is associated with the idea that women are moral, endorsing benevolent sexism might lead perceivers to give female targets (compared to male targets) the benefit of the doubt, raising the threshold for perceiving them as immoral. By contrast, benevolent sexism is also associated with the strong prescription that women *ought to* be moral. As such, endorsement of benevolent sexism can be associated with harsher judgements of female targets who behave immorally. I propose that the factor moderating these contrasting predictions is the extent to which the action to be judged is clearly immoral. I expect that endorsement of benevolent sexism is likely to be associated with harsher judgements when the action is clearly immoral, but with more lenient judgements when the action is not so immoral.

These proposals will be examined in the following chapters and discussed in the last chapter of this thesis.

3 Moral Biases in Minimal Groups

Ingroup favoritism has been widely examined in social psychology (Billig & Tajfel, 1973; Cikara et al., 2010; Tajfel et al., 1971; Turner, 1999; Tajfel, 1974; Valdesolo & DeSteneo, 2007), showing that people judge their in-groups as more sociable and more competent than out-groups. Here, we focus on in-group bias in the domain of morality, which has recently been recognized as one of the most important dimensions for one's evaluations of their group (Abele & Wojiszke, 2007; Brambilla & Leach, 2014; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Van Lange & De Bruin, 2000). Therefore, here we aimed to investigate whether ingroup biases occur in moral judgements.

The present research aims to examine biases in moral judgements in specific moral dilemmas (the footbridge and the switch) that have been widely used by moral psychology to investigate moral decision-making. A dimension that has been identified as crucial for moral judgments in moral psychology is whether the action requires active intervention or not when one is making their moral decision. Research has shown that people find harmful actions more acceptable if they do not require active intervention—impersonal moral action than if they require an active intervention—personal moral action (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001).

Research using these scenarios indicated that actions bearing on moral judgements can be supported by different processes in the brain, depending on characteristics of the action itself (e.g., whether the action is personal or impersonal). These studies showed that the characteristic of the action can lead to more automatic or more controlled judgements (Greene et al., 2001). Therefore, moral psychology suggests that these dilemmas tap into specific moral reasoning processes with the implication that responses tend to be easily predicted from the scenario itself. For example, pushing someone off a footbridge to stop a train, thereby killing the patient but saving the lives of five people—an action that is highly

personal—was judged as less acceptable than switching a train track, without directly acting upon someone, but causing the train to kill them and miss five other people. While these two actions have the same consequences (one-to-five trade off), the first action involves a direct personal force by the agent, which triggers an emotional automatic response, and is therefore judged as less acceptable by most participants (Greene et al., 2001).

A study examining the biases driven by political ideologies in these moral dilemmas (Uhlmann, Pizarro, Tannenbaum, & Ditto, 2009) found that American participants' moral judgements depended on the group membership of the victim. Specifically, they found that Conservatives were more likely to endorse the killing of civilians when these were Iraqi than when they were American, whereas liberals were not affected by the group membership of the victim. Although the study by Uhlmann and colleagues (2009) focused on an important intergroup context, it did not take into account variables that can function as important moderators in these contexts. Specifically, research on intergroup relations has suggested that identification is a powerful moderator of group-based evaluations, so it is possible that it also functions as a moderator of moral judgements in intergroup contexts.

Social identities derived from social group memberships are an important part of one's self-concept. Because people want to maintain a positive self-concept, they are inclined to positively evaluate their social groups in comparison with an out-group, and this tendency, in part, underlies in-group bias (Tajfel, 1978; Tajfel & Turner, 2004). A meta-analysis on intergroup bias by Bettencourt et al. (2001) showed that ingroup bias occurs even when people were merely categorized into different groups. In relation to morality, there have been studies investigating the effect of group membership directly on moral judgements (morally acceptability of an action). For example, Valdesolo and DeSteno (2007) looked at moral judgements of an immoral action (distributing resources unfairly to their own group) as a

function of group membership. They found that people rated ingroup targets as more moral than outgroup targets.

However, there is also evidence that people judge *outgroups* more favorably than their own membership group (Marques, Yzerbyt, & Leyens, 1988). Research has found that actions performed by ingroup members can significantly affect the ingroup's reputation, leading group members to derogate ingroup members who behave in negative ways even more than they derogate an outgroup member who displays the same behavior (Marques & Yzerbyt, 1988; Marques, Yzerbyt, & Leyens, 1988). This might be seen as a strategic expression of ingroup favoritism since it helps the ingroup to boost their positive image by distancing itself from deviant ingroup members (Marquez, Abrams, & Serodio, 2001). Although these two strategies stem from the same underlying motivation (to protect positive evaluations of one's own group), there might be another factor which causes these different responses of group members to ingroup deviants. Social identity theory (Turner & Brown, 1978) suggests that the fate of the ingroup is more important to individual group members who have more strongly internalized the group as part of their social identity. i.e., those who identify highly with the group (Turner, 1999). Indeed, research has suggested that low identifiers behave differently from high identifiers (Branscombe, Wann, Noel, & Coleman, 1993; Doosje, Ellemers, & Spears, 1995; Ellemers, Spears, & Doosje, 1997; Jetten, Spears, & Manstead, 1997; Sidanus, Pratto, & Mitchell, 1994; Terry & Hogg, 1996). These studies mostly indicated that low identifiers were less likely to behave normatively than group members who were highly identified with the ingroup. Therefore, the present study aims to investigate the function of identification on moral judgement.

Our research therefore examines the role of group membership in moral judgements in intergroup contexts, for personal vs. impersonal moral actions, using the footbridge and the switch scenarios as examples of personal (reflecting automatic processes) and impersonal

(reflecting controlled processes) actions. Research on bias in other dimensions has suggested a tendency to favor ingroup over outgroup (Billig & Tajfel, 1973; Cikara et al., 2010; Tajfel et al., 1971; Turner, 1999; Tajfel, 1972; Valdesolo & DeSteneo, 2007). Therefore, we hypothesized that the ingroup target's action is judged as more moral than the same action by an outgroup target (Hypothesis 1) – main effect of group membership. Second, we aim to examine whether this occurs for different actions (e.g., personal and impersonal). We do not find sufficient ground in the literature to draw specific predictions about this, so this will remain exploratory. Third, given that ingroup identification has been found to moderate many intergroup biases, we will examine whether that is also the case with regard to moral judgements. We expect that the effect of target's group membership on moral judgement is likely to be moderated by group identification, so that ingroup bias is likely to be stronger among high identifiers than among low identifiers (Hypothesis 3). Finally, we have the more exploratory aim to examine whether evaluations of moral actions by members of the ingroup/outgroup spill over to evaluations of the morality of the respective group as a whole. Here we reported four studies carried out to test these hypotheses.

3.1. Study 1

3.1.1. Method

Design. The design of this study was 2 (Target's membership: Ingroup vs Outgroup) X 2 (Moral scenarios: Switch vs. Footbridge) with identification as an additional independent variable. Moral scenarios varied within-participants and target's membership varied between-participants.

Participants. We ran a priori power analysis with an alpha = .05, power = .80 and identifying a small to medium effect size (f = .15) to estimate sample size with GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). The estimated sample size was 90. Therefore, data was collected via an online survey of 91 participants (65 males, 26 females; M_{age} = 29.33, *SD*

= 8.46). They were recruited through prolific.ac, a UK based crowdsourcing platform that supports research. The study took an average of 13 minutes to complete and, in exchange of their participation in the study, participants were paid £1.25 at the end of survey.

Procedure and materials. Participants were invited to participate in a study designated "Dominant Perception Style and Social Judgments." Participants who agreed to participate accessed the study through a web link, read basic information about the study, and provided informed consent (see appendix A for the full wording of the survey). Next, groups were formed, so as to enable the manipulation of ingroup vs. outgroup target.

First**,** we used a Navon task to form two groups: Detailed and global perceivers. Participants were shown fourteen Navon (Navon, 1977) stimuli, which consist of large letters constructed from a series of small letters. In this task, participants indicated which letter they saw on the screen after seeing each stimulus. Upon completion of this task, all participants were categorized as "detailed perceivers" regardless of their performance in the task, but they were led to believe that the task has divided participants into global and detailed perceivers. Then, participants indicated their identification with their assigned group with four questions (Ellemers, Spears, & Doosje, 1999) asking to what degree they identify as a detailed perceiver, if they are glad to be identified as a detailed perceiver, if they feel strong ties with global/detailed perceivers, and if they see themselves as a detailed/global perceiver. Participants indicated their responses on a 7-point scale, from 1 (*not all*) to 7 (*very much*). The group identification scale formed by averaging responses on all four items proved reliable ($\alpha = .88$).

Then, all participants were asked to make social judgements of targets involved in two scenarios: First the switch and then the footbridge scenario. All participants read two scenarios, either both involving ingroup targets or both involving outgroup targets. After reading each scenario (for scenarios see Appendix A), participants judged the permissibility

of the action on a 7-point scale (from 1 *forbidden* to 7 *permissible*—adapted from Cushman et al., 2006). These judgements constituted our measure of moral judgements. In the original texts, enactors' names are given; however, in the present study just the initials of the names were given in order to avoid providing additional categorical information about the enactor, like gender, nationality, or ethnicity. In addition, the original scenarios did not include any information about the enactor, which we added to disguise the goals of the study.

At this point, for exploratory purposes, participants were presented with a list of nine positive traits, encompassing competence, morality, and sociability (from Leach, Ellemers, & Barreto, 2007), and asked to indicate to what degree the ingroup and outgroup are characterised by each trait. The scales proved reliable both for ingroup perceptions (morality: $\alpha = .88$; sociability: $\alpha = .90$; and competence: $\alpha = .84$) and for outgroup perceptions (morality: $\alpha = .87$; sociability: $\alpha = .88$; and competence: $\alpha = .88$). Please see Appendix B.1 for exploratory analysis on whether evaluations of moral actions by members of the ingroup/outgroup spill over to evaluations of the morality of the respective group as a whole.

3.1.2. Results

Moral Judgements. We broke down this interaction by scenario, to examine whether the predicted interaction between target and identification would be limited to one of the scenarios. To do this, we used Model 1 of the PROCESS macro created by Hayes (2016) entering identification as a continuous factor. Identification was mean-centred and target's group membership was coded as 0 for the outgroup target and 1 for the ingroup target.

Table 3.1. Summary of means, standard deviations and correlations for variables in Study

Measures	***M***	***SD***	**1**	**2**	**3**
1.Footbridge morality score	3.53	1.89	___		
2.Switch morality score	5.16	1.75	.48**	___	
3.Identification	4.62	1.38	-09	-.07	___

Note: ** indicates significance at the level 0.01 (2-tailed).

Switch scenario. We tested whether the effect of target's group membership (X) on moral judgements (Y) is moderated by group identification (M). The overall model was statistically significant, $F(3,87) = 3.52$, $p = .018$, R^2=.108. There was a negative and significant effect of group identification, $b = -.65$, $t(87) = -2.88$, $p = .005$. The effect of target's group membership was not significant, $b = .33$, $t(87)= .94$, $p = .352$. Finally, the model shows that the predicted interaction between group identification and target's group membership was significant for this scenario, $b= .83$, $t(87) = 3.03$, $p =. 003$(see Figure 3.1). For low identifiers, there was no significant relationship between moral judgements and target's membership, $b = .41$, $t(87) = 1.57$, $p = .120$, that is, no intergroup bias. Similarly, for average identifiers, there was no statistically significant relationship between moral judgements and target's membership, $b = -.17$, $t(87) = -.94$, $p = .352$. However, for high identifiers, there was a statistically significant negative association between moral judgements and target's membership, $b = -.74$, $t(87) = -2.85$, $p = .005$. Therefore, the results indicate that, when making moral judgements in the switch scenario, high identifiers judged the outgroup actors' action as *less* permissible than when the same action was performed by an ingroup member. Simple slope analysis also revealed that the effect of identification on moral judgements was significant for outgroup targets, $b = -.65$, $t(87) = -2.89$, $p = .005$, but not for ingroup targets, $b = .18$, $t(87) = 1.15$, $p = .253$.

Footbridge scenario. Moderation analysis was conducted in the same way as described for the switch scenario, now focusing only on judgements for the footbridge

scenario. For this scenario the overall model was not statistically significant, $F(3,87) = .27$, $p = .848$, $R^2 = .009$. There was no statistically significant difference between judging an ingroup target or an outgroup target, $b = .10$, $t(87) = .24$, $p = .808$, level of group identification had no effect on judgements, $b = -.15$, $t(87) = -.58$, $p = .562$, and there was no statistically significant interaction between group identification and target's membership, $b = .03$, $t(87) = .11$, $p = .911$. Our hypotheses were therefore not supported or this scenario.

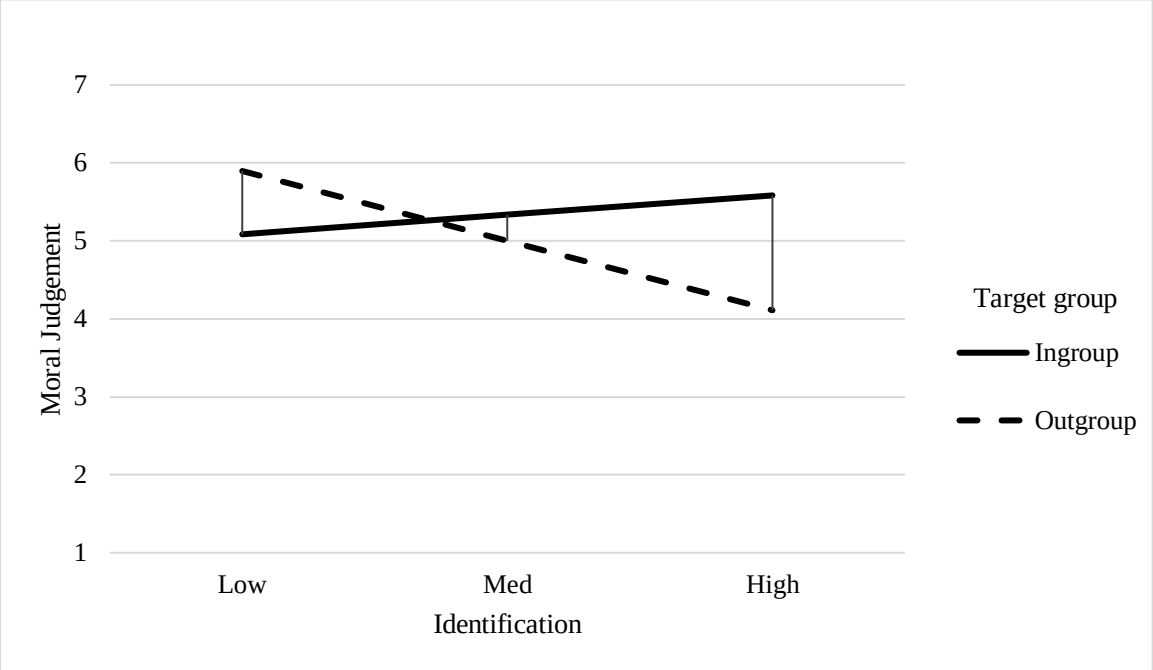

Figure 3.1. Moderation of the effect of target's group membership (X) on moral judgements (Y) by the levels of group identification (M) in the switch scenario in Study 1.

3.1.3. Discussion

This study provides some evidence for group biases in moral judgements. Specifically, we found that participants judged ingroup targets as more moral than outgroup targets, but only when ingroup identification was high and only for one of the scenarios investigated, i.e., the switch scenario. We did not find any significant effects of target group or identification on moral judgements on the footbridge scenario. This difference between the two scenarios can have multiple causes. First, it might be purely due to the fact that all participants in this study made their moral judgements first for the switch scenario and then

for the footbridge scenarios. It is possible that participants paid more attention to the first scenario, or allowed their bias to come through in that first scenario, losing interest in the second. This type of order effect is common (Wiegmann & Waldmann, 2013) and it will be further examined in the second study. However, note that participants could actually see both scenarios at the same time, so there is no guarantee they responded to the switch scenario first. Examining this by manipulating moral dilemma between participants will significantly contribute towards clarifying this effect.

A second and more theoretical possibility takes into account the distinction made by Greene et al. (2001) between the processes involved in making moral judgements in these two scenarios. If we take this perspective into account, one might propose that perhaps biases in moral judgements emerge mainly when processing is controlled, or that deontological judgements might not be so open to bias as those driven by controlled processes. While we are not ready to make such an inference—especially given the ample evidence for automatic biases in the extant literature—what our results do seem to suggest is that intergroup biases in moral judgements can be shaped by group membership and identification, but they can also ignore these factors. Finally, as in prior research, we found that actions in the switch scenario are perceived as more moral than actions in the footbridge scenario, irrespective of any other factor.

3.2. Study 2

Study 2 aimed to replicate and extend Study 1 in two ways. First, by varying scenario between participants, we aimed to avoid order effects and interference between the judgements of the two scenarios. Second, by adding a scenario that is also personal but that is typically regarded as permissible (because the patient is saved, rather that sacrificed), we aimed to further explore the role of scenario on moral judgements. In addition, unlike Study 1, to take into account the possibility that people might find one group label more attractive than the

other, we counterbalanced the group labels, by labelling half of the participants as detailed and the other half as global perceivers. We again hypothesized that moral judgements would be more favorable if the target was an ingroup than an outgroup member (Hypothesis 1), in particular for high identifiers (Hypothesis 2).

3.2.1. Method

Design. The design of this study was 2 (Target's membership: Ingroup vs Outgroup) X 3 (Moral scenario: Footbridge vs. Switch vs. Implied consent) between-participants factorial with identification as an additional continuous independent variable in the model.

Participants. A priori power analysis with an alpha = .05, power = .80, and identifying a small to medium effect size (calculated from effect size (η_p^2 = .07) of the significant three-way interaction between target's membership, identification, and moral scenario in Study 1 as f = .27) estimated 263 total sample size. Therefore, data was collected via an online survey of 279 participants (170 men, 105 women, 4 preferred not to disclose; M_{age} = 29.89, *SD* = 10.50). Participants were recruited through prolific.ac, a UK based crowdsourcing platform that supports research. The study took an average of 8.90 (*SD* = 8.25) minutes to complete and, in exchange for their participation in the study, participants were paid £1.00 at the end of the survey.

Procedure. Participants were invited to participate in a study designated "Dominant Perception Style and Social Judgments". Participants who had already completed Study 1 were not able to participate in this study because they were already familiar with the procedure. To do so, we blocked prolific ID's of those who already participated in the first study. Participants who agreed to participate accessed the study through a web link provided, read basic information about the study, and provided informed consent. The procedure for group formation and ingroup identification were done the same way in the previous study.

The identification was again found to be a reliable scale (α = .86). However, new to this study we added the Implied Consent scenario (see Appendix C for the scenario and the response scale used in this study) developed by Mikhail (2009). In this scenario the agent is causing bodily harm, but this results in a positive outcome for the victim. Therefore, although it can be described as personal, it tends to be seen as permissible (even obligatory). Thus, in this study, we changed the anchors for moral judgements since the act in the implied consent scenario might be considered as praiseworthy. Therefore, all participants were asked to make social judgements of actions by targets involved in one of three scenarios, which varied according to the experimental condition to which they were assigned, on a 7-point scale, from 1*"Forbidden,"* to 4 *"Permissible,"* to 7 *"Praiseworthy."* These judgements constituted our measure of moral judgements.

We measured ingroup and outgroup perceptions the same way we did in Study 1. Please see Appendix B.2 for exploratory analysis on group perceptions in Study 2.

3.2.2. Results

To test our predictions, we conducted a 2 (Target's membership: Ingroup vs Outgroup) X 3 (Moral scenario: Switch vs. Footbridge vs. Implied consent) between-participants ANCOVA with identification as an additional continuous predictor to test our predictions. Counterbalanced ingroup identity had no significant main or interaction effects, so to simplify the analyses reported here we exclude this factor.

Table 3.2. Summary of means, standard deviations and correlations for variables in Study 2.

Measures	*M*	*SD*	1	2	3	4
Footbridge morality score	2.85	1.80	__			
Switch morality score	4.95	1.48	*N/A*	__		
Implied consent morality score	6.08	1.26	*N/A*	*N/A*	__	
Identification	4.51	1.30	.17	-.042	.18	__

Most relevant to our hypotheses, the main effect of target's group membership was not statistically significant, $F(1,267) = 0.01$, $p = .918$, $\eta_p^2 < .001$. The interaction between the target's group and identification was also not significant, $F(1,267) = 0.08$, $p = .778$, $\eta_p^2 < .001$. Additional results revealed a significant effect of moral scenarios on moral judgements, $F(2,267) = 12.40$, $p < .001$, $\eta_p^2 = .09$, indicating that targets in the implied consent scenario were judged as more moral ($M = 6.07$, $SE = .17$) than those who in the switch ($M = 4.91$, $SE = .15$) and in the footbridge scenario ($M = 2.81$, $SE = .16$). There was a significant effect of group identification, $F(1,267) = 4.63$, $p = .032$, $\eta_p^2 = .02$ and a significant interaction between moral scenario and target's group membership, $F(2,267) = 3.41$, $p = .035$ $\eta_p^2 = .03$, indicating that outgroup targets were judged as more moral ($M = 3.22$, $SE = .23$) than the ingroup targets ($M = 2.40$, $SE = .22$) only in the footbridge ($p < .01$) but not in the switch (ingroup target: $M = 5.01$, $SE = .22$; outgroup target: $M = 4.81$, $SE = .22$, $p = .512$) and the implied consent scenarios (ingroup target: $M = 6.20$, $SE = .24$; outgroup target: $M = 5.93$, $SE = .24$, $p = .422$). There were no other significant main or interaction effects [interaction moral scenario X identification, $F(2,267) = 1.65$, $p = .194$, $\eta_p^2 = .01$; three-way interaction, $F(2,267) = 1.93$, $p = .147$, $\eta_p^2 = .01$].

3.2.3. Discussion

This study provided evidence for ougroup bias in moral judgements. Unlike Study 1, we did not find group identification to moderate this effect, which contradicts Hypothesis 2. Importantly, this time we found this effect only in the footbridge scenario. Although it is

unclear why the condition where bias is revealed differs between the two studies, what is clear is that both studies show intergroup bias in moral judgements. In addition, our results show that these biases can emerge both in the switch (Study 1) and in the footbridge (Study 2) scenario—and again, judgements of the switch scenario were more positive than judgements of the footbridge scenario. While we found an ingroup bias in the switch scenario, where the agent's action was judged to be more permissible (among high identifiers), we found outgroup bias in a clearly immoral scenario in Study 2. Indeed, there is research showing that people are stricter to in-group members than outgroup members when they misbehave (Gino, Gu, & Zhong, 2009). Therefore, this result might be seen as providing support for the idea that group biases are designed to protect the moral reputation of the ingroup in both cases.

Besides, if judgements in these scenarios indeed differ in the extent to which they are more automatic or more deliberate, this would be consistent with research on other types of intergroup biases, which shows intergroup biases both at the more explicit and the more implicit levels (Dovidio, Brigham, Johnson, & Gaertner, 1996; Dovidio, Kawakami, & Beach, 2003).

Why might the two studies differ with regard to the moral dilemma that elicited intergroup bias? It is possible that this is a result of the methodological differences between these studies. Specifically, Study 1 varied the scenarios within participants, with participants being able to see both scenarios at the same time, and Study 2 varied the scenarios between participants. This means that judgements in Study 1 were made in a context allowing for comparisons or management of biases between scenarios, whereas judgements in Study 2 were not. To further examine this, we conducted a third study where we replicate the design of Study 1, but this time manipulate scenario between (rather than within) participants.

3.3. Study 3

In this study, we aimed to replicate Study 1 with a between-participants design where participants judged either an ingroup or an outgroup target either in the Switch or the Footbridge scenario. We again expected to observe ingroup bias (i.e., more positive judgements of the ingroup than of the outgroup target; Hypothesis 1). We again tested whether group identification moderated this effect and, despite the results of study 2, given past evidence for the role of group identification in group biases, we predicted that the intergroup bias would be stronger for high identifiers (Hypothesis 2). Given the results of Study 2, which used a between participants design, we further expected an interaction between target group, identification, and moral dilemma, so that the bias would be stronger for high identifiers and in the Switch rather than the Footbridge scenario (Hypothesis 3).

3.3.1. Method

Design. The design of this study was 2 (Target's membership: Ingroup vs Outgroup) X 2 (Moral scenarios: Switch vs. Footbridge) between-participants with identification as an additional continuous independent variable.

Participants. A priori power analysis with an alpha = .05, power = .80 identifying a and identifying a small to medium effect size (calculated from effect size (η_p^2 = .07) of the significant three-way interaction between target's membership, identification, and moral scenario in Study 1 as f = .27) indicated a total sample size of 107. Therefore, data was collected via an online survey of 129 participants (62 males, 67 females; M_{age} = 35.12, *SD* = 12.06). They were recruited through prolific.ac, a UK based crowdsourcing platform that supports research. The study took an average of 8.21 (*SD* = 2.81) minutes to complete and, in exchange of their participation in the study, participants were paid £1.00 at the end of survey.

Procedure. The procedure for group formation and group identification was done as in the previous studies. The group identification scale was found to be reliable (α = .90). The procedure was the same as in Study 1 except that scenario variable varied between participants, therefore each participant judged either an ingroup or an outgroup target either in switch or footbridge scenario. The response scale for moral judgments was the same as in the Study 1, from 1 (*Forbidden*) to 7 (*Permissible*). Participants who took part in the previous studies were unable to participate to this study. Please see Appendix A for full wording of the surveys used in Studies 1-4.

3.3.2. Results

To test our hypotheses, we conducted a regression analysis entering target group (1 = ingroup; 0 = outgroup), moral dilemma (1 = switch; 0 = footbridge), and group identification (continuous, centred) in step 1, the two-way interaction terms in step 2, and the three-way interaction term in step 3 (see Table 3.2).

Table 3.3. Summary of means, standard deviations and correlations for variables in Study 3.

Measures	***M***	***SD***	**1**	**2**	**3**
1. Footbridge morality score	2.19	1.53	__		
2. Switch morality score	4.71	1.84	*N/A*	__	
3. Identification	4.60	1.37	-.06	-.11	__

Relevant to hypothesis 1, we did not find overall ingroup favoritism, as the main effect of target's group membership was not significant, b = -.31, $t(125)$ = -1.03, p = .304. Hypothesis 2 was also not supported, since the interaction between target's group membership and identification was not significant either, b = -.24, $t(121)$ = -.75, p = .460. Finally, Hypothesis 3 also received no support, as the three-way interaction between target's group membership, moral scenario and identification was also not significant, b = -.29, $t(121)$ = -.64, p = .525.

Table 3.4. Regression analysis predicting moral judgements in Study 3

Predictor	*b*	*b* 95% CI [LL, UL]	*beta*	*r*	Fit	Difference
(Intercept)	2.32	[1.80, 2.85]				
Target Group	-0.31	[-0.91, 0.29]	-0.07	-.05		
Moral Scenario	2.59**	[1.99, 3.19]	0.62	-60**		
Identification	-0.13	[-0.35, 0.10]	-0.08	.04		
					R^2 = .370** 90% CI [0.26, 0.47]	
(Intercept)	2.28	[1.66, 2.89]				
Target Group	-0.27	[-1.12, 0.59]	-.06	-.05		
Moral Scenario	2.62**	[1.74, 3.52]	0.62	-60**		
Identification	0.11	[-0.28, 0.51]	0.03	.04		
Target Group x Moral Scenario	-0.06	[-1.28, 1.16]	-0.01	.32**		
Target Group x Identification	-0.38	[-0.83, 0.08]	-0.12	-.05		
Moral Scenario x Identification	-0.04	[-0.48, 0.41]	0.06	.01		
					R^2 = .385** 90% CI [0.28, 0.49]	ΔR^2 = .015 90% CI [-0.02, 0.05]
(Intercept)	2.27	[1.66, 2.89]				
Target Group	-0.23	[-1.10, 0.63]	-.06	-.05		
Moral Scenario	2.60**	[1.70, 3.49]	0.62	-60**		
Identification	0.04	[-0.41, 0.49]	0.03	.04		
Target Group x Moral Scenario	-0.05	[-1.27, 1.17]	-0.01	.32**		
Target Group x Identification	-0.24	[-0.87, 0.39]	-0.12	-.05		
Moral Scenario x Identification	0.14	[-0.56, 0.83]	0.06	.01		
Target Group x Moral Scenario x Identification	-0.29	[-1.20, 0.62]	-0.11	-.10		
					R^2 = .387** 90% CI [0.28, 0.49]	ΔR^2 = .002 90% CI [-0.01, 0.01]

Note: A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

There was a main effect of moral dilemma, $b = 2.59$, $t(125) = 8. .52$, $p < .001$, indicating that participants judged the switch scenario as more moral than the footbridge, as in Studies 1 and 2. No other effects were significant: Main effect of identification, $b = -.13$, $t(125) = -1.11$, $p = .270$; two-way interaction between target's group membership and moral scenario, $b = -.05$, $t(121) = -.08$, $p = .94$; interaction between moral scenario and identification, $b = .14$, $t(121) = .39$, $p = .700$.

Given the effects found in Studies 1 and 2, we chose to continue by probing the interaction between target group membership and identification for each scenario. To do so, we conducted moderation analysis by using PROCESS macro model 3 (Hayes, 2016), to see if the relationship between target's group membership (X) and moral judgements (Y) was moderated by identification level (M) and moral scenario (W).

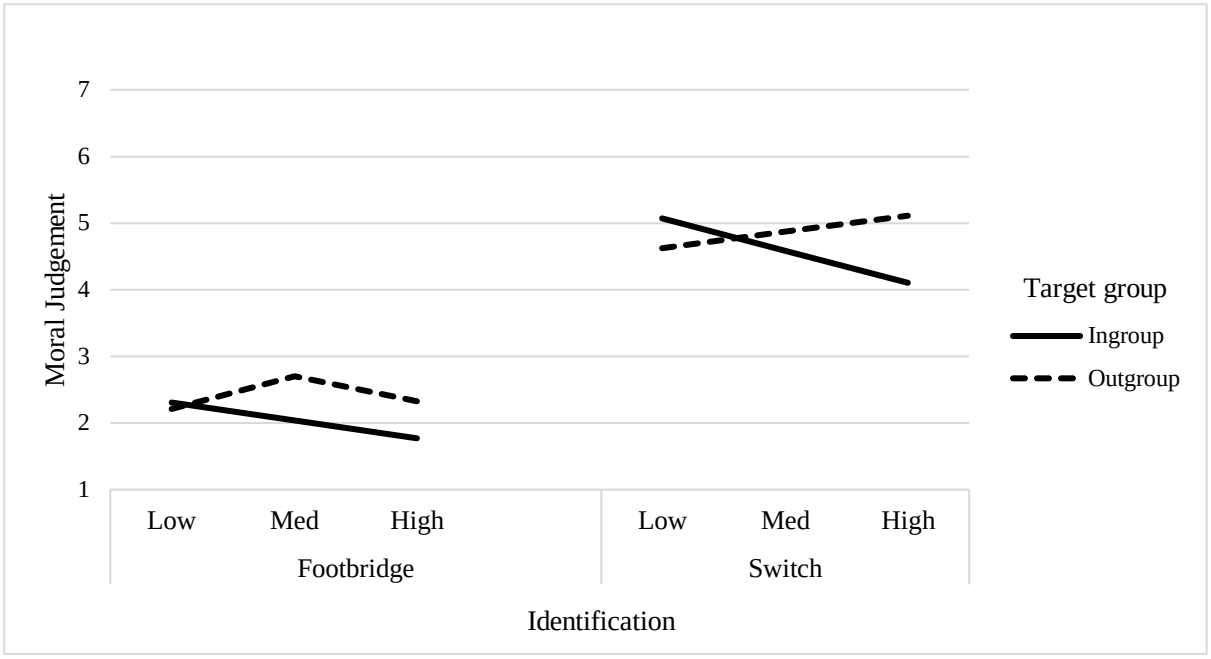

Figure 3.2. Moral judgement by target's group membership and group identification: Footbridge and Switch scenarios in Study 3.

The analyses revealed the same effects already reported. Looking at conditional effects, only the conditional effect of target's group membership for high identifiers in the switch scenario approached significance, b = -1.01, $t(121)$ = -1.82, p = .07. As displayed in Figure 3.2, participants evaluated the outgroup target as more moral than the ingroup target, but only in the switch dilemma and when they were highly identified with the group (see Figure 2 for patterns in the Footbridge dilemma).

3.3.3. Discussion

In sum, although we did not find any support for our specific hypotheses, the results suggest an outgroup bias in moral judgements that is moderated by ingroup identification and specific to the switch scenario. Unfortunately, the results are not sufficiently strong to allow

us to draw clear conclusions. We decided to conduct one more study to once again test these effects, as well as to compare moral judgements made in isolation, to those made across multiple moral dilemmas.

3.4. Study 4

In the research reported in this chapter we used within-participants' manipulations, as well as between-participants' manipulations of moral dilemmas. The results are inconsistent and those that appear to converge are not strong. The aim of Study 4 is to once again test these effects, but also to examine whether moral judgements made when different behaviors are compared (as those made with the within participants design of Study 1) reveal stronger biases in the Switch dilemma than when only one behavior is being judged at the time (as with the between participants design employed in Study 3). This could be the case because there is evidence in support of the idea that biases are more likely to be expressed when actions are less extreme (and the behavior in the Footbridge scenario is clearly more frowned upon) and providing a comparison between more and less extreme behavior might facilitate the expression of bias.

To keep the design manageable, this time we only examine judgements of outgroup targets, so we cannot test ingroup bias, as such, but only the extent to which the manipulations affect judgements of outgroup targets. As in studies 1 and 3, we expected moral judgements of outgroup targets to be affected by group identification, with evaluations of outgroup targets becoming less positive as group identification increased (Hypothesis 1). Given the results of Studies 1 and 3, we expected to find this effect to be stronger in the Switch scenario than in the Footbridge scenario (Hypothesis 2), and stronger in the Switch scenario when it was judged at the same time as the Footbridge scenario than when it was judged in isolation (Hypothesis 3).

3.4.1. Method

Design. The study compares moral judgements when participants judged an action either only in the footbridge dilemma, or in the switch dilemma, to when participants judged actions both in the footbridge and the switch dilemmas. Therefore, we had three different groups: Switch only vs. Footbridge only vs. both. Again, moral judgements constituted our dependent variable. Identification was again entered as a continuous variable.

Participants. A priori power analysis to estimate sample size with GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), with an alpha = .05, power = .95, and identifying a medium to large effect size (f = .27) indicated a total sample size of 175. Therefore, data was collected via an online survey of 198 participants; however, 15 participants were excluded due to absent moral judgement scores. Therefore, our analyses include 183 participants (85 males, 97 females, 1 prefer not to disclose; M_{age} = 33.11, *SD* = 21.92). They were recruited through prolific.ac, a UK based crowdsourcing platform that supports research. The study took an average of 8.17 minutes (*SD* = 3.28) to complete and, in exchange of their participation in the study, participants were paid £1.00 at the end of survey.

Procedure and materials. Participants were invited to participate in a study designated "Dominant Perception Style and Social Judgments." Participants who agreed to participate accessed the study through a web link provided, read basic information about the study, and provided informed consent except those who had already taken part in previous studies. Next, minimal groups were formed, so as to enable the manipulation of ingroup vs. outgroup target. Minimal groups were formed in the same way as in the prior studies. Group identification was also measured in the same way as before (α = .88).

Unique to this study, participants were presented with either only the footbridge dilemma, or only the switch dilemma, or both the footbridge and the switch dilemmas

(participants were able to see both scenarios at the same time as in Study 1), depending on condition. Please see Appendix A for the full wording of the survey used in Studies 1-4.

3.4.2. Results

To test whether identification affected moral judgements, we started by examining the correlation between identification and moral judgements in the different scenarios. As displayed in Table 3.5, identification was not significantly associated with moral judgements in either scenario. Next, we aimed to explore the differences between judgments made singly or in combination and whether the link between identification and moral judgements depended on the comparative context in which these were made (i.e., whether scenarios were presented individually or in combination).

Table 3.5. Summary of means, standard deviations and correlations for variables in Study 4.

Measures	***M***	***SD***	**1**	**2**	**3**	**4**	**5**
1.Footbridge morality score[a]	2.46	1.48	__				
2.Footbridge morality score[b]	2.26	1.58	*N/A*	__			
3.Switch morality score[a]	5.31	1.48	*N/A*	*N/A*	__		
4.Switch morality score[b]	4.48	1.88	*N/A*	.36**	*N/A*	__	
5.Identification	4.56	1.25	.06	.06	.012	-.02	__

Note: * indicates significance at the level 0.05, **indicates significance at the level 0.01 (2-tailed).

To test this, we first created two new variables comparing moral judgements in different contexts. The first variable (MJf) compares moral judgements in the footbridge and the switch where they were judged singly vs. moral judgement in the *footbridge* where it was judged in the context of the switch. The other (designated as MJs) compares moral judgements in the footbridge and the switch where they were judged singly vs. moral judgement in the *switch* where it was judged in the context of the footbridge.

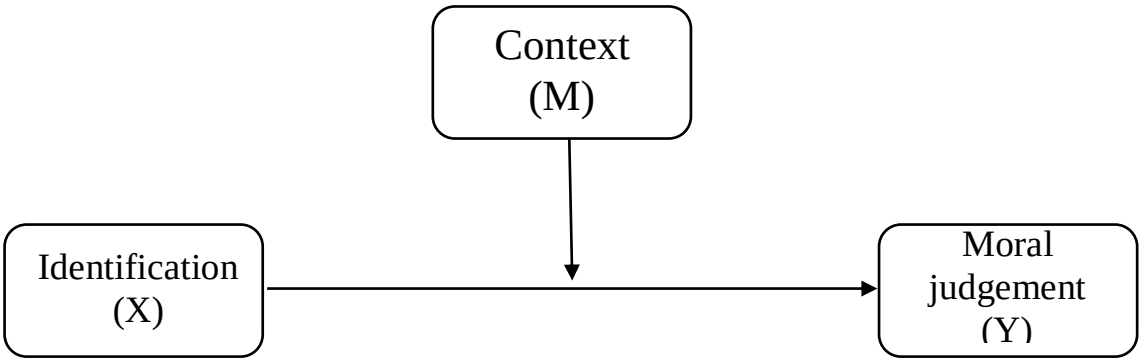

Figure 3.3. The moderation effect of evaluative context (M) on the relationship between ingroup identification (X) and moral judgement of an outgroup target (Y) in Study 4.

To test whether the direct effect of identification (X) on moral judgements (Y) was moderated by the context (M), we ran moderation analyses by using the PROCESS macro model 1 (Hayes, 2016; see Figure 3). This was done twice, once focusing on the footbridge comparisons (MJf) as the outcome variable and one focusing on the switch comparisons (MJs) as the outcome variable. Also, our moderator was dummy coded since it had three levels (see Table 3.3). Our baseline group for dummy coding was the one where participants judged both scenarios (mixed). Thus, D1 contrasts judgements in mixed (0) with judgements in the footbridge (1) scenario. D2 contrasts judgements in the mixed (0) with judgements in the switch (1) scenario. Additionally, we computed two new variables to represent the interaction between identification and the second dummy variable and entered the corresponding dummy and the interactions in the analyses as covariates.

Table 3.6. Dummy coding for the context (M) variable in Study 4.

	D1	D2
Footbridge single	1	0
Switch single	0	1
Mixed	0	0

Moral judgements in the footbridge (MJf) as the outcome. The first analyses focused on moral judgements in footbridge as the outcome variable (D1: footbridge vs. mixed), with D2 and identification X D2 as covariates. This analysis revealed a significant

moderation model, $F(5,177) = 30.70$, $p < .001$, R^2=.46. The direct effect of identification was not statistically significant, $b = .07$, $t(177) = .65$, $p = .51$. Also, the direct effect of D1 (footbridge vs. footbridge in mixed) was not statistically significant, $b = .21$, $t(177) = .75$, $p = .45$ indicating that judgements of the target in the footbridge scenario were unaffected by whether these were made in isolation or in combination with judgements in the switch scenario. The interaction between identification and D1 was also not significant, $b = -.01$, $t(177) = -.05$, $p = .96$, indicating identification did not affect moral judgements in the footbridge scenario more when these were made in isolation or together with the switch scenario. However, the direct effect of D2 (switch vs. mixed) on moral judgements was positive and significant, $b = 3.33$, $t(177) = 3.03$, $p = .003$, indicating that moral judgements in the switch scenario when it is judged singly, were more positive than moral judgements in the footbridge in the mixed condition. The interaction between identification and D2 was not statistically significant, $b = -.06$, $t(177) = -.26$, $p = .79$.

Next, we entered D2 (switch vs. mixed) into the analysis as our moderator, and D1 and id*D1 as covariates. The rest was the same as above. The overall moderation model was statistically significant, $F(5,177) = 30.70$, $p < .001$, R^2=.46. The direct effect of identification on moral judgement was not statistically significant, $b = .06$, $t(177) = .48$, $p = .63$. The direct effect of D2 on moral judgement was positive and significant, $b = 3.05$, $t(177) = 11.06$, $p < .001$, indicating that moral judgements in the switch when it was judged singly were more positive higher than moral judgments in the footbridge when it was in the mixed condition. The interaction between identification and D2 was not statistically significant, $b = -.06$, $t(177) = -.26$, $p = .79$. The statistics for D1 and the interaction between identification and D1 was the same as when D1 was the moderator (see above).

In sum, these analyses indicate that participants provided higher morality ratings of the target in the switch scenario than of the target in the footbridge scenario, when the latter

was judged together with the switch. This was unaffected by identification. Importantly, judgements in the footbridge scenario were no different when made singly or in combination with judgements in the switch scenario.

Moral judgements in the switch (MJs) as the outcome. We ran the same analyses with moral judgements in the switch as the outcome variable (D1: footbridge vs. mixed), with D2 and identification X D2 as covariates. The overall model was statistically significant, $F(5,177) = 19.68$, $p < .001$, R^2=.36. The direct effect of identification on moral judgement was not significant, $b = .007$, $t(177) = .06$, $p = .95$. The direct effect of D1 (footbridge vs. mixed) was negative and significant, $b = -2.01$, $t(177) = -6.78$, $p < .001$, indicating that the agent in the footbridge was judged to be less moral (when it was judged singly) than the agent in the switch (when it was judged after the footbridge). The interaction between identification and D1 was not statistically significant, $b = .09$, $t(177) = .39$, $p = .70$. The direct effect of D2 (switch vs. mixed) was also not significant, $b = .66$, $t(177) = .56$, $p = .58$. Finally, the interaction between identification and D2 was not significant, $b = .04$, $t(177) = .15$, $p = .88$.

Next, we entered D2 (switch vs. mixed) in to analyses as our moderator, and D1 and identification X D1 as covariates. The overall model was statistically significant, $F(5,177) = 19.68$, $p < .001$, R^2=.36. The direct effect of identification on moral judgement was not statistically significant, $b = -.01$, $t(177) = -.08$, $p = .94$. The direct effect of D2 on moral judgement was positive and statistically significant, $b = .83$, $t(177) = 2.82$, $p = .005$, indicating that the agent in the switch scenario was judged to be statically more moral when it was judged singly than when the agent in the switch was judged after the footbridge (mixed condition). Nevertheless, the interaction between identification and D2 was not statistically significant, $b = .04$, $t(177) = .15$, $p = .88$. The statistics for D1 and the interaction between identification and D1 was the same as when D1 was the moderator (see above).

In sum, evaluations of moral actions in the switch scenario were affected by whether it was presented singly or in combination with the footbridge scenario, with the actor being judged as more moral when the switch scenario was evaluated on its own than when it was evaluated in combination with the footbridge scenario. Again, the target in the switch scenario was always evaluated as more moral than the actor in the footbridge scenario, irrespective of whether the switch scenario was presented on its own or in combination with the footbridge scenario. Group identification played no role in these evaluations.

3.4.3. Discussion

The aims of this study were twofold: To again test the link between group identification and moral judgements and to examine whether moral evaluations on the two dilemmas would be affected by making single versus repeated judgements. In contrast to our prior results, we did not find any significant effect involving group identification in this study.

As in the prior studies and extant research, we found that moral judgements in the switch scenario were more positive than in the footbridge scenario. Most importantly for this study, we found that this was the case irrespective of whether the switch scenario was presented on its own or with the footbridge scenario. In addition, moral judgements in the footbridge scenario were unaffected by whether it was presented on its own or together with the switch scenario. By contrast, moral judgements in the switch scenario were more positive when it was presented on its own than when it was presented with the footbridge scenario. This is not what we expected, based on a comparison of the findings of Studies 1 and 3, but that comparison was imperfect, as it relied on effects of identification.

While the absence of effects of identification means that these results do not clarify the patterns of our earlier studies, they are interesting in their own right. Specifically, they

add to the evidence of the prior studies by directly showing that moral judgements are context dependent—not only do they vary depending on the action, but also depending on whether more than one action is being judged. Exactly why this happens we cannot say with certainty, but future research might wish to focus on understanding this in more detail.

3.5. Conclusion

The research reported in this chapter aimed to examine ingroup biases in moral judgements in experimentally created groups. We specifically aimed to investigate whether moral judgments would be influenced by group membership where there is no prior information, beliefs, or emotions associated with the target group. As moral psychology literature suggested (Greene et al., 2001), we varied the type of action being judged, to investigate whether biases were more likely to emerge for actions that are typically considered more or less permissible. Finally, we also examined whether group identification moderated these effects. We expected that participants would judge the ingroup target as more moral than the outgroup target and this would be the case particularly for high identifiers.

In Study 1, we varied moral scenario variable within participants: each participant read two versions of trolley dilemma (switch and footbridge) where the target was either an ingroup or an outgroup target. We found that there is an ingroup bias in switch scenario (more permissible) particularly for high identifiers. However, there was no bias in footbridge scenario. The results indicated that ingroup biases on moral judgments would emerge even in minimal group contexts. This also suggested that the type of action being judged and the group identification moderated the group biases in moral judgments.

In Study 2, we aimed to examine whether these findings could be replicated. However, moral scenario was varied between participants in order to avoid interference

between the moral judgments of two scenarios. In this study, we also added another moral scenario that is rated as heroic to further investigate the effect of moral scenario on moral judgements. To allow participants indicate their moral responses we ranged the response scale from forbidden to permissible to praiseworthy. The results did not replicate those of Study 1. However, they again revealed group biases in moral judgments. Specifically, we found outgroup bias in the footbridge scenario where the agent shoves a man off the footbridge to save five other people.

We conducted another study to clarify the inconsistency between Studies 1 and 2. This study also followed a between participants design as in Study 2, with the same two trolley dilemmas as in Study 1. The response scale was the same as Study 1. The results did not support our hypothesis. However, there was outgroup bias in the switch scenario only for high identifiers when we further analysed the interaction between target group membership and identification for each scenario. This direction was opposite to what we found in prior studies.

We conducted a final study to examine whether this inconsistencies across the three studies could be explained by methodological differences. In Study 4, we manipulated the context where moral judgments were made, specifically varying whether participants made isolated judgments of one target in one scenario, or of multiple targets in multiple scenarios (i.e., within versus between participants). We did not find effect of identification in this study. However, the results indicated that context had an influence on moral judgments. Specifically, we found that the switch scenario was judged more positively when it was presented on its own than when it was presented with the footbridge scenario. However, there was no effect of context on moral judgments in the footbridge scenario across conditions. Therefore, this study suggested that moral judgments are context dependent. That is not only

do they vary depending on the action, but also depending on whether more than one action being judged.

In sum, identification moderated ingroup bias in Study 1, but did not emerge as a significant moderator in Studies 2 and 3. Again, Study 4 did not reveal any main effect or moderation of identification. We thus provide very little evidence for the role of identification in moral biases, in experimentally created groups.

Even though we found mixed and somewhat inconclusive evidence for group biases in moral judgments, these studies suggest that group membership has a substantial effect on moral judgments even in minimal group settings where there is no prior beliefs, thoughts or emotions regarding the target group.

4 Moral Bias as a Function of Sexist Ideology

Studies on moral decision-making have suggested that people's moral judgements of a particular action might be affected by various factors, e.g. characteristics of the action itself, the decision maker's group, political ideology, and religious views (Ellemers, von der Toorn, Paunov, & van Leeuwen, 2019; Iyer, Koleva, Graham, Ditto, & Haidt, 2012; Uhlmann, Pizarro, & Ditto, 2009; Valdesolo & DeSteno, 2007). More importantly, the identity of the agent and the victim of the action has been shown to have a substantial impact on moral judgements (Genezy, 2005; Mealy, Stephan & Urrutia, 2007; Mifune, Hashimoto & Yamagishi, 2010; Valdesolo, & DeSteno, 2007; Steinel & de Dreu, 2004). In this sense, the moral principles applied by the decision-maker are to some extent flexible and depend on the context of the action being judged.

If moral judgements are flexible, it becomes possible to conceive that they might also be biased. When biases shape moral judgements these, in turn, can form the basis of discriminatory treatment, such as unequal penalties (Albonetti, 1997; Cameron, Payne, & Knobe, 2010; Steffenameier, Painter-Davis, & Ulmer, 2017). These biases can also emerge when the action in question is a criminal act (Barkan, & Cohn, 1994; Blair, Judd, & Cheapleau, 2004; Daly & Tonry, 1997; Tonry & Melewski, 2008). Undoubtedly, one expects that every individual should be treated equally in the justice system, regardless of their demographic characteristics. However, there is evidence showing that discrimination can influence criminal sentencing. For instance, a study by Blair, Judd, and Cheapleau (2004) focused on whether feature-based stereotyping might play a role in criminal sentencing decisions of Black and White defendants. The results showed that across both Black and White race categories, those with more Afrocentric facial features faced harsher sentences than those with less Afrocentric features. One might think that the White-Black polarization is a particularly problematic issue in the US, but this is not accurate. For example, recently

Labour MP David Lammy (2017) pointed to the pervasiveness of racial discrimination against Black, Asian, and minority ethnic ('BAME') people in the British criminal justice system, highlighting how important it is to examine the relationship between prejudice and moral judgements in other national contexts too. Indeed, a recent project on racial bias, led by The Guardian (Booth, Mohdin & Levett, 2019), has shown that racial bias towards BAME people is also a widespread issue in Britain. To illustrate the issue, findings from the project indicated that BAME respondents (47%) were almost two times more likely to be treated like a shoplifter in a shop when they hadn't done anything wrong compared to white respondents (22%).

The relationship between prejudice and moral judgements is not limited to race. There has been evidence from research on gender and criminology suggesting that offender's gender plays a role on sentencing decisions (Doerner & Demuth, 2010; Koons-Witt, 2002; Tillyer, Hartley, & Ward, 2015; Steffensmier & Demuth, 2006; Steffensmier, Kramer, & Streifel, 1993). Indeed, research has also suggested that sexism is linked to moral judgements. Research examining the relationship between sexism and moral judgements has mostly focused on judgements of victims (Masser, Lee, & McKimmie, 2010) and perpetrators (Viki, Abrams, & Masser, 2004) of sexual abuse, and on attitudes toward women engaged in premarital sex (Sakalli-Ugurlu & Glick, 2003). The general finding is that sexist beliefs are associated with harsher treatment of victims and more lenient treatment of perpetrators in the context of sexual crimes.

The existing research suggests that it is particularly associated with moral judgments of women. Therefore, we are particularly interested in whether *benevolent* sexist beliefs moderate target gender effects on moral judgements. Benevolent sexism is rooted in ideas about the complementarity of men and women, with men being afforded power and competence, and women idealized as demure, domestic, and moral (Glick & Fiske, 1996,

1997, 2011; Glick, Diebold, Bailey-Werner, & Zhu, 1997; Glick et al., 2000; Jackman, 1999; Krys et al., 2018; Leach, Carraro, Garcia, & Kang, 2015; Shields, 2007). It is therefore a form of sexism that inherently involves both the expectation that women are moral (descriptive stereotype) and the imperative that they should be so (prescriptive stereotype).

Existing evidence for the link between benevolent sexism and moral judgements is, however, quite mixed. In one study, Herzog and Oreg (2008) showed that endorsement of benevolent sexism was linked to more lenient punishment of a female offender. This was, however, only the case when the female offender behaved consistently with gender roles—when the offender was a woman who behaved non-traditionally, benevolent sexism did not play a role in the suggested punishment of the offender (while hostile sexism predicted harsher punishment). A different study, however, found that benevolent sexism was negatively related to evaluations of a woman whose behaviour could be seen as violating traditional gender roles (whereas hostile sexism was not; Viki, Massey, & Masser, 2005). In sum, though research appears to support the idea that sexism moderates moral judgements of female targets, results are somewhat inconsistent and in need of further exploration.

Research in this area has tended to focus on how sexism might affect judgements of behaviours such as sexual abuse, or behaviour that violates traditional gender roles. These behaviours have direct bearings on the power relations between men and women (reflecting it and reinforcing it) and therefore also on the gender hierarchy. Would benevolent sexism have similar effects when the target action does not necessarily have such implications? When actions have a bearing on the gender hierarchy, the immorality of the action is in itself subjective and shaped by sexist beliefs. Since benevolent sexism idealizes women as morally pure and superior, any immoral action by a female might be, by definition, considered as breaching gender roles. However, it is as yet unclear whether benevolent sexism shapes

moral judgements by making it less hard to detect immorality in women, or by making immoral behaviour seem more problematic.

The first goal of the studies reported here is therefore to clarify the role of benevolent sexism in moral judgements by examining whether benevolent sexism also moderates moral judgements of women's actions when these do not have direct bearings on the gender hierarchy. The second goal is to shed light on how benevolent sexism might play this role, by investigating whether benevolent sexism affects moral judgements by raising the threshold for identifying women's actions as immoral or by leading to harsher judgements of immoral actions enacted by women. To examine this, we varied whether moral judgements of actions by women were affected by the interplay between benevolent sexism and the moral extremity of the action (i.e., how immoral it is). Indeed, one of the ways in which studies with inconsistent results for the effects of sexism on moral judgements have differed is precisely the moral extremity of the action being judged (from stealing to murder; Herzog & Oreg, 2008; Viki, Massey, & Masser, 2005). We expected that high endorsement of benevolent sexism would be more likely to be associated with a higher threshold for identifying immoral behaviour when this behaviour is morally ambiguous, as this allows for sexist beliefs to shape whether or not the behaviour is detected as immoral. By contrast, we expected benevolent sexism to be associated with harsher moral judgements when the behaviour is clearly immoral. In line with this reasoning, we predicted an interaction between benevolent sexism and moral extremity so that benevolent sexism would have opposite effects on moral judgements of female targets, depending on the moral extremity of the action.

Overview of the research

We report three studies. In Study 1, we examine how benevolent sexism affects moral judgements of female targets when actions were morally ambiguously vs. clearly immoral. Study 2, maintains this design, but adds a manipulation of the gender of the target, to

examine whether or not the effect is restricted to female targets. In Study 3, we introduce an action that contradicts gender roles, to examine whether the effects of benevolent sexism, participant gender, and target gender can be replicated in this context.

4.1. Pilot Study

The central aim of Studies 1 and 2 is to examine how benevolent sexism affects moral judgments of women in contexts where the action does not directly bear on the gender hierarchy. That is, Studies 1 and 2 use scenarios that do not reflect the power relationships between men and women. In addition, we aimed to vary the moral extremity of the action, so we conducted a pilot study to select two hypothetical scenarios that were perceived as varying in the extent to which they were seen as moral (ambiguous vs. clearly immoral). We also assessed how certain participants were of their moral judgements to explore whether or not this was related to morality perception, and how typical the action was of men or women, to control for gender stereotypicality. We aimed to select two scenarios that varied in moral extremity, but not in gender stereo-typicality.

We used Prolific academic to recruit 40 participants (15 males, 24 females, 1 preferred not to say; M_{age} = 30.60, *SD* = 10.06, all resident in the UK) for an online survey. Participants who agreed to participate accessed the study through a web link provided, read basic information about the study, and provided informed consent.

Next, participants read four scenarios and answered some questions about each scenario. There were two pairs of scenarios (see appendix D for the full wording of the scenarios). Two scenarios described a target crashing their car because they were driving without their glasses (scenario 1; clearly immoral) or because they were in a hurry to help someone get to the hospital (scenario 2; morally ambiguous). A second pair described a target stealing medication from a pharmacy because they were addicted to drugs (scenario 3; clearly immoral) or because they were terminally ill (scenario 4; morally ambiguous).

After each scenario, participants indicated the extent to which they thought the action was acceptable (from 1 *Not at all* to 7 *Very much*), the extent to which they thought that the target was moral (from 1 *Not at all* to 7 *Very much*), and the extent to which they were sure about their judgements (from 1 *Not at all* to 7 *Very much*). Finally, participants indicated whether they thought that the target in the scenario was more likely to be a man, more likely a woman or equally likely to be either (from 1 *Strongly disagree* to 7 *Strongly agree*). Participants where then thanked for their participation in the study and received £1 as a token of appreciation.

4.1.1. Results and Discussion

As can be seen in Table 4.4, one pair of scenarios differs in the ideal way for our purposes: The two scenarios involving the car crash (scenarios 1 and 2) differ in the extent to which the action was perceived as moral, but not in the extent to which the actor is expected to be more typically male or female (note that the scenarios were perceived as slightly more typically female than male). The scenarios do not differ in how certain participants were of their judgements. Scenario 2 was judged as clearly immoral (see Table 4.4 and note that SDs place all participants' judgements on the immoral side of the scale) whereas judgements of scenario 1 were more ambiguous, as they ranged from slightly immoral to slightly moral, around the mid-point of the scale—though all participants were equally certain of their moral judgements, the scenarios themselves varied in how stable those moral judgements were across participants. Since this is precisely the difference we wanted to create, we chose to proceed by using these two scenarios involving the car crash in the subsequent studies.

Table 4.7. Means and standard deviations for the car crash scenarios (left panel) and stealing medication scenario (right panel) in Pilot Study.

	Scenario 1	**Scenario 2**	**Scenario 3**	**Scenario 4**
Measures	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Acceptability of Action	5.28 (1.36)[a]	2.43 (1.28)[b*]	4.30 (1.71)[a]	1.98 (1.14)[b*]
Certainty of Judgement	5.53 (1.40)	5.77 (1.03)	5.07 (1.44)[a]	5.90 (1.32)[b*]
Morality of Target	5.90 (1.01)[a]	3.50 (1.45)[b*]	4.60 (1.22)[a]	2.70 (1.18)[b*]
Certainty of Judgement	5.83 (1.17)	5.35 (1.25)	4.98 (1.49)	5.45 (1.41)
Target Gender				
Likely Man	3.98 (1.35)	3.85 (1.56)	4.68 (1.36)	4.53 (1.40)
Likely Woman	4.10 (1.45)	3.93 (1.75)	3.05 (1.11)	3.15 (1.31)
Equally likely either	5.55 (1.81)	5.45 (1.96)	4.63 (2.09)	4.60 (2.23)

Note: Means with * differ from the mid-point of the scale (4) and means with different subscripts within panels differ with $p < .05$.

4.2. Study 5

In this study, we examined whether benevolent sexism predicted moral judgements of female targets differently depending on whether the action was morally ambiguous or immoral. We hypothesised that a female target displaying an ambiguous action (crashing into someone when helping another) would be perceived as more moral than a female target committing an immoral action (crashing someone because of not wearing their glasses; Hypothesis 1). Most importantly, we also anticipated an interaction between moral extremity and benevolent sexism, so that higher endorsement of benevolent sexism would be associated with judgements of the female target as more moral when the action was morally ambiguous (Hypothesis 2a), but with judgments of the female target as less moral when the action was clearly immoral (Hypothesis 2b). We included both male and female participants and additionally explored whether or not participant gender affected moral judgements.

4.2.1. Method

Design. The study had a 2 (Moral extremity: Morally ambiguous vs. Clearly immoral) X 2 (Participant's gender: Female vs. Male) between participants experimental design. All targets were female. In addition, participants' benevolent sexism score was measured on a continuous scale, and added to the model as a group-mean centred moderator. Hostile sexism[1] was measured but not included in the analyses. Our dependent variable was moral judgement of the action.

Participants. A priori power analysis, using G*Power (Faul, et al., 2007), based on the two-way interaction between moral extremity and benevolent sexism with a small to medium effect size f = .15, alpha = .05 and power = .80 suggested a total sample size of 432 was needed. Therefore, data was collected via an online survey from 438 participants (234 females, 204 males; M_{age} = 31.94, *SD* = 13.23) who were recruited through prolific.ac, a UK based crowdsourcing platform. In exchange for their participation in the study, participants were paid £1 at the end of the survey. Participants who participated in this study did not participate in studies 5 or 7.

Procedure and materials. Participants were invited to participate in a study designated as "Social Judgements and Relationships." Participants who agreed to participate accessed the study through a web link, read basic information about the study, and provided informed consent. Next, participants were assigned to one of the two experimental conditions, where they read a scenario with a female target committing either a morally ambiguous action or a clearly immoral action (see Appendix E.1 for full wording of scenarios and measures). After reading the hypothetical scenario, they were asked to judge the morality of the action, the target's moral character, and their overall impression of the target. Judgement of the action was measured with two items asking to what extent the action was

[1] See appendix F.1 for regression analysis with hostile sexism added as a moderator in the model in Study 5.

"acceptable" and "moral" on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*). Evaluations of the target's character, were assessed by asking participants to indicate to what extent they thought that the target was moral, reliable, and trustworthy with three different items on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*). Finally, participants indicated their overall impressions of the target on a 7-point scale, from 1 (*Very negative*) to 7 (*Very positive*). Though these different measures are often examined separately, we chose to collapse our three dependent variables into one morality score because they were highly correlated ($rs \geq .62$, $ps \leq .01$), and together constituted a highly reliable morality scale ($\alpha = .95$).

At this point, participants completed the Ambivalent Sexism Inventory developed by Glick and Fiske (1996). Participants rated to what extent they agree with 22 items in the ASI on a 6-point Likert scale, from 0 (*Strongly disagree*) to 5 (*Strongly agree*). The benevolent sexism scale was reliable ($\alpha = .86$) and unaffected by the manipulation of moral extremity, $F(1,434) = .06$, $p = .81$. However, the effect of participant gender on benevolent sexism was significant, $F(1,437) = 21.92$, $p < .001$, indicating that male participants ($M = 2.01$, $SD = .06$) scored higher on benevolent sexism than female participants ($M = 1.59$, $SD = .06$) did. The interaction between the two did not affect benevolent sexism scores, $F(1,437) = 2.48$, $p = .12$. Before finishing the study, participants were asked to provide some demographic information (age, gender and ethnicity). Finally, they were informed about the actual aims of the research, fully debriefed, and paid for their participation in the survey.

4.2.2. Results

We conducted a hierarchical regression analysis in which all predictors were entered in the first step, their two-way interactions in the second step, and the three-way interaction in the third step (see Table 4.5). Moral extremity was coded as 1 = ambiguously immoral, 0 = clearly immoral and participant's gender coded as 1 = female, 0 = male. Benevolent sexism

scores were group mean-centred, where we subtracted the mean within the corresponding participant gender group from their individual benevolent sexism scores.

Table 4.8. Summary of means, standard deviations and correlations for variables in Study 5.

Measures	***M***	***SD***	**1**	**2**	**3**
1. Moral Judgement	3.96	1.50	__		
2. Benevolent Sexism	1.78	0.96	.07	__	
3. Hostile Sexism	1.94	1.11	-.48	.40**	__

Note: * indicates significance at the level 0.05, **indicates significance at the level 0.01 (2-tailed).

In line with Hypothesis 1, the main effect of moral extremity was positive and significant, $b = 2.13$, $t(430) = 20.85$, $p < .001$, suggesting targets were judged as more moral when the action was morally ambiguous than when it was clearly immoral. The results also revealed a significant main effect of benevolent sexism on moral judgment of female targets that was positive and significant, $b = .13$, $t(430) = 2.42$, $p = .016$, suggesting more positive moral judgements of the female target as benevolent sexism increased. The interaction between moral extremity and benevolent sexism on moral judgements was negative and significant, $b = -.25$, $t(430) = -2.26$, $p = .024$. Simple slope analysis were computed through SPSS. Higher endorsement of benevolent sexism indicates +1SD above the mean and low endorsement of benevolent sexism indicates -1SD below the mean. Simple slopes suggest that higher endorsement of benevolent sexism was positively associated with judgements of female targets as more moral when the action was clearly immoral (contrary to Hypothesis 2b), $b = .26$, $t(434) = 3.35$, $p < .001$, but not when the action was morally ambiguous (Hypothesis 2a), $b = .01$, $t(434) = .15$, $p = .884$. While Hypotheses 2a and 2b predicted an interaction between moral extremity and benevolent sexism, this is not the predicted pattern.

Aside from these, we also found a negative and significant interaction between participant's gender and benevolent sexism, $b = -.22$, $t(430) = -2.02$, $p = .044$, with simple slope analyses showing that endorsement of benevolent sexism was positively associated

with male participants' moral judgments, b = .32, $t(434)$ = 2.99, p =.003, but not with female participants' moral judgements, b = -.10, $t(434)$ = -.88, p = .378.

Table 4.9. Regression analysis predicting moral judgements in Study 5.

Predictor	*b*	*b* 95%CI [LL, UL]	*beta*	*r*	Fit	Difference
(Intercept)	3.00**	[2.79, 3.14]				
Benevolent Sexism	.13*	[0.02, 0.24]	.08	.07		
Moral Ambiguity	2.13**	[1.93, 2.33]	.71	.70**		
Participant Gender	-.21	[-0.41, -0.01]	-.07	.003		
					R^2 = .503** 90% CI [0.45, 0.56]	
(Intercept)	3.00**	[2.80, 3.20]				
Benevolent Sexism	.37**	[0.19, 0.55]	.23	.07		
Moral Ambiguity	2.13**	[1.83,2.42]	.71	.70		
Participant Gender	-.21	[-0.50, 0.07]	-.07	.003		
Benevolent Sexism x Moral Ambiguity	-.25**	[-0.46, -0.32]	-.11	-.003		
Benevolent Sexism x Participant Gender	-.22**	[-0.44, -0.01]	-.10	-.04		
Moral Ambiguity x Participant Gender	-.03	[-0.43, 0.37]	-.01	.41**		
					R^2 = .514** 90% CI [0.46. 0.57]	ΔR^2 = .012* 90% CI [-0.005, 0.03]
(Intercept)	3.01**	[2.81, 3.20]				
Benevolent Sexism	.46**	[0.25, 0.66]	.28	.07		
Moral Ambiguity	2.13**	[1.84, 2.42]	.71	.70		
Participant Gender	-.21	[-0.49, 0.07]	-.07	.003		
Benevolent Sexism x Moral Ambiguity	-.43**	[-0.73, -0.13]	-.19	-.003		
Benevolent Sexism x Participant Gender	-.41**	[-0.72, -0.10]	-.18	-.04		
Moral Ambiguity x Participant Gender	-.03	[-0.49, 0.07]	-.01	.41**		
Benevolent Sexism x Moral Ambiguity x Participant Gender	.36	[-0.06, 0.79]	.12	-.03		
					R^2 = .518** 90% CI [0.46, 0.57]	ΔR^2 = .003 90% CI [-0.01, 0.0

Note: A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance (p < .05) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates p < .05. ** indicates p < .01.

Finally, we found a marginally significant three-way interaction between moral extremity, participant's gender, and benevolent sexism, $b = .37$, $t(430) = 1.67$, $p = .096$, which revealed that the pattern described above was significant for male participants, but not for female participants (see appendix F.3 for details).

4.2.3. Discussion

The results reveal an interaction between benevolent sexism and moral extremity, but with a different pattern from what was expected. Specifically, endorsement of benevolent sexism was positively associated with (more lenient) moral judgements of a female target who engaged in a clearly immoral action, but there was no association between benevolent sexism and moral judgements of the female target when the action was morally ambiguous. Though this supports the idea that benevolent sexism also moderates moral judgements in scenarios that do not have a direct bearing on the gender hierarchy, this pattern was surprising and requires further examination.

The results must be understood by reference to the finding that, though the moral judgements of the ambiguous scenario were around the mid-point of the scale in the pilot study, they were actually clearly above the mid-point of the scale in Study 1. That is, the ambiguous scenario was not really ambiguous, but moral. In turn, the clearly immoral scenario was perceived as immoral, but the extent to which this was the case depended on participants' benevolent sexism. Participants high in benevolent sexism perceived the female target as more moral than did participants low in benevolent sexism, but only when the action was immoral. Seen in this way, the results make sense and question our ability to distinguish ambiguous from clearly immoral scenarios when moral judgements are so subjective. In addition, these results raise the possibility that our pilot study with 40 participants did not provide sufficient precision to estimate ambiguity defined as ratings around mid-point of the scale.

Given the need to replicate these results, we conducted a second study, in which we additionally manipulated the gender of the target. Since benevolent sexism aims to regulate women's behaviour, we expect that these effects of benevolent sexism are likely to only emerge for female targets.

4.3. Study 6

In this study, we aimed to examine whether or not the results of Study 1 can be replicated, especially given the unexpected interaction pattern. We also aimed to extend our focus to examine whether the interplay between benevolent sexism and moral extremity was restricted to female targets. Given that benevolent sexism aims to regulate women's behaviour, we expected this to be the case. That is we hypothesised that the interaction between benevolent sexism and moral extremity would emerge for female targets, but not for male targets (Hypothesis 3).

4.3.1. Method

Design. The study had a 2 (Moral extremity: Ambiguously immoral vs. Clearly immoral) X 2 (Participant's gender: Female vs. Male) X 2 (Target's gender: Female vs. Male) between participants quasi-experimental design. Participants' benevolent sexism score was measured on a continuous scale, and added to the model as a group-mean centred moderator. Our dependent variable was moral judgement of the action by female and male targets.

Participants. A priori power analysis**,** using G*Power (Faul, et al., 2007), with a small to medium effect size f = .15, alpha = .05 and power = .80 based on the hypothesised three-way interaction between benevolent sexism, moral extremity and target's gender, suggested a total sample size of 830. Therefore, data was collected via an online survey from 827 participants (455 females, 372 males; M_{age} = 35.15 *SD* =15.88) who were recruited

through prolific.ac, a UK based crowdsourcing platform. In exchange of their participation in the study, participants were paid £1 at the end of the survey.

Procedure and materials. Participants were invited to participate in a study designated as "Social Judgements and Relationships." Participants who already participated in the previous study were not allowed to participate again. Participants who agreed to participate accessed the study through a web link provided, read basic information about the study, and provided informed consent. Next, participants were assigned to one of the four experimental conditions, where they read a scenario with either a female or a male target committing either a morally ambiguous or a clearly immoral action (see Appendix E for materials). After reading the hypothetical scenario, participants judged the action, the target's moral character, and provided their overall impression of the target. Judgement of the action was measured with two items asking to what extent the action was acceptable and moral on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*). Participants also indicated to what extent they thought that the target was moral, reliable, and trustworthy with three different items on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*). Finally, participants indicated their overall impressions of the target on a 7-point scale, from 1 (*Very negative*) to 7 (*Very positive*). As in Study 1, these items were highly inter-correlated ($rs \geq .68$, $ps \leq .01$), and together constituted a reliable scale ($\alpha = .95$) so they were averaged for further analyses.

At this point, participants completed the 22 items of the Ambivalent Sexism Inventory (Glick & Fiske, 1996, using 6-point Likert-type scales, from 0 (*Strongly disagree*) to 5 (*Strongly agree*). The benevolent sexism ($\alpha = .87$) scale was reliable and remained unaffected by the manipulations of moral extremity, $F(1,819) = .09$, $p = .77$, and of target gender, $F(1,819) = .001$, $p = .97$, or the interaction between two, $F(1,819) = 2.480$, $p = .12$. The main effect of participant gender on benevolent sexism score was significant, $F(1,819) = 43.13$, $p < .001$, indicating that male participants ($M = 2.12$, $SD = .05$) scored higher on benevolent

sexism scale than female participants (*M* =1.68, *SD* = .05) did. The interaction between participant gender and target gender on benevolent sexism was not significant, *F*(1,819) = 1.59, *p* =.21, the interaction between participant gender and moral extremity was not significant, *F*(1,819) = .98, *p* = .32. The three-way interaction between three was not significant, *F*(1,819) =.02, *p* = .88. Before finishing the study, participants were asked to provide some demographic information (age, gender, and ethnicity). Finally, they were informed about the actual aims of the present research, fully debriefed, and paid for their participation to the survey.

4.3.2. Results

We conducted a regression analysis (see Table 4.6) in which all predictors were entered in the first step, and their interactions in the second step. Moral extremity was coded as 1 = morally ambiguous, 0 = clearly immoral, target's gender was coded as 1 = female, 0 = male, and participant's gender was coded as 1 = female, 0 = male. Benevolent sexism scores were group mean-centred. (see Appendix F.2 for analysis with HS)

Table 4.10. Summary of means, standard deviations and correlations for variables in Study 6.

Measures	***M***	***SD***	**1**	**2**	**3**
1. Moral Judgement	3.92	1.57	__		
2. Benevolent Sexism	1.87	0.97	.09*	__	
3. Hostile Sexism	1.98	1.12	.06	.50**	__

Note: * indicates significance at the level 0.05, **indicates significance at the level 0.01 (2-tailed).

Table 4.11. Regression analysis predicting moral judgments in Study 6.

Predictor	*b*	*b* 95%CI [LL, UL]	*beta*	*r*	Fit	Difference
(Intercept)	3.00**	[2.86, 3.14]				
Benevolent Sexism	.14**	[0.06, 0.21]	.08	.07*		
Moral Ambiguity	2.35**	[2.21, 2.49]	.74	.74**		
Participant Gender	-.16*	[-0.231, -0.02]	-.08	-.05		
Target Gender	-.04	[-0.18, 0.11]	-.01	-.01		
					R^2 = .557** 90% CI [0.52, 0.60]	
(Intercept)	2.97**	[2.78, 3.16]				
Benevolent Sexism	.09	[-0.06, 0.23]	.05	.07*		
Moral Ambiguity	2.35**	[2.10,2.61]	.74	.74**		
Participant Gender	-0.17	[-0.41, 0.06]	-.06	-.05		
Target Gender	0.11	[-0.14, 0.36]	.04	-.01		
Benevolent Sexism x Moral Ambiguity	-0.02	[-0.18, 0.13]	-.01	.04		
Benevolent Sexism x Participant Gender	0.08	[-0.08, 0.23]	.03	.09**		
Benevolent Sexism x Target Gender	0.04	[-0.12, 0.19]	.02	.09**		
Moral ambiguity x Target Gender	-0.20	[-0.50, 0.09]	-.05	.40**		
Participant Gender x Target Gender	-0.11	[-0.40, 0.18]	-.03	-.03		
Moral Ambiguity x Participant Gender	0.15	[-0.14, 0.44]	.04	.46**		
					R^2 = .560** 90% CI [0.52, 0.60]	ΔR^2 = .003 90% CI [-0.003, 0.01]
(Intercept)	2.99**	[2.79, 3.19]				
Benevolent Sexism	0.23*	[0.04, 0.42]	.14	.07*		
Moral Ambiguity	2.28**	[1.98, 2.57]	.72	.74**		
Participant Gender	-0.22	[-0.48, 0.04]	-.07	-.05		
Target Gender	0.06	[-0.23, 0.34]	.02	-.01		
Benevolent Sexism x Moral Ambiguity	-0.32*	[-0.60, -0.05]	-.13	.04		
Benevolent Sexism x Participant Gender	-0.15	[-0.39, 0.09]	-07	.09**		
Benevolent Sexism x Target Gender	-0.04	[-0.30, 0.23]	-02	.09**		
Moral ambiguity x Target Gender	-0.04	[-0.47, 0.39]	-01	.40**		
Participant Gender x Target Gender	0.01	[-0.38, 0.39]	.00	-.03		
Moral Ambiguity x Participant Gender	0.30	[-0.10, 0.70]	.08	.46**		
Benevolent Sexism x Moral Ambiguity x Participant Gender	0.47**	[0.16, 0.78]	.15	.09**		
Benevolent Sexism x Moral Ambiguity x Target Gender	0.08	[-0.23, 0.38]	.02	.06*		
Moral Ambiguity x Target Gender x Participant Gender	-0.33	[-0.91, 0.25]	-.07	.26**		
Benevolent Sexism x Participant Gender x Target Gender	0.06	[-0.25, 0.37]	.02	.09**		
					R^2 = .565** 90% CI [0.53, 0.60]	ΔR^2 = .006* 90% CI [-0.002, 0.01]

Table 4.6 (cont.)						
(Intercept)	2.99	[2.79, 3.19]				
Benevolent Sexism	0.24	[0.04, 0.44]	.14	.07*		
Moral Ambiguity	2.28	[1.98, 2.57]	.72	.74**		
Participant Gender	-0.22	[-0.48, 0.04]	-.07	-.05		
Target Gender	0.06	[-0.23, 0.34]	.02	-.01		
Benevolent Sexism x Moral Ambiguity	-0.35	[-0.67, -0.04]	-.14	.04		
Benevolent Sexism x Participant Gender	-0.17	[-0.44, 0.10]	-.08	.09**		
Benevolent Sexism x Target Gender	-0.07	[-0.37, 0.24]	-.03	.09**		
Moral ambiguity x Target Gender	-0.04	[-0.47, 0.40]	-.01	.40**		
Participant Gender x Target Gender	0.004	[-0.38, 0.39]	.001	-.03		
Moral Ambiguity x Participant Gender	0.30	[-0.09, 0.70]	.08	.46**		
Benevolent Sexism x Moral Ambiguity x Participant Gender	0.52	[0.10, 0.93]	.16	.09**		
Benevolent Sexism x Moral Ambiguity x Target Gender	0.14	[-0.33, 0.61]	.04	.06*		
Moral Ambiguity x Target Gender x Participant Gender	-0.33	[-0.91, 0.25]	-.07	.26**		
Benevolent Sexism x Participant Gender x Target Gender	0.11	[-0.30, 0.53]	.03	.09**		
Benevolent Sexism x Moral Ambiguity x Participant Gender x Target Gender	-0.11	[-0.74, 0.51]	-.03	.09**		
					R^2 = .565** 90% CI [0.53, 0.60]	ΔR^2 < .001 90% CI [-0.002, 0.02]

Note: A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

The results revealed a significant positive association between benevolent sexism and moral judgements, b =.14, $t(811) = 3.49$, p =.001, indicating that higher benevolent sexism was associated with judgements of the target as more moral. Contrary to Hypothesis 2 (and Study 5), the interaction between moral extremity and benevolent sexism on moral judgements was not significant, b = -.02, $t(811) = -0.31$, p = .760. Contrary to Hypothesis 3, the interaction between target's gender, moral extremity, and benevolent sexism was not significant, b =.14, $t(811)$ =.59, p =.557. However, the results revealed an unexpected significant three-way interaction between moral extremity, benevolent sexism, and participants' gender, b =.47, $t(811) = 2.94$, p = .003. This reflects the finding that the two-

way interaction between benevolent sexism and moral extremity is significant for male participants, but not for female participants.

This was revealed by two separate moderation analysis testing whether moral extremity (X) predicted moral judgements (Y) moderated by benevolent sexism (M) separately for male and female participants. We used Model 1 in the PROCESS macro created by Hayes (2016) entering benevolent sexism as a centred continuous variable. The overall model was significant for male participants, $F(3,368) = 144.95$, $p < .001$, $R^2 = .54$. The main effect of benevolent sexism was positive and significant, $b = .21$, $t(368) = 2.74$, $p = .006$. The main effect of moral extremity was also positive and significant, $b = 2.26$, $t(368) = 20.67$, $p < .001$. Additionally, the two-way interaction between moral extremity and benevolent sexism was negative and significant, $b = -.29$, $t(368) = -2.40$, $p = .017$, qualifying the effect of benevolent sexism, but not the main effect of moral extremity. That is, the interaction shows that all participants perceived the clearly immoral scenario as less moral than the morally ambiguous scenario—and, as in Study 5, in reality the morally ambiguous scenario was in fact fairly consensually perceived as moral; whereas the clearly immoral scenario was perceived as immoral, though to different degrees depending on participants' benevolent sexism. Specifically, for the clearly immoral scenario, moral judgements of the target were positively associated with benevolent sexism, $b = .21$, $t(368) = 2.74$, $p = .006$, whereas this relationship was not significant in the ambiguously moral scenario, $b = -.08$, $t(368) = -.85$, $p = .396$ (see Figure 4.5). These findings, replicate the results of Study 1, among male participants and again suggest that we failed to create a morally ambiguous condition, having instead created a clearly moral and a clearly immoral condition, the latter of which was more harshly judged by participants low in benevolent sexism.

For female participants, the overall model was also significant, $F(3,451) = 204.71$, $p < .001$, $R^2 = .58$, but the only significant effect was that of moral extremity, $b = 2.42$, $t(451) =$

24.44, $p < .001$, suggesting that the clearly immoral action was rated as less moral by female participants than the morally ambiguous action. The main effect of benevolent sexism, $b = .09$, $t(451) = 1.21$, $p = .228$, and the two-way interaction between moral extremity and benevolent sexism, $b = .17$, $t(451) = 1.65$, $p = .099$, were not significant (see Figure 4.5).

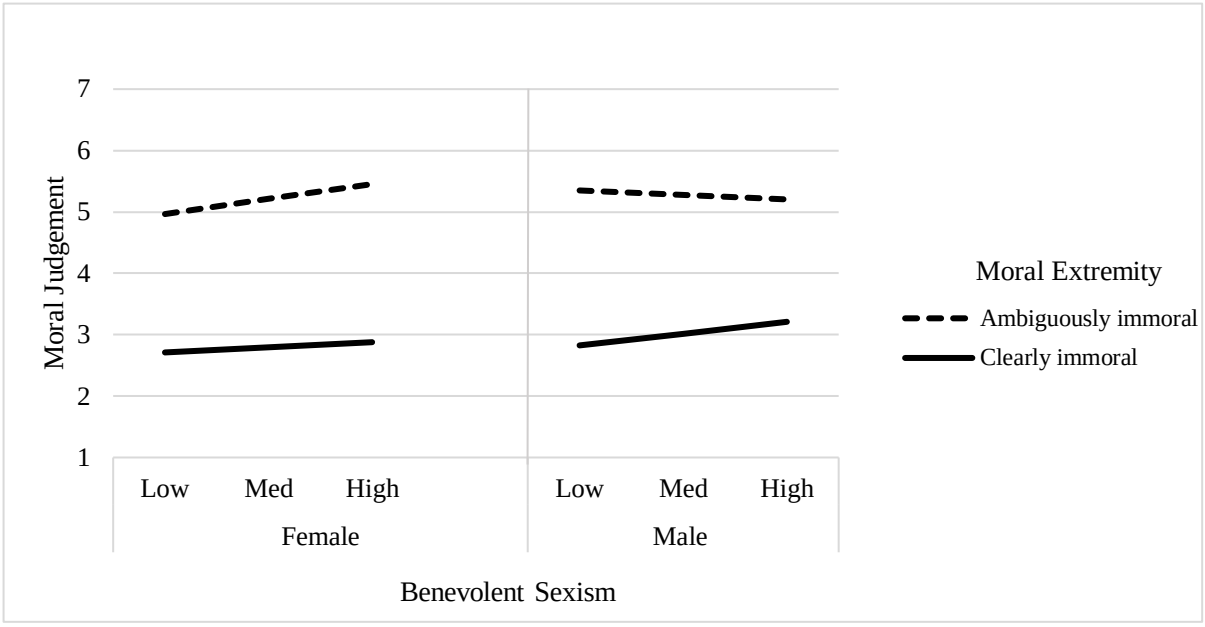

Figure 4.4. Moderating effect of benevolent sexism on moral judgements at the levels of moral extremity for female and male participants in Study 6.

4.3.3. Discussion

Study 6 replicated some findings from Study 5. Specifically, we replicated the interaction between benevolent sexism, moral extremity and participant gender already shown (though not significant) in Study 5. This enhances the idea that benevolent sexism plays a role in moral judgements, even when the context has no bearing on the gender hierarchy. However, the pattern found was again unexpected (though the same as in Study 1), in that moral judgements were only affected by benevolent sexism—and positively so—when the action was clearly immoral. These results were only revealed for male participants—an effect that was already suggested in Study 5. That is, we again found that the higher the benevolent sexism, the more male participants perceived the target as moral when the action was clearly immoral.

Surprisingly, we did not find support for the effect of target's gender on this interaction between benevolent sexism and moral extremity (Hypothesis 4). We had originally expected that the interplay between benevolent sexism and moral extremity would be likely to only emerge for female targets (but not for male targets) because benevolent sexism aims to regulate women's behaviour. Therefore, in this study, we are unable to conclude that benevolent sexism regulates women's behaviours but not men's. That it, our results suggest that endorsement of benevolent sexism might lead to overall leniency towards people who engage in clearly immoral acts, regardless of their gender.

In sum, the previous two studies showed that when the action has no direct bearing on the gender hierarchy, moral judgements—especially men's—were moderated by benevolent sexism. This is an important addition to the literature, where effects of sexism on moral judgements have focused on highly gendered actions. However, our findings did not reveal a moderation by target gender, suggesting that the role of benevolent sexism on moral judgments that are not highly gendered is not target gender specific. That is, our findings suggest that when the action is in itself not highly gendered, benevolent sexism does not seem to differentiate between male and female targets. This is odd, if one thinks of the functions of sexism as rooted in the differentiation between men and women, but it is possible that what we are capturing here is the benevolence aspect of benevolent sexism, rather than sexism *per se*.

It remains possible, however, that target gender is an important factor when the action is gendered. To test this, we conducted a final study.

4.4. Study 7

This study aims to test the moderating effect of sexism on moral judgements when the action being judged is highly gendered. To do this, we chose an action that is counter-stereotypical for women—prioritising work over child care—which is expected to elicit harsher moral judgements when it is enacted by female targets and judged by perceivers high in sexism (Hypothesis 4). That is, in this case, target gender should play a crucial role, as the action is only likely to be considered immoral (as a function of sexism) for female targets, not for male targets. This principle might of course apply to men performing in a feminine domain. However, for this study, we focused only on judgments of women. Therefore, this study examines whether sexism moderates the relationship between target gender and moral judgments when the action being performed is counter-stereotypical for women.

While in the prior studies we expected benevolent sexism to be the relevant moderator, given its idealization of women as moral, when it comes to gendered actions we considered that hostile sexism might play a more important role. This is because hostile sexism functions so as to sanction women who deviate from gendered expectations (Glick, Diebold, Bailey-Werner, & Zhu, 1997). Therefore, this time we will test whether women who choose to prioritise work over child care are considered particularly immoral by individuals high in hostile sexism (Hypothesis 4 i.e., interaction between hostile sexism and target gender). We will also explore whether benevolent sexism plays a similar role (i.e., interaction between benevolent sexism and target gender). Finally, given the work context used in this study, we chose to measure perceived target competence, in addition to target morality, so as to assess whether or not prioritising work over children paid off for women in terms of competence evaluations. Indeed, women have been shown to have to choose between warmth and competence in professional contexts (Crosby, Williams & Biernat, 2004; Cuddy, Fiske &

Glick, 2004; Fiske, Cuddy, Glick & Xu, 2002), so it is very possible that this is also reflected in a trade-off between morality and sociability.

4.4.1. Method

Design. The study had a 2 (Target's gender: female vs. male) X 2 (Participant's gender: Female vs. Male) between participants quasi-experimental design. Participants' benevolent and hostile sexism scores were continuously measured, centred, and added to the model as moderators. Our core dependent variable was again moral judgements, but this time we also added competence judgements.

Participants. A priori power analysis**,** using G*Power (Faul, et al., 2007), based on the two-way interaction between sexism and target gender, with a small to medium effect size f = .15, alpha = .05 and power = .80 suggested that a total sample size of 430 was required. Therefore, data was collected via an online survey from 432 participants (244 females, 188 males; M_{age} =32.28, *SD*=12.26) who were recruited through prolific.ac, a UK based crowdsourcing platform that supports research. In exchange of their participation in the study, participants were paid £1 at the end of the survey.

Procedure and materials. Participants were invited to participate in a study designated as "Social Judgements and Relationships." Participants who had participated in studies 1 and 2 did not participate in this one. Participants who agreed to participate accessed the study through a web link provided, read basic information about the study, and provided informed consent. Next, male and female participants were randomly assigned to one of the two experimental conditions, where they read a scenario with either a female or a male target. We changed target's name and pronouns to indicate the target's gender. The scenario was read as follows:

> *Mary [John], who is a mother [father] of two children, has been a full-time employee in the finance department of an energy company, where she [he] has been working for five years. Mary [John] has recently been promoted to a management role, which resulted in a very hectic schedule for her [him]. Therefore, she [he] is not able to spare as much time for her [his] family as she [he] used to. For instance, she [he] missed her [his] kid's school play last week since she [he] had to travel to another city for an important meeting with new business clients on the same day.*

After reading the hypothetical scenario, they were asked to judge the action, target's moral character, and their overall impression of the target separately. Judgement of the action was measured with two items asking to what extent the action was acceptable and moral on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*). Participants also indicated to what extent they thought that the target was moral, reliable, and trustworthy with three different items on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*). Finally, participants indicated their overall impressions of the target on a 7-point scale, from 1 (*Very negative*) to 7 (*Very positive*). As in the previous studies, we collapsed our dependent variables into one morality score as they were highly correlated ($rs \geq .50$, $ps \leq .01$), and together formed a reliable scale ($\alpha = .79$). In addition, participants rated to what extent the target seemed competent on a 7-point scale, from 1 (*Not at all*) to 7 (*Very much*).

To measure participants' benevolent and hostile sexism scores, participants completed the Ambivalent Sexism Inventory. Both benevolent sexism ($\alpha = .86$) and hostile sexism ($\alpha = .90$) formed reliable sub-scales and were unaffected by the manipulation of target gender, $F(1,428) = .03$, $p = .86$, $F(1,428) = .50$, $p = .48$ respectively. However, the effect of participant gender on benevolent sexism and hostile sexism was significant, $F(1,428) = 13.25$, $p < .001$, $F(1,428) = 13.25$, $p < .001$, indicating that male participants scored higher on benevolent sexism ($M = 2.20$, $SD = .07$) than female participants ($M = 1.87$, $SD = .06$), also the males were higher on hostile sexism score ($M = 2.47$, $SD = .07$) than female participants ($M = 1.66$, $SD = .06$). The interaction between target gender and participant gender did not affect benevolent sexism and hostile sexism scales, $F(1,428) = .000$, $p = .99$, $F(1,428) = .22$,

p = .64, respectively. Before finishing the study, participants were asked to provide some demographic information (age, gender and ethnicity). New to this study, given the work-related nature of the scenario, we also asked participants their employment, relationship and parental status to control for their effect on moral judgements. Finally, they were informed about the actual aims of the present research, fully debriefed, and were paid for their participation to the survey.

4.4.2. Results

Moral judgements. We started by conducting a regression analysis (see Table 4.8) in which all predictors were entered in the first step and their interactions in the second step. This time, this was done with hostile sexism as the moderator first and then, for exploratory purposes, repeated with benevolent sexism as a moderator in the model. We also ran the same analyses for competence ratings. Target's gender was coded as 1 = female, 0 = male and participant's gender was coded as 1 = female, 0 = male. Benevolent and hostile sexism scores were group mean-centred as in previous studies.

Table 4.12. Summary of means, standard deviations and correlations for variables in Study 7.

Measures	***M***	***SD***	***1***	***2***	***3***	***4***
1.Moral Judgement	5.17	0.97	__			
2.Competence Judgement	5.69	1.05	.62**	__		
3.Benevolent Sexism	2.01	0.94	-.08	-.06	__	
4.Hostile Sexism	2.02	1.04	-.12*	-.16**	.42**	__

Note: * indicates significance at the level 0.05, **indicates significance at the level 0.01 (2-tailed).

Contrary to Hypothesis 4, the interaction between hostile sexism and target's gender on moral judgements was not significant, b = -.05, t(419) = -0.49, p = .622. However, unexpectedly, the three-way interaction between hostile sexism, target's gender, and participant's gender was negative and significant, b = -.51, t(419) = -2.62, p = .009, so we decomposed this effect by participant gender. To do so, we ran a moderation analysis testing whether the effect of target's gender (X) on moral judgements (Y) was moderated by hostile

sexism (M) for male and female participants separately. The results revealed that the predicted two-way interaction between target gender and hostile sexism was significant for female participants only.

Table 4.13. Regression analysis predicting moral judgements when hostile sexism is a moderator in the model in Study 7.

Predictor	*b*	*b* 95% *CI* [*LL*, *UL*]	*beta*	*r*	Fit	Difference
(Intercept)	5.11	[4.81, 5.41]				
Hostile Sexism	-0.09	[-0.19, 0.01]	-0.09	-.08		
Target Gender	-0.02	[-0.20, 0.16]	-0.01	-.02		
Participant Gender	0.17	[-0.22, 0.36]	0.07	.12**		
					R^2 = .061** 90% CI [0.03, 0.10]	
(Intercept)	5.15	[4.83, 5.47]				
Hostile Sexism	-0.05	[-0.22, 0.12]	-.05	-.08		
Target Gender	-0.10	[-0.38, 0.17]	-0.05	-.02		
Participant Gender	0.09	[-0.17, 0.36]	0.05	.12**		
Hostile Sexism x Target Gender	-0.05	[-0.24, 0.14]	-0.03	-.08		
Hostile Sexism x Participant Gender	-0.04	[-0.23, 0.15]	-0.03	-.07		
Target Gender x Participant Gender	0.15	[-0.22, 0.51]	0.07	.07		
					R^2 = .063** 90% CI [0.03, 0.10]	ΔR^2 = .002 90% CI [-0.005, 0.01]
(Intercept)	5.15	[4.84, 5.47]				
Hostile Sexism	-0.18	[-0.38, 0.02]	-0.18	-.08		
Target Gender	-0.10	[-0.37, 0.17]	-0.05	-.02		
Participant Gender	0.08	[-0.18, 0.35]	0.04	.12**		
Hostile Sexism x Target Gender	0.22	[-0.06, 0.50]	0.16	-.08		
Hostile Sexism x Participant Gender	0.23	[-0.04, 0.51]	0.17	-.08		
Target Gender x Participant Gender	0.15	[-0.22, 0.51]	0.07	-.07		
Hostile Sexism x Target Gender x Participant Gender	-0.51**	[-0.88, -0.13]	-0.28	-.12**		
					R^2 = .079** 90% CI [0.04, 0.12]	ΔR^2 = .015** 90% CI [-0.004, 0.03]

Note. A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

Specifically, for female participants, the overall model was marginally significant, $F(8,235) = 2.12$, $p = .053$, $R^2 = .06$, and only the two-way interaction between target's gender and hostile sexism was significant, $b = -.28$, $t(235) = -2.35$, $p = .040$ (main effect of hostile sexism, $b = .05$, $t(235) = .55$, $p = .580$, main effect of target's gender, $b = .04$, $t(235) = .34$, $p = .732$). This indicates that endorsement of hostile sexism was negatively associated with moral judgements of female targets as less moral, $b = -.22$, $t(235) = -2.45$, $p = .015$. However, hostile sexism was not associated with moral judgements of male targets, $b = .05$, $t(235) = .55$, $p = .579$ (see Figure 4.6).

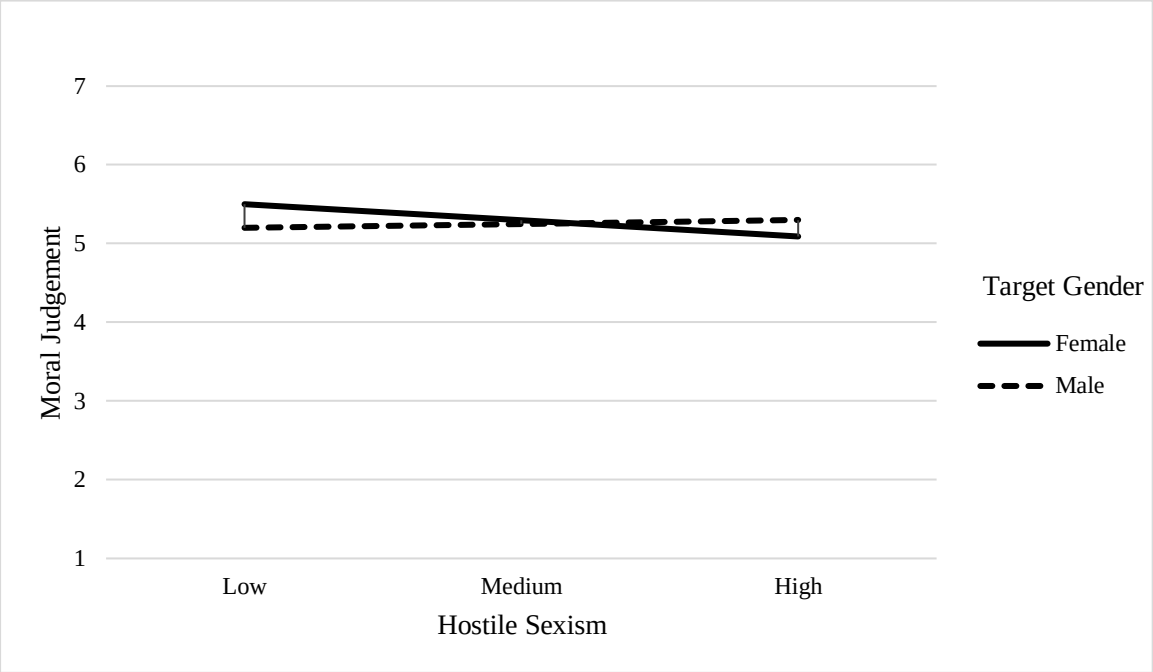

Figure 4.5. Moderating effect of hostile sexism on moral judgements at the levels of target gender for female participants.

For male participants, the overall model was significant, $F(8,179) = 2.05$, $p = .043$, $R^2 = .08$, and only the effect of hostile sexism on moral judgements was marginally significant, $b = -.19$, $t(179) = -1.90$, $p = .058$ (main effect of target's gender, $b = -.08$, $t(179) = -0.58$, $p = .565$; two-way interaction between target's gender and hostile sexism, $b = .22$, $t(179) = 1.58$, $p = .115$).

We then considered the contribution of benevolent sexism to these relationships. We therefore conducted the same regression analysis (see Table 4.9) on moral judgment but this time benevolent sexism (group mean-centred) was entered as the moderator instead of hostile sexism. There were no significant main or interaction effects on moral judgements when benevolent sexism was a moderator in the model.

Table 4.14. Regression analysis predicting moral judgments when benevolent sexism is a moderator in the model in Study 7.

Predictor	*b*	*b* 95% CI [LL, UL]	*beta*	*r*	Fit	Difference
(Intercept)	5.12	[4.82, 5.42]				
Benevolent Sexism	-0.08	[-0.18, 0.02]	-0.07	-.06		
Target Gender	-0.02	[-0.20, 0.17]	-0.01	-.02		
Participant Gender	0.17	[-0.02, 0.36]	0.09	.12**		
					R^2 = .059** 90% CI [0.02, 0.10]	
(Intercept)	5.17	[4.85, 5.49]				
Benevolent Sexism	0.004	[-0.17, 0.18]	0.004	-.06		
Target Gender	-0.10	[-0.37, 0.18]	-0.05	-.02		
Participant Gender	0.10	[-0.17, 0.36]	0.05	.12**		
Benevolent Sexism x Target Gender	-0.07	[-0.27, 0.12]	0.05	-.07		
Benevolent Sexism x Participant Gender	-0.08	[-0.27, 0.12]	-0.06	-.07		
Target Gender x Participant Gender	0.14	[-0.22, 0.51]	0.07	.07		
					R^2 = .063** 90% CI [0.03, 0.10]	ΔR^2 = .004 90% CI [-0.01, 0.01]
(Intercept)	5.17	[4.85, 5.49]				
Benevolent Sexism	-0.01	[-0.22, 0.20]	-0.01	-.06		
Target Gender	-0.10	[-0.37, 0.18]	-0.05	-.02		
Participant Gender	0.09	[-0.17, 0.36]	0.05	.12**		
Benevolent Sexism x Target Gender	-0.05	[-0.35, 0.25]	-0.03	-.07		
Benevolent Sexism x Participant Gender	-0.06	[-0.34, 0.23]	-0.04	-.07		
Target Gender x Participant Gender	0.14	[-0.22, 0.51]	0.07	.07		
Benevolent Sexism x Target Gender x Participant Gender	-0.04	[-0.44, 0.36]	-0.28	-.06		
					R^2 = .063** 90% CI [0.03, 0.10]	ΔR^2 < .001 90% CI [-0.004, 0.01]

Note: A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

Competence Judgements. As indicated, given that the scenario focused on a work context we chose to also explore competence judgements. We repeated the analyses for competence ratings, first with hostile sexism as moderator and then benevolent sexism as the moderator. We started with a regression analysis (see Table 4.10) on competence judgment in which all predictors were entered in the first step, and their interactions in the second step.

Table 4.15. Regression analysis predicting competence judgments when hostile sexism is a moderator in the model in Study 7.

Predictor	*b*	*b* 95% CI [*LL*, *UL*]	*beta*	*r*	Fit	Difference
(Intercept)	5.63	[5.31, 5.96]				
Hostile Sexism	-0.16**	[-0.26, 0.05]	-0.14	-.14**		
Target Gender	0.13	[-0.07, 0.32]	0.06	.06		
Participant Gender	0.10	[-0.11, 0.30]	0.05	.07		
					R^2 = .039* 90%CI [0.01, 0.07]	
(Intercept)	5.59	[5.24, 5.94]				
Hostile Sexism	-0.07	[-0.24, 0.11]	-.24	-.14**		
Target Gender	0.21	[-0.09, 0.51]	0.10	.06		
Participant Gender	0.17	[-0.12, 0.46]	0.08	.07		
Hostile Sexism x Target Gender	-0.05	[-0.26, 0.16]	-0.03	-.12**		
Hostile Sexism x Participant Gender	-0.10	[-0.31, 0.10]	0.16	-.07**		
Target Gender x Participant Gender	-0.15	[-0.55, 0.26]	-0.06	.05		
					R^2 = .043* 90%CI [0.01, 0.07]	ΔR^2 = .004 90%CI [-0.01, 0.01]
(Intercept)	5.59	[5.25, 5.94]				
Hostile Sexism	-0.24*	[-0.38, 0.02]	-.22	-.14**		
Target Gender	0.21	[-0.09, 0.51]	0.10	.06		
Participant Gender	0.16	[-0.13, 0.44]	0.07	.07		
Hostile Sexism x Target Gender	0.28	[-0.06, 0.50]	0.19	-.12**		
Hostile Sexism x Participant Gender	0.23	[-0.07, 0.53]	0.15	-.13**		
Target Gender x Participant Gender	-0.15	[-0.54, 0.25]	-0.07	.05		
Hostile Sexism x Target Gender x Participant Gender	-0.63**	[-1.04, -0.21]	-0.31	-.18**		
					R^2 = .063* 90%CI [0.03, 0.10]	ΔR^2 = .020 90%CI [-0.001, 0.04]

Note: A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

The three-way interaction between target's gender and hostile sexism and participant gender on competence ratings was negative and significant, b = -.63, t(419) = -2.98, p = .003. We decomposed this three way interaction by participant gender.

For male participants, the overall model was not significant, F(8,179) = 1.88, p = .217, R^2 = .06. There was a significant main effect of hostile sexism, b = -.25, t(179)= -2.20, p = .030, and a marginally significant interaction between target's gender and hostile sexism, b = .28, t(179) = 1.74, p = .084. This revealed that the endorsement of hostile sexism was not associated with competence judgements of female targets who prioritised work over child care, b = .03, t(179) = .24, p = .814. Surprisingly, a male target who prioritised work over child care was perceived as less competent as hostile sexism increased, b = -.25, t(179)= -2.20, p = .030 (see Figure 4.7). The main effect of target's gender was not significant, b =.23, t(179) = 1.44, p = .151.

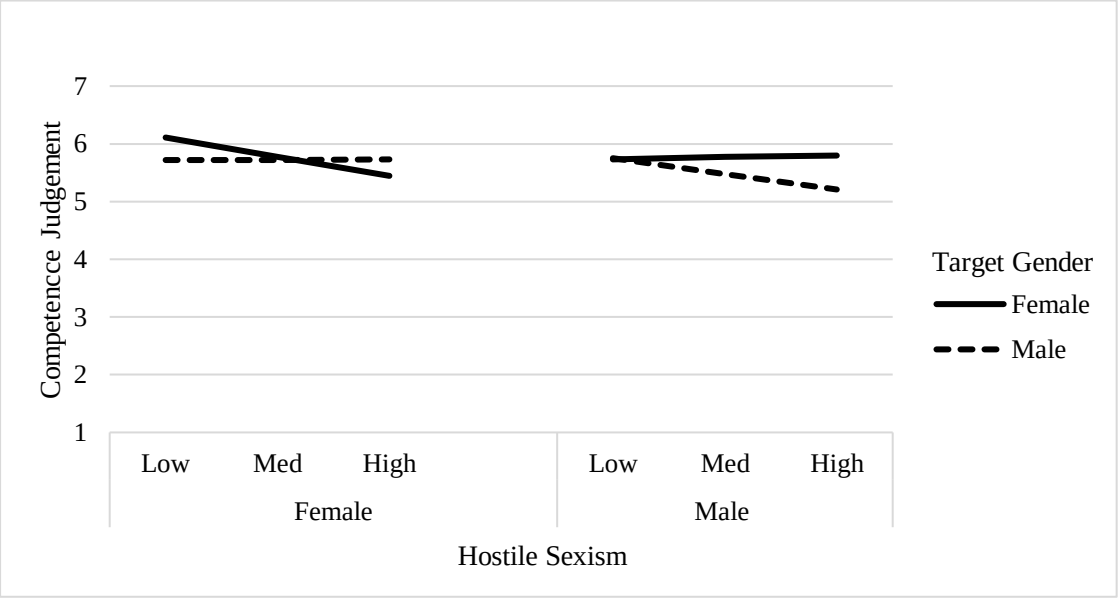

Figure 4.6. Moderating effect of hostile sexism on competence judgements at the levels of target gender for female and male participants in Study 7.

For female participants, the overall moderation model was significant, $F(8,235) = 2.50$, $p = .013$, $R^2 = .08$ and only the two-way interaction between target's gender and hostile sexism was significant, $b = -.32$, $t(235) = -2.32$, $p = .021$. This revealed that the endorsement of hostile sexism was negatively associated with competence judgement of female targets as less competent, $b = -.33$, $t(235) = -3.55$, $p < .001$, whereas it did not have an effect on competence judgements of male targets, $b = -.01$, $t(235) = -.10$, $p = .918$ (see Figure 4.7). The main effect of hostile sexism was not significant, $b = -.01$, $t(235) = -.10$, $p = .918$. The main effect of target's gender was not significant, $b = .07$, $t(235) = .53$, $p = .595$.

We then replaced hostile sexism with benevolent sexism in the same moderation model (see Table 4.11).

Table 4.16. Regression analysis predicting competence judgments when benevolent sexism is the moderator in the model in Study 7.

Predictor	*b*	*b* 95% CI [LL, UL]	*beta*	*r*	Fit	Difference
(Intercept)	5.66	[5.33, 5.99]				
Benevolent Sexism	-0.06	[-0.17, 0.05]	-0.05	-.05		
Target Gender	0.13	[-0.07, 0.33]	0.06	.06		
Participant Gender	0.10	[-0.11, 0.31]	0.05	.07		
					R^2 = .022 90% CI [-0.001, 0.04]	
(Intercept)	5.62	[5.27, 5.97]				
Benevolent Sexism	0.17	[-0.02, 0.37]	0.15	-.05		
Target Gender	0.22	[-0.08, 0.52]	0.10	.06		
Participant Gender	0.18	[-0.11, 0.46]	0.08	.07		
Benevolent Sexism x Target Gender	-0.27*	[-0.48, -0.05]	-0.01	-.12**		
Benevolent Sexism x Participant Gender	-0.17	[-0.38, 0.05]	-0.11	-.08*		
Target Gender x Participant Gender	-0.15	[-0.55, 0.25]	-0.07	.05		
					R^2 = .042 90% CI [0.01, 0.07]	ΔR^2 = .020* 90% CI [-0.01, 0.04]
(Intercept)	5.63	[5.28, 5.98]				
Benevolent Sexism	0.05	[-0.18, 0.28]	0.04	-.05		
Target Gender	0.22	[-0.08, 0.52]	0.10	.06		
Participant Gender	0.17	[-0.12, 0.46]	0.08	.07		
Benevolent Sexism x Target Gender	-0.02	[-0.35, 0.31]	-0.01	-.12**		
Benevolent Sexism x Participant Gender	0.05	[-0.26, 0.35]	0.03	-.08*		
Target Gender x Participant Gender	-0.15	[-0.55, 0.25]	-0.07	.05		
Benevolent Sexism x Target Gender x Participant Gender	-0.43	[-0.86, 0.01]	-0.21	-.16**		
					R^2 = .051* 90% CI [0.02, 0.08]	ΔR^2 = .009 90% CI [-0.01, 0.02]

Note: A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

The interaction between target gender and benevolent sexism was negative and significant, $b = -.27$, $t(419) = -2.43$, $p = .016$. Simple slope analysis revealed that the endorsement of benevolent sexism was negatively associated with competence judgements of female targets, $b = -.19$, $t(419) = -2.50$, $p = .013$, but not associated with competence

judgements of male targets, b = .09, t(419) =1.15, p = .251. The three-way interaction between target's gender, participant's gender and benevolent sexism was marginally significant, b = -.43, t(419) =-1.94, p = .053. For male participants, the overall model was not significant, F(8,179) = .74, p =.660, R^2 = .03. The main and interaction effects were not significant (main effect of benevolent sexism, b = .03, t(188) = .26, p = .795; main effect of target gender, b = .24, t(188) = 1.47, p = .143; two-way interaction between benevolent sexism and target gender, b=-.04, t(188) = -.23, p = .816).

For female participants, the overall model was marginally significant, F(8,235) = 2.49 p =.013, R^2 = .08. The two-way interaction between target's gender and benevolent sexism was significant, b = -.43, t(235) = -3.11, p = .002, indicating that endorsement of benevolent sexism was negatively associated with competence judgements of female target as less competent, b = -.33, t(235) = -3.38, p<.001, whereas it was not associated with competence judgements of male targets, b = .10, t(235) = 1.03, p = .306 (see Figure 4.8).

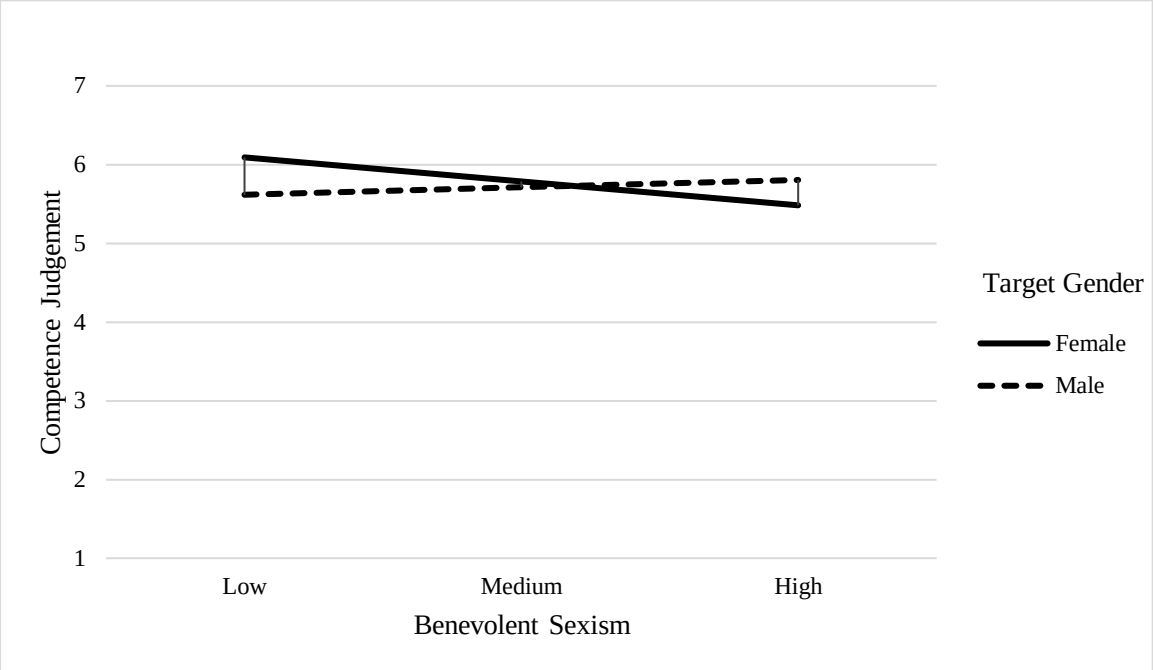

Figure 4.7. Moderating effect of benevolent sexism on competence judgements at the levels of target gender for female participants in Study 7.

4.4.3. Discussion

The aim of this final study was to examine whether target's gender would moderate the effect of sexism on moral judgements when actions are highly gendered. In this study, male and female targets engaged in the same behaviour, which was counter-stereotypical for women but not for men—i.e., favouring work over child care. We found the predicted interaction between target gender and hostile sexism, but only among female participants, not among male participants. In this study, benevolent sexism did not have a moderating effect.

As for our exploratory analysis regarding competence ratings, the effects differed for male and female participants. Male participants high in hostile sexism derogated the male target who favoured work over child care, whereas, female participants derogated the female target who displayed the same behaviour. This suggests that each gender group made harsher competence judgements of ingroup members who favoured their work but neglected their child. It might be that when explicitly asked, those who scored high on hostile sexism were concerned about being perceived as sexist about their judgments of female targets. Namely, male participants with high hostile sexism might have over-corrected their judgements in such a manner that they would not appear as sexist (overcorrection effect). Indeed, a study by Park, Smith, and Correll (2010) showed that although implicit associations of moms as caretakers and of dads as breadwinners were prominent, participants were cautious about their explicit judgments of female targets. Moreover, for female perceivers, both hostile and benevolent sexism have significant effect on competence judgments of female target whereas sexism does not affect competence judgements when the target is a male. This supports the idea that benevolent and hostile sexism aim to regulate women's behaviour to keep them within traditional gender roles.

In sum, this study suggests that when a woman performs a counter-stereotypical action for women, moral judgments are affected by hostile sexism in particular. However,

both benevolent and hostile sexism affect competence judgments of women when they violate the gender expectations (i.e., prioritizing work over child care).

4.5. Conclusion

The literature has shown that sexist attitudes towards women have an impact on how women are evaluated—specifically when women violate the traditional female role (Herzog & Oreg, 2008; Masser, Lee, & McKimmie, 2010; Viki, Abrams, & Masser, 2004; Viki, Massey, & Masser, 2005; Sakalli-Ugurlu & Glick, 2003). However, it was to date unclear whether sexism also moderated moral judgements when actions did not directly bear on the power relationship between men and women. Therefore, this research aimed to answer whether moral judgements were moderated by sexism when the actions were less gendered than in prior research.

The studies we conducted provided the support for the notion that moral judgements are affected by sexist attitudes towards women even though additional gendered information was not provided. However, the patterns found, though some were consistent across studies, were not in line with our hypotheses and therefore require particularly cautious consideration.

In studies 5 and 6 we created one scenario that was perceived as moral and one that was perceived as immoral. The extent to which the target in the immoral scenario was perceived as immoral was shaped by benevolent sexism, particularly for male participants. That is, for male participants, the higher their benevolent sexism, the more they perceived the female (Study 5) and both the male and the female (Study 6) targets as moral. This is not what we expected, so why might this be? Our argument was that benevolent sexism would shape moral judgements when actions were sufficiently ambiguous to allow room for sexist beliefs to play a role. We expected this would be the case in the scenario that we labelled ambiguous, as a result of the ratings obtained in the pilot study. However, in the main studies, with a much larger sample size, this scenario was not perceived as ambiguous at all—all

participants rated this scenario as moral. The fact that we did not find any role for sexist beliefs in this scenario, thus, is perhaps not surprising.

The clearly immoral scenario, in turn, was expected to leave less room for interpretation, and therefore also less room for effects of sexism. This scenario was indeed perceived as immoral, as planned, but we nevertheless found effects of sexism on moral judgements—high benevolent sexism led participants (especially male participants) to perceive targets as more moral when the action was clearly immoral. This would suggest that the role of benevolent sexism is restricted to immoral actions, driving more lenient judgements. However, the fact that target gender did not moderate this effect (Study 6) suggests that this is not as much an effect of sexism as an effect of benevolence. The fact that participant gender did moderate this effect does however raise further questions about how to interpret this finding

In Study 7 we explored whether sexism moderated moral judgements when the actions were clearly gendered. Whereas studies 5 and 6 focused on the 'positive' effects of benevolent sexism, as improving evaluations of (women's) immoral actions, Study 7 focused on how hostile sexism can make these evaluations more negative. We found that, for female participants, endorsement of hostile sexism was associated with harsher moral judgements of female targets who favoured work over child care, whereas they did not affect judgements of men who engaged in the same behaviour. Competence judgements were differently affected, with men suffering greater cost to their competence evaluations when they favoured work over child care, when they were evaluated by male participants high in hostile sexism. Benevolent sexism, in turn, was negatively associated with moral judgements of women who favoured work over child care as less competent only for female participants.

Overall, the present research contributes to existing literature in several ways. First, it suggests that the role of benevolent sexism can be triggered even in settings that do not bear

on the gender hierarchy (Study 5 and 6). Second, the present studies indicate that the interplay between benevolent sexism and the extent to which an action is clearly (im)moral occurs in different patterns for male and female perceivers. While endorsement of benevolent sexism was associated with more lenient moral judgements from male participants in clearly immoral scenario, it predicted more lenient judgements of targets from female participants in ambiguously immoral scenarios. Finally, the present research, in line with existing research, suggests that benevolent sexism suggest a more positive moral judgements of females (and males) whereas hostile sexism functions in a more negative way to judge women who deviates from traditional roles.

5 General Discussion

The present research aimed to investigate moral biases in minimal groups and as a function of ideologies. Previous research in moral psychology indicates that moral judgments are not controlled, consistent, objective and indiscriminative, but are substantially influenced by intuition, personal preferences and affective responses (Haidt, 2001). This suggests that moral decision making is open to intergroup biases, which often contribute to intuitive judgements. Indeed, the evidence suggests that even though people perceive their moral judgements as objective and unbiased they in fact make biased moral decisions that are then *rationalized* as objective (e.g., Haidt, Rosenberg, & Horn, 2003; Uhlmann et al., 2009).

In this thesis, I identify important gaps in this literature and report research with which I aimed to address some of these gaps. Previous research on intergroup biases mostly focused on judgements of individuals and groups on warmth and competence, but scholars have highlighted that it is crucial to examine moral judgements as a third fundamental dimension. (Leach et al., 2007). Here, we do precisely that, and examine the effect of intergroup biases on judgments of others' morality. Prior research that has looked at the effects of intergroup factors on moral judgements has mainly included pre-existing groups (e.g., race, ethnicity, gender) that are associated with certain beliefs and emotions, making it hard to determine whether observers' judgments of the group's morality arise from observers' knowledge of group stereotypes, or from ideological biases held by observers. Separating these issues is one of the central aims of the work presented in this thesis. I specifically aimed to investigate whether or not moral judgments are influenced by group membership where there is no prior information, beliefs, or emotions associated with the target group (i.e., in a minimal group setting). As in much prior research on intergroup biases (Branscombe, Wann, Noel, & Coleman, 1993; Doosje, Ellemers, & Spears, 1995; Ellemers, Spears, & Doosje, 1997; Jetten, Spears, & Manstead, 1997; Sidanus, Pratto, & Mitchell, 1994), I also examined

whether or not group identification moderated the effect of group membership on moral judgement (Chapter 3). Next, I aimed to build in group-related ideologies. To do so, I focused on gender categories and examined whether sexism moderated moral judgements of women's actions, since sexism towards women primarily aims to regulate women's behaviour (Chapter 4).

Although there is already some research that is relevant to the understanding of the link between sexism and moral judgements—such as research examining the penalties recommended for women who have committed certain crimes (e.g., Viki, Massey, & Masser, 2005)—this research does not directly measure moral judgements. In addition, existing research in this area tends to focus on actions that are highly gendered and have a specific bearing on the power relation between men and women, such as women's deviation from gender role prescriptions (Herzog & Oreg, 2008; Sakalli-Ugurlu & Glick, 2003; Masser, Lee, & McKimmie, 2010; Viki, Massey, & Masser, 2005). Therefore, when examining the role of sexism in moral judgements, I aimed to examine whether it would play a similar role when the action to be judged does not directly bear on the power relationship between men and women (e.g., careless driving). Finally, I chose to focus specifically on the possible moderating effect of benevolent sexism, since it is a form of sexism that directly refers to women's morality. Specifically, benevolent sexism idealizes women as morally pure and superior to men (Glick & Fiske, 1995). As such, for people who endorse benevolent sexism, an immoral action displayed by a woman would be both unexpected (due to the descriptive aspect of gender stereotypes) and particularly problematic (due to the prescriptive aspect of gender stereotypes; Rudman & Glick, 2010).

My goal in this thesis was therefore to examine whether moral judgments would be influenced by biases based on group membership and associated ideologies. I examined this both with experimentally created (Chapter 3) and pre-existing social groups (Chapter 4).

5.1. Summary of Findings

5.1.1. Moral Biases in Minimal Groups

In Chapter 3, I reported four studies looking at the effect of (minimal) group membership on moral judgments. These studies aimed to provide evidence that group biases in moral judgements emerge even in minimal group conditions. Therefore, the studies in this chapter involved experimentally created groups. I additionally varied the type of action being judged, to examine whether biases were more likely to emerge for actions that are typically—i.e., in the absence of group biases—considered more or less permissible. Finally, we examined whether group identification moderated these effects. I expected that participants would judge the in-group agent as more moral than the out-group agent and that this would be the case in particular for high identifiers.

In Study 1, each participant read two versions of the trolley dilemmas, where the moral agent was either an ingroup or an outgroup member. The dilemmas varied in the extent to which they are typically judged as permissible, with the switch scenario (where the agent causes the death of one person to save others by switching the track on which an uncontrolled train rides) being typically judged as more morally permissible than the footbridge scenario (where a similar result is obtained by pushing a person from a footbridge onto the train track). After reading each dilemma, participants indicated their moral judgement of the agent. The results revealed an *ingroup bias* in the switch scenario when group identification was high, but no bias in the footbridge dilemma. This suggested that ingroup biases on moral judgements do emerge in these minimal contexts, that they can be moderated by the type of action being judged, and that group identification plays an important role in this process.

In Study 2, we aimed to examine whether these findings could be replicated and this time manipulated moral dilemma between participants, so as to avoid strategic decisions participants could make about when biases could be better expressed. Additionally, we aimed

to examine whether similar biases would emerge when the action is actually praiseworthy or heroic. Therefore, we added another version of the trolley dilemma where the agent causes a death of someone while trying to save another one. The findings of this study did not replicate the findings of Study 1. The results did reveal group biases in moral judgements, but the pattern was different this time. Specifically, this time the results showed an *outgroup bias* in the footbridge scenario that is seen as clearly wrong. This suggest that biases are designed to protect the moral reputation of the ingroup. However, no group bias emerged in the other scenarios.

The first two studies have already identified two different ways in which intergroup biases might affect moral judgments: judging ingroup more favourable in ambiguous situations (amongst high identifiers), and judging ingroup more harshly when their actions are clearly wrong. Therefore, we conducted a third study to check whether strategic concerns played a role in moral judgments. This third study also followed a between-participants design (as Study 2), with the same two trolley dilemmas as in Study 1. The response scale, which had been adjusted in Study 2 to allow participants to praise the agent's action, was again the same as in Study 1. The agent's group membership was varied between participants. This time, the results revealed an *outgroup bias* in the switch scenario and no moderation of group identification. So, again, we found evidence in favour of the idea that group biases can emerge under minimal conditions, but the precise pattern of results was again different and therefore inconclusive.

We opted to conduct a final study to examine whether this inconsistency across the three studies could be explained by methodological differences between studies. In Study 4, we manipulated the context where moral judgements were made, specifically varying whether participants made isolated judgements of one dilemma/agent at the time, or of multiple dilemmas/agents at the same time (i.e., within versus between participants). The

results indicated that moral judgments in the switch scenario were affected by whether it was presented in an isolated or of with the footbridge scenario (i.e., the switch scenario was rated more positively when it was presented on its own than when it was presented with the footbridge scenario). However, there was no effect of identification revealed in this study. Although these results did not clarify the patterns of earlier studies, they contributed the literature by directly showing that moral judgments are context dependent—not only do they vary depending on the action, but also depending on whether more than one action being judged.

In sum, across four experimental studies with minimal groups, we found evidence for group biases, but the specific patterns obtained varied greatly. As such, these results are inconclusive and must be taken with care. However, they suggest that group membership has an influence on moral judgments even in minimal group settings.

5.1.2. Moral Biases as a Function of Sexist Ideology

It is possible that one of the reasons why the results of my first line of work were inconclusive is precisely because they isolated group membership from its deeper meaning and associated ideologies. As such, the context provided room for participants to choose to go with the flow, or to strategically adjust their responses, as determined by other considerations, such as individual differences in the importance of fairness and egalitarianism. Though I did not address this, I chose to proceed my enquiries by focusing on pre-existing social groups and relevant ideologies.

In Chapter 4 of this thesis, I report 3 studies that aimed to investigate whether moral judgments are influenced by endorsement of certain ideologies. I focused on judgements of women's morality and on sexist ideologies towards women, specifically on a type of sexism that directly refers to women's morality, i.e., benevolent sexism. Gender context is particularly interesting for moral psychology because sexist ideologies refer directly to

morality of women (e.g., women are morally pure). I hypothesized that benevolent sexism would predict moral judgements of women's actions, but given the scarcity of research in this area it was unclear whether this would happen because benevolent sexism would make it harder to perceive women's actions as immoral (as women are expected to be moral), or because immoral actions enacted by women would be more strongly rejected (as moral actions are prescribed for women). To examine this, I opted to vary the moral extremity of the action and expected that participants high in benevolent sexism would be less ready to see women as immoral when they enact actions that are seen as only slightly immoral (or about which people are more divided). By contrast, I expected that participants high in benevolent sexism would be particularly harsh towards women who enacted clearly immoral actions.

These predictions were tested in the studies reported in Chapter 4 of this thesis. The results of Study 5 revealed, as predicted, an interaction between moral extremity and benevolent sexism, but the pattern was very different from what we expected. Specifically, we found that benevolent sexism did not affect moral judgements of women who displayed morally ambiguous actions, but led to *more lenient* judgements of women who displayed *clearly immoral* actions. A marginally significant three way interaction suggested that this was particularly the case among male participants. It is possible that men with higher benevolent sexism respond to power relations in the case of judging morality of women with a show of niceness. Research suggests that the stable high status groups might not need to discriminate against a low status group because their position in the social hierarchy is already securely prestigious. This secure position might even sometimes lead to outgroup bias (noblesse oblige; Leach, Snider, & Iyer, 2002; Shepherd, Spears, & Manstead, 2013). However, it is also worth nothing at this point that the morality ratings made by participants in the study differed from those made by pilot participants, in that participants in the main study judged the scenario intended to be ambiguous as actually moral. In that sense, the

results can be rephrased as showing that benevolent sexism moderates judgements of the female target's morality only when the action was immoral, but not when it was moral, which is more in line with our original reasoning.

Give the unpredicted results, we sought to replicate these findings in an additional Study 6. We also opted to add male targets to the female targets examined in Study 5, to examine whether the effects of benevolent sexism were specific to cases where participants judged women's moral behaviour, as ambivalent sexism theory would lead us to expect (Glick & Fiske, 1995). Indeed, we expected that effects of benevolent sexism would be specific to female targets, since benevolent sexism is proposed to be a sexist ideology that aims to regulate women's behaviour. The results again revealed a significant effect of moral extremity and a significant interaction between benevolent sexism and moral extremity. The patterns replicated those found in Study 5. As before benevolent sexism moderated moral judgements when the action was clearly immoral but not when it was (ambiguously) moral suggesting that benevolent sexism moderated judgments of immoral, but not of moral actions.

This time, the three way interaction between benevolent sexism, moral extremity, and participant gender was actually significant, revealing that the two-way interaction between moral extremity and benevolent sexism was significant only for male participant, not for female. Finally, the study did not reveal any effects of target gender, which was also unexpected—that is, for male participants judging a clearly immoral action, the higher their benevolent sexism the more they judged the target as moral, irrespective of the target's gender. This suggests that perhaps the measure of benevolent sexism, at least in this context, taps into a more general benevolence, rather than on to the concept of benevolent sexism per se. This is slightly troubling, given all the research and knowledge that has relied on this concept and measure, but it needs to be considered. Alternatively, it is possible that benevolent sexism towards women overlaps with benevolent sexism towards men (Glick &

Fiske, 1999; Glick et al., 2004) and that what we are seeing in these studies is the shared influence these concepts might have on moral judgements of male and female targets. However, benevolent sexism towards men, as originally conceptualized, does not involve the same type of assumed moral superiority in men as benevolent sexism towards women does, so this is unlikely to be the case. Research on sexism towards men is, however, very scarce and might require further development if we are to understand these patterns more completely.

To close my examination of these issues, I conducted a final study (Study 7, also reported in Chapter 4) to examine whether sexism would also moderate moral judgements of men and women who enact gendered actions. I was interested in judgements of women in particular and whether sexism would determine the extent to which a counter-stereotypical action (e.g., prioritising work over child care) would be regarded as immoral, especially as compared to when a man displays the very same action (which is then not considered counter-stereotypical). In this case, we also focused on the possible effect of hostile sexism, given that—though it does not explicitly refer to women's (im)morality—it expresses antagonism towards women who behave in counter-stereotypical ways.

The results of Study 7 revealed that, for female participants patterns were as expected: endorsement of hostile sexism was associated with harsher moral judgements of female targets who favoured work over child care. This was not the case when female participants evaluated a male target who displayed the exact same behaviour. The judgements made by male participants were only affected by hostile sexism (and not the target's) gender, with higher hostile sexism leading to more negative moral judgements of both male and female targets. Finally, benevolent sexism had no significant effects on moral judgements in this study.

In sum, we found that hostile sexism negatively affected the moral judgements of women who behaved counter-stereotypically, but not of men who displayed the same (stereotypically male) behaviour, but this was only the case for female participants. It is however important to note that the effects were small and judgements of male and female targets were all above the mid-point of the scale.

5.3. Key Contributions of the Present Thesis

The present research contributes to existing literature in several ways. First, broadly, it supports the idea that biases in moral judgments are influenced by group membership and intergroup biases, as well as gender-based ideologies like benevolent and hostile sexism, and contextual factors. This adds to prior knowledge on intergroup biases, which tended to focus on other domains such as competence, or warmth more generally. My research takes on the finding that morality is the most important dimension in the positive evaluations of individuals and groups (Brambilla, Sacchi, Pagliaro, & Ellemers, 2013; Cottrell, Neuberg, & Li, 2007; Leach, Ellemers, & Barreto, 2007; Pagliaro, Ellemers, Barreto, & Di Cesare, 2016; van der Lee, Ellemers, Scheepers, & Rutjens, 2017) and reveals that group biases can be expected also here. This finding in itself has implications across a variety of areas where moral judgements are made, ranging from judgments within criminal justice systems to job applications. With regards to the effect of intergroup biases on moral judgments, this research evidences intergroup biases in moral judgments even when *minimal* groups were used, allowing us to isolate effects of categorization from ideological or historical factors. Though the specific patterns we found in this line of work were inconsistent, we did consistently find (in- or out-) group biases in all studies. This suggests that future research will need to delve more deeply into when and why these biases emerge, but also that we need to be alert for the emergence and use of such biases in contexts where newly formed groups come into contact, such as during organizational mergers.

Fitting this research into broader trends in the literature, and following from the above, the present research brings together accounts of moral judgements from moral psychology and the social psychology of intergroup relations. Indeed, this research supports both the idea that group membership and related ideologies affect moral judgements, producing intergroup biases in moral judgements, and the idea that the particular action being judged plays an important role and can over-rule group biases. Combining these insights from different literatures is an important aspect of good science and allows new insights come to flourish.

With regard to the specific ideologies I examined, these findings extend existing research by showing that the effect of benevolent sexism can be triggered in settings that do not directly bear on the power relations between men and women, such as when judging careless driving. It is true that driving is regarded as a stereotypical male skill (Ozkan & Lajunen, 2006), and therefore poor driving is a stereotypically female display, but driving is not something that men do to women, or that women do to men. This can be contrasted with other gendered behaviours such as rape or domestic violence, which directly speak of the power men hold over women. The fact that benevolent sexism also affects moral judgements in such domain that are not bearing on gender-hierarchy suggests that its effects are more widespread. At the same time, we found that benevolent sexism was positively associated with the judgements or female *and* of male targets, which raises questions regarding what this concept means, or at least what the measure taps into. Clearly, then, these findings contribute to raising questions as to what these measures represent in these settings and to highlight the need for more research on the effects of benevolent sexism on judgements of men.

I also contribute to existing knowledge by showing that sexism (this time hostile sexism) determines the moral judgement female participants made of other women. This is suggestive of system justification beliefs that are shared by minority and disadvantaged

groups (Jost & Kay, 2005; Pratto, Sidanus, & Levin, 2006) and contributes to a growing area of research examining the effects of sexism on relationships amongst women (Van Breen, Barreto, Darden. & Dimitriou, in preparation). It is unclear, however, why these effects did not emerge for male participants, as system justification (and social dominance) beliefs tend to operate similarly for majority and minority groups. Future research will need to delve into this more deeply to shed further light on these findings.

5.4. Limitations and Future Directions

A rather striking limitation of my research is that the findings that emerged from my first line of research—on intergroup biases in moral judgments in minimal groups—revealed inconsistent patterns across studies (Chapter 3). One possible reason for this inconsistency might be the sample sizes were not large enough to generate reliable and replicable results. Future research should aim to replicate these studies with better powered designs.

Another limitation of the studies reported in Chapter 3 is the use of trolley dilemmas. Though we used trolley dilemmas to ensure that actions were not gendered, and though trolley dilemmas are a useful tool to test research questions regarding morality because they are easily modifiable (Greene, 2008; Mikhail, 2011), research using these dilemmas has important limitations. For example, trolley dilemmas lack mundane and psychological realism and might therefore not elicit the same moral processes as real moral encounters (Bauman, McGraw, Bartel, & Warren, 2014). They also involve single item responses, the reliability of which cannot be checked. This in itself might explain why the effects were so different across studies. Future research using trolley dilemmas might take steps to address these concerns increasing its validity.

Another limitation of these studies is the use of an online data collection method, which restricted our methodological options. Different versions of trolley dilemmas have been shown to tap into different moral decision making processes and these can be evidence

by examining activity in different brain areas, or by examining response times. Unfortunately, by using online surveys, it was not possible to reliably assess participants' response times, which did not allow us to delve more deeply into these different processes. Future research might take response times into account when investigating group biases in moral decision making, to further validate the difference between processes tapped into by the different scenarios and shed further light on the circumstances under which additional variables (like group membership or ideology) might affect moral decision making.

With regard to the research reported in Chapter 4, a methodological shortcoming is that the manipulations were not perceived by participants in the main study in exactly the same way as by pilot participants. This might have simply been a result of the small sample size used in the pilot study. However, it is also important to consider that, as my studies also show, moral judgements are subjective and determined by context, so it is possible that the mere presence of group membership cues alters moral judgements by making these less ambiguous. This is a possibility that can be further examined by future research.

It is possible that the scenarios we used in these studies tapped more directly into competence rather than morality. That is, it is possible that provoking an accident was perceived as more dumb than immoral. Although it is clear that the scenarios were perceived as different in morality, there is evidence indicating that an action can often be construed both in competence and morality terms (Wojcizke, 1994). Without teasing these domains apart, especially when examining gender biases, we are unlikely to achieve a precise understanding of these processes.

Finally, in the studies reported in Chapter 4 we also saw that benevolent sexism had similar effects for male and female targets, which was surprising and raises questions either about the validity of the benevolent sexism measure, or about the validity of the benevolent sexism concept, or both. Manipulating the gender relevance of the scenario—e.g., one

scenario including behaviour that enhances the patriarchy and another not—might well assist in shedding further light on this issue. More generally, however, research in this area has not paid sufficient attention to how benevolent sexism might affect judgements of men, a knowledge which our findings suggest is needed.

Apart from all these methodological limitations, the present research is an attempts to integrate literatures from moral and intergroup psychology. Indeed, combining insights from different literature is an important aspect of good science, and is well placed to give new insights. However, drawing on different theoretical traditions could bring potential drawbacks to the research. In the account of this research, the intersection between moral psychology and intergroup psychology could be particularly tricky because both morality and group considerations are very basic drivers of human cognition and behaviour, which might make investigating group biases in moral judgements difficult. Therefore, the unexpected results in the present research should also be considered as the early-stage problems of integrating these different literatures that can subsequently lead to novel insights in the future. It is important to further explore the intersection between moral psychology and intergroup psychology

5.5. Conclusion

My goal in this thesis was to investigate moral biases in an intergroup context. I have done this across seven experimental studies using minimal groups as well as real social categories. Even though my findings are inconclusive, they contribute to the existing evidence by revealing that group categorizations and ideologies interfere with moral judgements, further underlining their flexibility and subjectivity, and uncover various paths for the future research.

References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751-763.

Albonetti, C. A. (1997). Sentencing under the federal sentencing guidelines: Effects of defendant characteristics, guilty pleas, and departures on sentence outcomes for drug offenses. Law and Society Review, 31(4), 789-822.

Arnett, J., & Balle-Jensen, L. (1993). Cultural bases of risk behavior: Danish adolescents. *Child Development, 64*(6), 1842-1855.http://dx.doi.org/10.2307/1131473

Balliet, D., Mulder, L. B., & Van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin, 137*(4), 594-615. http://dx.doi.org/10.1037/a0023489

Barkan, S. E., & Cohn, S. F. (1994). Racial prejudice and support for the death penalty by Whites. *Journal of Research in Crime and Delinquency, 31*(2), 202-209. http://dx.doi.org/10.1177/0022427894031002007

Batson, C & Turk, C., Shaw, L., & Klein, T. (1995). Information Function of Empathic Emotion: Learning That We Value the Other's Welfare. Journal of Personality and Social Psychology. 68. 300-313. 10.1037/0022-3514.68.2.300.

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass, 8*(9), 536-554. http://dx.doi.org/10.1111/spc3.12131

Bettencourt, B. A., Charlton, K., Dorr, N., & Hume, D. L. (2001). Status differences and in-group bias: A meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological Bulletin, 127*(4), 520–542.

https://doi.org/10.1037/0033-2909.127.4.520

Biernat, M., & Thompson, E. R. (2002). Shifting Standards and Contextual Variation in Stereotyping. *European Review of Social Psychology*, *12*(1), 103–137. https://doi.org/10.1080/14792772143000030

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology*, *3*(1), 27-52.http://dx.doi.org/10.1002/ejsp.2420030103

Björklund, F., Haidt, J., & Murphy, S. (2000). *Moral dumbfounding: when intuition finds no reason*. (Lund psychological reports; Vol. Vol 1 no 2). Department of Psychology, Lund University.

Blair, I. V, Judd, C. M., & Chapleau, K. M. (2004). Facial Features in Criminal Sentencing. *Psychological Science*, *15*(10), 674–679. https://doi.org/10.1111/j.0956-7976.2004.00739.x

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*(1), 1-45. http://dx.doi.org/10.1037/0033-2909.88.1.1

Blasi, A. (1983). Moral cognition and moral action: A theoretical perspective. *Developmental Review*, *3*(2), 178-210. http://dx.doi.org/10.1016/0273-2297(83)90029-1

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition*, *32*(4), 397-408. http://dx.doi.org/10.1521/soco.2014.32.4.397

Brambilla, M., & Riva, P. (2017). Predicting pleasure at others' misfortune: Morality trumps sociability and competence in driving deservingness and schadenfreude. *Motivation and Emotion*, *41*(2), 243–253. https://doi.org/10.1007/s11031-016-9594-2

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The

primary role of morality (vs. sociability and competence) in information gathering. European Journal of Social Psychology, 41, 135-143. doi:10.1002/ejsp.744

Brambilla, M., Sacchi, S., Menegatti, M., & Moscatelli, S. (2016). Honesty and dishonesty don't move together: Trait content information influences behavioral synchrony. *Journal of Nonverbal Behavior, 40*(3), 171-186. http://dx.doi.org/10.1007/s10919-016-0229-9

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. https://doi.org/10.1016/j.jesp.2013.04.005

Branscombe, N. R., Wann, D. L., Noel, J. G., & Coleman, J. (1993). In-Group or Out-Group Extemity: Importance of the Threatened Social Identity. Personality and Social Psychology Bulletin, 19(4), 381–388. https://doi.org/10.1177/0146167293194003

Boehm, C. (2008). Purposive Social Selection and the Evolution of Human Altruism. Cross-Cultural Research, 42(4), 319–352. https://doi.org/10.1177/1069397108320422

Booth, R., Mohdin, A., & Levett, C. (2019). Bias in Britain: explore the poll results. *The Guardian*. Retrieved from https://www.theguardian.com/uk-news/ng-interactive/2018/dec/02/bias-in-britain-explore-the-poll-results

Cameron, C. D., Payne, B. K., & Knobe, J. (2010). Do theories of implicit race bias change moral judgments? *Social Justice Research*, *23*(4), 272–289. https://doi.org/10.1007/s11211-010-0118-z

Chesney-Lind, M. (1999). Challenging Girls' Invisibility in Juvenile Court. The ANNALS of the American Academy of Political and Social Science, 564(1), 185–202. https://doi.org/10.1177/000271629956400111

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*(4), 404–413. https://doi.org/10.1093/scan/nsq011

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*(6), 1314–1329. https://doi.org/10.1037/0022-3514.83.6.1314

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, *92*(2), 208–231. https://doi.org/10.1037/0022-3514.92.2.208

Crosby, F. J., Williams, J. C., & Biernat, M. (2004). The maternal wall. *Journal of Social Issues*, *60*(4), 675-682. http://dx.doi.org/10.1111/j.0022-4537.2004.00379.x

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2004). When Professionals Become Mothers, Warmth Doesn't Cut the Ice. *Journal of Social Issues*, *60*(4), 701-718. http://dx.doi.org/10.1111/j.0022-4537.2004.00381.x

Daly, K. and Tonry, M. (1997) Gender, Race and Sentencing. In: Tonry, M., Ed., Crime and Justice: A Review of Research, University of Chicago Press, Chicago, Vol. 22, 201-242. https://doi.org/10.1086/449263

Darley, J. M., & Berscheid, E. Y. (1967). Increased liking as a result of the anticipation of personal contact. *Human Relations*, *20*(1), 29-40. http://dx.doi.org/10.1177/001872676702000103

De Bruin, E. N. M., & van Lange, P. A. M. (2000). What People Look for in Others: Influences of the Perceiver and the Perceived on Information Selection. Personality and Social Psychology Bulletin, 26(2), 206–219. https://doi.org/10.1177/0146167200264007

Doerner, J. K., & Demuth, S. (2010). The independent and joint effects of race/ethnicity, gender, and age on sentencing outcomes in U.S. federal courts. *Justice Quarterly*, *27*(1),

1–27. https://doi.org/10.1080/07418820902926197

Doosje, B., Branscombe, N. R., Spears, R,, & Manstead, A. S. R. (1998). Guilty by association: When one's group has a negative history. Journal of Personality and Social Psychology, 75(4), 872-886.

Doosje, B., Ellemers, N., & Spears, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of Experimental Social Psychology, 31*(5), 410-436. http://dx.doi.org/10.1006/jesp.1995.1018

Dovidio, J.F., Brigham, J.C., Johnson, B.T., & Gaertner, S.L. (1996). Stereotyping, prejudice and discrimination: Another look. In C.N. Macrae, C. Stangor, & M. Hewstone (Eds.), Stereotypes and stereotyping (pp. 276-319). New York: Guilford.

Dovidio, J., Kawakami, K., & Beach, K. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. L. Gaertner (Eds.), Blackwell handbook of social psychology: Vol. 4. Intergroup relations (pp. 175–197). Oxford, England: Blackwell.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking Deathworthy: Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes. Psychological Science, 17(5), 383–386. https://doi.org/10.1111/j.1467-9280.2006.01716.x

Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology, 72*(3), 617-626. http://dx.doi.org/10.1037/0022-3514.72.3.617

Ellemers, N., Spears, R., & Doosje, B. (Eds.). (1999). Social identity: Context, commitment, content. Oxford, England: Blackwell Science

Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The Psychology of

Morality: A Review and Analysis of Empirical Studies Published From 1940 Through 2017. *Personality and Social Psychology Review.* https://doi.org/10.1177/1088868318811759

Faul, F., Erdfelder, E., Lang, AG. et al. Behavior Research Methods (2007) 39: 175. https://doi.org/10.3758/BF03193146

Fehr, E. & Gächter, S. (2002). Altruistic Punishment in Humans. Nature. 415. 137-40. 10.1038/415137a.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2018). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Social Cognition: Selected Works of Susan Fiske*, *82*(6), 163–214. https://doi.org/10.4324/9781315187280

Gilligan, C. (1982). In a different voice: Psychological theory and women's development. Cambridge, MA, US: Harvard University Press.

Gino, F., Gu, J., & Zhong, C.-B. (2009). Contagion or restitution? When bad apples can motivate ethical behavior. *Journal of Experimental Social Psychology, 45*(6), 1299-1302. http://dx.doi.org/10.1016/j.jesp.2009.07.014

Glick, Peter; Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. https://doi.org/10.1111/ppl.12599

Glick P., Fiske S. T. (1999). The Ambivalence Toward Men Inventory: Differentiating hostile and benevolent beliefs about men. Psychology of Women Quarterly, 23, 519–536. doi:10.1111/j.1471–6402.1999.tb00379.x

Glick, P., Diebold, J., Bailey-Werner, B., & Zhu, L. (1997). The two faces of Adam: Ambivalent sexism and polarized attitudes toward women. *Personality and Social Psychology Bulletin*, *23*(12), 1323–1334. https://doi.org/10.1177/01461672972312009

Glick, P., & Fiske, S. T. (2006). Hostile and Benevolent Sexism. *Psychology of Women Quarterly*, *21*(1), 119–135. https://doi.org/10.1111/j.1471-6402.1997.tb00104.x

Glick, P., & Fiske, S. T. (2011). Ambivalent Sexism Revisited. *Psychology of Women Quarterly*, *35*(3), 530–535. https://doi.org/10.1177/0361684311414832

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., … López, W. L. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology*, *79*(5), 763–775. https://doi.org/10.1037/0022-3514.79.5.763

Glick P., Lameiras M., Fiske S. T., Eckes T., Masser B., Volpato C., Glick P. (2004). Bad but bold: Ambivalent attitudes toward men predict gender inequality in 16 nations. Journal of Personality and Social Psychology, 86, 713–728. doi:10.1037/0022–3514.86.5.713

Glick, P. & Rudman, L. (2010). Sexism. In J. F. DovidioM. Hewstone & P. Glick *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 328-344). London: SAGE Publications Ltd doi: 10.4135/9781446200919.n20

Glick, P., Wilkerson, M., & Cuffe, M. (2015). Masculine identity, ambivalent sexism, and attitudes toward gender subtypes: Favoring masculine men and feminine women. *Social Psychology*, *46*(4), 210-217. http://dx.doi.org/10.1027/1864-9335/a000228

Gneezy, U. 2005. Deception: The role of consequences. Amer. Econom. Rev. 95(1) 384–394.

Gollwitzer, M., & Keller, L. (2010). What you did only matters if you are one of us: Offenders' group membership moderates the effect of criminal history on punishment severity. *Social Psychology*, *41*(1), 20-26. http://dx.doi.org/10.1027/1864-9335/a000004

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029-1046. http://dx.doi.org/10.1037/a0015141

Graham J, Nosek BA, Haidt J (2012) The Moral Stereotypes of Liberals and Conservatives:

Exaggeration of Differences across the Political Spectrum. PLoS ONE 7(12): e50092. https://doi.org/10.1371/journal.pone.0050092

D. Greene, Joshua. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. Journal of Experimental Social Psychology. 45. 581-584. 10.1016/j.jesp.2009.01.003.

Greene, J, Sommerville, R, Nystrom, L, Darley, J, Cohen, J. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. Science (New York, N.Y.). 293. 2105-8. 10.1126/science.1062872.

Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037//0033-295X.

Haidt, J. (2008). Morality. Perspectives on Psychological Science, 3(1), 65–72. https://doi.org/10.1111/j.1745-6916.2008.00063.x

Haidt, J. and Graham, J. (2007). When Morality Opposes Justice: Conservatives Have Moral Intuitions that Liberals may not Recognize. *Soc Just Res*, 20(1), pp.98-116.

Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*(1), 191-221. http://dx.doi.org/10.1111/j.1559-1816.2001.tb02489.x

Haidt, J., & Joseph, C. (2004). Jonathan Haidt & Craig Joseph. *Daedalus*, *133*(4), 55–66.

Haidt, J., & Kesebir, S. (2010). Morality. In S. T. Fiske, G. Lindzey, & D. T. Gilbert (Eds.), Handbook of social psychology (pp. 797-832). Hoboken, NJ: Wiley.

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*(4), 613-628. http://dx.doi.org/10.1037/0022-3514.65.4.613

Haidt, J., Rosenberg, E., & Hom, H. (2003), Differentiating diversities: Moral diversity is not

like other kinds. *Journal of Applied Social Psychology*, 33, 1-36. doi:10.1111/j.1559-1816.2003.tb02071.x

Hauser, M. (2006). Moral minds: How nature designed our universal sense of right and wrong. New York, NY, US: Ecco/HarperCollins Publishers.

Hauser, M., Cushman, F., Young, L., Jin, R. K.-X., & Mikhail, J. (2007). A Dissociation Between Moral Judgments and Justifications. *Mind & Language, 22*(1), 1-21. http://dx.doi.org/10.1111/j.1468-0017.2006.00297.x

Hayes, A. F. (2016). The PROCESS macro for SPSS and SAS (Version 2.15) [Computer software].

Heidensohn F. (1989) Gender and Crime. In: Crime and Society. Sociology for a Changing World. Palgrave, London

Herzog, S., & Oreg, S. (2008). Chivalry and the moderating effect of ambivalent sexism: Individual differences in crime seriousness judgments. *Law and Society Review, 42*(1), 45–74. https://doi.org/10.1111/j.1540-5893.2008.00334.x

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS ONE, 7*(8). Article ID e42366. http://dx.doi.org/10.1371/journal.pone.0042366

Jetten, J., Spears, R., & Manstead, A. S. R. (1997). Strength of identification and intergroup differentiation: The influence of group norms. *European Journal of Social Psychology, 27*(5), 603-609. http://dx.doi.org/10.1002/(SICI)1099-0992(199709/10)27:5<603::AID-EJSP816>3.0.CO;2-B

Jost, J., & Kay, A. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. Journal of Personality and Social Psychology. 88. 498-509. 10.1037/0022-3514.88.3.498.

Kohlberg, L. (1963). The Development of Children's Orientations Toward a Moral

Order. *Human Development*, 6(1-2), pp.11-33.

Kohlberg, L. (1969). Stage and sequence: The cognitive development approach to socialization. In D. A. Goslin (Ed.). *Handbook of socialization theory* (pp. 347-480). Chicago, IL, Rand McNally.

Koons-Witt, B. A. (2002). The Effect of Gender on the Decision To Incarcerate Before and After the Introduction of Sentencing Guidelines. *Criminology*, *40*(2), 297–328. https://doi.org/10.1111/j.1745-9125.2002.tb00958.x

Krys, K., Capaldi, C. A., van Tilburg, W., Lipp, O. V., Bond, M. H., Vauclair, C. M., … Ahmed, R. A. (2018). Catching up with wonderful women: The women-are-wonderful effect is smaller in more gender egalitarian societies. *International Journal of Psychology*, *53*, 21–26. https://doi.org/10.1002/ijop.12420

Lammy, D. (2017). The racial bias in our justice system is creating a social timebomb. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2017/sep/08/david-lammy-review-bame-children-face-prejudice-flawed-criminal-justice-system

Leach, C. W., Carraro, L., Garcia, R. L., & Kang, J. J. (2017). Morality stereotyping as a basis of women's in-group favoritism: An implicit approach. Group Processes & Intergroup Relations, 20(2), 153–172. https://doi.org/10.1177/1368430215603462

Leach, C., 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. 234-49. 10.1037/0022-3514.93.2.234.

Leach, C. W., Snider, N., & Iyer, A. (2002). "Poisoning the consciences of the fortunate": The experience of relative advantage and support for social equality. In I. Walker & H. J. Smith (Eds.), *Relative deprivation: Specification, development, and integration* (pp. 136-163). New York, NY, US: Cambridge University Press.

Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, *31*(4), 443–453. https://doi.org/10.1177/0146167204271651

Levine, M., & Thompson, K. (2004). Identity, place, and bystander intervention: Social categories and helping after natural disasters. *The Journal of Social Psychology*, 144(3), 229-245. https://doi.org/10.3200/SOCP.144.3.229-245

Marques, J., 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. http://dx.doi.org/10.1037/0022-3514.81.3.436

Marques, J. M. and Yzerbyt, V. Y. (1988), The black sheep effect: Judgmental extremity towards ingroup members in inter-and intra-group situations. Eur. J. Soc. Psychol., 18: 287-292. doi:10.1002/ejsp.2420180308

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. http://dx.doi.org/10.1002/ejsp.2420180102

Masser, B., Lee, K., & McKimmie, B. M. (2010). Bad woman, Bad victim? Disentangling the effects of victim stereotypicality, gender stereotypicality and benevolent sexism on acquaintance rape victim blame. *Sex Roles*, *62*(7–8), 494–504. https://doi.org/10.1007/s11199-009-9648-y

Mealy, M., Stephan, W., & Carolina Urrutia, I. (2007). The acceptability of lies: A comparison of Ecuadorians and Euro-Americans. *International Journal of Intercultural Relations*, *31*(6), 689–702. https://doi.org/10.1016/j.ijintrel.2007.06.002

Mifune, N., Hashimoto, H., & Yamagishi, T. (2010). Altruism toward in-group members as a reputation mechanism. Evolution and Human Behavior. 31. 10.1016/j.evolhumbehav.2009.09.004.

Mikhail, J. (2009). Chapter 2 Moral Grammar and Intuitive Jurisprudence: A Formal Model of Unconscious Moral and Legal Knowledge. Psychology of Learning and Motivation. 50. 27-100. 10.1016/S0079-7421(08)00402-7.

Mikhail, J. (2011). Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment. New York, NY, US: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511780578

Monin, B. & Jordan, A. (2009). The Dynamic Moral Self: A Social Psychological Perspective.

Narvaez, D., & Lapsley, D. K. (Eds.). (2009). Personality, identity, and character: Explorations in moral psychology. New York, NY, US: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511627125

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353-383. http://dx.doi.org/10.1016/0010-0285(77)90012-3

Nowak, M.A. & Sigmund, K. (2005) Evolution of Indirect Reciprocity. *Nature*, 437, 1291-1298. http://dx.doi.org/10.1038/nature04131

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues, 46*(1), 1-20. http://dx.doi.org/10.1111/j.1540-4560.1990.tb00268.x

Özkan, T., & Lajunen, T. (2006). What causes the differences in driving between young men and women? The effects of gender roles and sex on young drivers' driving behaviour and self-assessment of skills. *Transportation Research Part F: Traffic Psychology and Behaviour*, *9*(4), 269-277. http://dx.doi.org/10.1016/j.trf.2006.01.005

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. https://doi.org/10.1177/0146167211406906

Pagliaro, S., Ellemers, N., Barreto, M., & Di Cesare, C. (2016). Once dishonest, always dishonest? The impact of perceived pervasiveness of moral evaluations of the self on motivation to restore a moral reputation. *Frontiers in Psychology*. 7. 10.3389/fpsyg.2016.00586.

Park, B., Smith, J. A., & Correll, J. (2010). The persistence of implicit behavioral associations for moms and dads. *Journal of Experimental Social Psychology*, *46*(5), 809–815. https://doi.org/10.1016/j.jesp.2010.04.009

Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology*, *17*, 271-320. http://dx.doi.org/10.1080/10463280601055772

Rai, T. S. (2012). Moral Psychology is Relationship Regulation. *UCLA*. ProQuest ID: Rai_ucla_0031D_10414. Merritt ID: ark:/13030/m59c6z6h. Retrieved from https://escholarship.org/uc/item/8pn9p2n7

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. http://dx.doi.org/10.1037/a0021867

Sachdeva, S., Singh, P., & Medin, D. (2011). Culture and the quest for universal principles in moral reasoning. *International Journal of Psychology*: Journal international de psychologie. 46. 161-76. 10.1080/00207594.2011.568486.

Sakalh-Uğurlu, N. & Glick, P.(2003) Ambivalent sexism and attitudes toward women who engage in premarital sex in Turkey, The Journal of Sex Research, 40:3, 296-

302, DOI: 10.1080/00224490309552194

Shepherd, L., Spears, R., & Manstead, A. S. (2013). The self-regulatory role of anticipated group-based shame and guilt in inhibiting in-group favoritism. *European Journal of Social Psychology, 43*(6), 493-504.

Sidanius, J., & Pratto, F. (1999). Social dominance: An intergroup theory of social hierarchy and oppression. New York, NY, US: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139175043

Shweder, R. A., Mahapatra, M., & Miller, J. G. (1990). Culture and moral development. In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 130-204). (Reprinted from "The Emergence of Morality in Young Children," Jerome Kagan and Sharon Lamb, eds. (Chicago & London: University of Chicago Press, 1987))New York, NY, US: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139173728.005

Shweder, R. A. (1991). Thinking through cultures: Expeditions in cultural psychology. Cambridge, MA, US: Harvard University Press.

Sidanius, J., Pratto, F., & Mitchell, M. (1994). In-group identification, social dominance orientation, and differential intergroup social allocation. *The Journal of Social Psychology, 134*(2), 151-167. http://dx.doi.org/10.1080/00224545.1994.9711378

Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass, 4*(4), 267-281. http://dx.doi.org/10.1111/j.1751-9004.2010.00254.x

Sleebos, E., Ellemers, N., & de Gilder, D. (2006). The Carrot and the Stick: Affective Commitment and Acceptance Anxiety as Motives for Discretionary Group Efforts by Respected and Disrespected Group Members. *Personality and Social Psychology Bulletin*, 32(2), 244–255. https://doi.org/10.1177/0146167205282147

Smith, I. H., Aquino, K., Koleva, S., & Graham, J. (2014). The Moral Ties That Bind . . .

Even to Out-Groups: The Interactive Effect of Moral Identity and the Binding Moral Foundations. Psychological Science, 25(8), 1554–1562. https://doi.org/10.1177/0956797614534450

Steffensmeier, D., & Demuth, S. (2006). Ethnicity and Judges' Sentencing Decisions: Hispanic-Black-White Comparisons. *Criminology*, *39*(1), 145–178. https://doi.org/10.1111/j.1745-9125.2001.tb00919.x

Steffensmeier, D., Kramer, J., & Streifel, C. (1993). Gender and Imprisonment Decisions. *Criminology*, *31*(3), 411–446. https://doi.org/10.1111/j.1745-9125.1993.tb01136.x

Steffensmeier, D., Painter-Davis, N., & Ulmer, J. (2017). Intersectionality of Race, Ethnicity, Gender, and Age on Criminal Punishment. *Sociological Perspectives*, *60*(4), 810–833. https://doi.org/10.1177/0731121416679371

Steinel, W., & De Dreu, C. K. W. (2004). Social Motives and Strategic Misrepresentation in Social Decision Making. *Journal of Personality and Social Psychology*, *86*(3), 419-434. http://dx.doi.org/10.1037/0022-3514.86.3.419

Tajfel, H. (1974). Social identity and intergroup behaviour. Information (International Social Science Council), 13(2), 65–93. https://doi.org/10.1177/053901847401300204

Tajfel, H. (Ed.). (1978). Differentiation between social groups: Studies in the social psychology of intergroup relations. Oxford, England: Academic Press.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149-178. http://dx.doi.org/10.1002/ejsp.2420010202

Tajfel, H., & Turner, J. C. (2004). The Social Identity Theory of Intergroup Behavior. In J. T. Jost & J. Sidanius (Eds.), *Key readings in social psychology. Political psychology: Key readings* (pp. 276-293). New York, NY, US: Psychology Press. http://dx.doi.org/10.4324/9780203505984-16

Terry, D. J., & Hogg, M. A. (1996). Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification. Personality and Social Psychology Bulletin, 22(8), 776–793. https://doi.org/10.1177/0146167296228002

Tillyer, R., Hartley, R. D., & Ward, J. T. (2015). Differential Treatment of Female Defendants. *Criminal Justice and Behavior*, *42*(7), 703–721. https://doi.org/10.1177/0093854814560624

Tonry, M. & Melewski, M. (2008). The Malign Effects of Drug and Crime Control Policies on Black Americans. Crime and Justice. 37. 1-44. 10.1086/588492.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge [Cambridgeshire]: Cambridge University Press.

Turner, J. C. (1999), Studies in self-categorization and minority conversion: The in-group minority in intragroup and intergroup contexts. British Journal of Social Psychology, 38: 115-134. doi:10.1348/014466699164086

Uhlmann, E., Pizarro, D. A., Tannenbaum, D., & Ditto, P. (2009). The motivated use of moral principles. *Judgment and Decision Making*, *4*(607), 479–491.

Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology*, *44*(5), 1334–1338. https://doi.org/10.1016/j.jesp.2008.03.010

van Breen, J.A., Barreto, M., Darden, S., & Dimitriou, E.(2019). Social Affiliation Following Experiences of Sexism and Sexual Harassment at Work. *Under Review*

van der Lee, R., Ellemers, N., Scheepers, D., & Rutjens, B. T. (2017). In or out? How the perceived morality (vs. competence) of prospective group members affects acceptance and rejection. *European Journal of Social Psychology*, *47*(6), 748-762. http://dx.doi.org/10.1002/ejsp.2269

Viki, G. T., Abrams, D., & Masser, B. (2004). Evaluating stranger and acquaintance rape:

The role of benevolent sexism in perpetrator blame and recommended sentence length. *Law and Human Behavior, 28*(3), 295-303. http://dx.doi.org/10.1023/B:LAHU.0000029140.72880.69

Wheeler, M., McGrath, M., & Haslam, N. (2019). Twentieth century morality: The rise and fall of moral concepts from 1900 to 2007. *PLOS ONE*, *14*(2), e0212267. doi: 10.1371/journal.pone.0212267

Wiegmann, A., & Waldmann, M. R. (2013). Transfer effects between moral dilemmas: A causal model theory. *Cognition*, *131*, 28-43. Retrieved from http://dx.doi.org/10.1016/j.cognition.2013.12.004.

Wilkerson, L. (2014). How sexism makes the man: Examining the relationship between masculinity, ambivalent sexism, and gender stereotyping. *Lawrence University Honors Projects*. 55. https://lux.lawrence.edu/luhp/55

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology, 67*(2), 222-232. http://dx.doi.org/10.1037/0022-3514.67.2.222

Wu, J., Balliet, D. P., & van Lange, P. A. M. (2016). Reputation, gossip, and human cooperation. *Social and Personality Psychology Compass*, *10*(6), 350-364. https://doi.org/10.1111/spc3.12255

Yamagishi, T., & Mifune, N. (2008). Does Shared Group Membership Promote Altruism? *Rationality and Society*, *20*(1), 5–30. https://doi.org/10.1177/1043463107085442

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151-175. http://dx.doi.org/10.1037/0003-066X.35.2.151

Appendices

Appendix A: Materials Used in Studies 1-4 with Full Wording

Dominant Perception Style and Social Judgements

This study aims to examine how people's visual perception styles can affect how they perceive social situations. In the first part of the study, you will be asked to perform a brief task to establish what your dominant visual perception style is. In the second part, you will read two short stories and will be asked to answer some questions about them. Finally, you will be asked a few demographic questions. The study will take approximately 20 minutes to complete. At the end of the study, you will be awarded £x as a token of appreciation.

In this study, participants will be asked to read some simple scenarios that involve individuals making decisions that affect other people's lives. Reading these scenarios might be upsetting to some people. If you think you might find such scenarios upsetting, it is recommended that you do not participate in this study. If you do choose to participate and, nevertheless, experience some distress, please note the resources we list at the end of the survey, which you can access to seek help.

Please also note that your participation in this study is completely voluntary. You may discontinue your participation at any time without prejudice. You do not have to answer any questions that you may rather not answer—however, if you leave questions unanswered we will not be able to use your responses, so please do try to answer all questions. Note also that all your answers are anonymous and all data collected will be treated confidentially. All data will be stored electronically and will only be used for research purposes.

Upon completion of this study, you will be provided with more information regarding the aims and expected outcomes of this research. If you have further questions or concerns regarding your participation in this study, or if you wish to receive a brief summary of the results of this study at a later point in time, please contact Esra Dasci (PhD candidate) at ed368@exeter.ac.uk or Prof Barreto at m.barreto@exeter.ac.uk.

This study has been reviewed and approved by the Ethics Review Board of School of Psychology at the University of Exeter. *If you are happy with the above information and would like to participate in this study, please click "Continue" to start. If you do not wish to participate, you can just close your browser. You can also go directly to the last page of this study, if you wish to see the list of resources we provide for individuals struggling with their psychological wellbeing.*

Part 1: Dominant perception task

Research has shown that there are two main styles of visual perception: Global and detailed perception. While everyone uses both perception styles, each individual has a tendency to use one more than the other.

Individuals who tend to use a global perception style more often are designated as **Global Perceivers.** This means that, when observing an object or situation, they most immediately identify its general or global features. Their main focus is the overall image, or situation—i.e., the 'whole'. Global perceivers usually process information by starting from general elements and proceeding to an analysis of more local details.
Individuals who tend to use a detailed perception style are designated as **Detailed Perceivers**. This means that, when observing an object or situation, they tend to first direct their attention to details and local features. Therefore, their main focus on the specific details–i.e., the 'parts' that make the whole. Detailed perceivers usually process information by starting from detailed elements and proceeding to an analysis of the whole.

Next, you will perform a task that can identify which is your dominant perception style. In this task, you see either an S or an H (smaller or larger) on the computer screen. Your task is to press one of two keyboard buttons to indicate which letter you have seen first on the screen. If you see an S first, you should press the key "s". If you see an H first, you should press the key "h".

When you are ready, please use the mouse to click on the 'start' key.
***
(After completion of Navon task, a clock will appear for 15 seconds and says)
"Your responses are being scored by the computer. You will get to know your dominant perception style soon!"
***
Your predominant perception style is: Global/Detailed!
***
Now take a moment to think what it means to you to be a global/detailed perceiver. How do you imagine that this affects your daily life—how you perceive the world around you, how you think, and how you make decisions, among other things. After you have given this some thought, please write below a description of what being a global/detailed perceiver means to you. (max 100 words)

***
We would now like to ask you to indicate how you feel about being a global/detailed perceiver. To do so, please read each of the following questions and provide your answer by clicking on the number that corresponds to your response.

Not at all 1 2 3 4 5 6 7 Very much

1. To what degree do you identify as a global/detailed perceiver?
2. To what extent are you glad to be a global/detailed perceiver?
3. To what degree do you feel strong ties with other global/detailed perceivers?
4. To what degree do you see yourself as a global/detailed perceiver?

In the second part of this study we would like to examine how individual's perception styles affect how they perceive social situations. Prior research has shown that these perception styles affect a variety of basic cognitive processes, but so far we know very little about how they affect perceptions of social situations. To help us examine this, we ask you to read two short stories and to answer some questions about them. There are no right or wrong answers to these questions—what matters to us is your personal opinion. Please do not think too long about your answers—we are interested in your initial opinion.

Scenario 1 (Switch)

D is a 21 year-old university student, who is currently studying History. D, who is a global/detailed perceiver, enjoys travelling in their free time. One day, D is a passenger on a train whose driver has just shouted that the train's brakes have failed, and subsequently fainted of the shock. On the track ahead are five people; the banks are so steep that these five people will not be able to get off the track in time. The track has a side track leading off to the right, and D can turn the train onto it. Unfortunately there is one person on this right hand track. D can turn the train to this right hand track, killing the one; or can refrain from turning the train, letting the five die.

To what extent do you think that it is morally permissible for D to switch the train to the side track? (Drag the cursor to indicate your answer)

Forbidden 1 2 3 4 5 6 7 Permissible

Scenario 2 (Footbridge)

F is a 25 year-old graduate whose favorite hobbies are reading and doing sports. F, who is a global/detailed perceiver, is on a footbridge over the train tracks. F knows trains and can see that the one approaching the bridge is out of control. On the track under the bridge there are five people; the banks are so steep that they will not be able to get off the track in time. F knows that the only way to stop an out of control train is to drop a very heavy weight into its path. But the only available, sufficiently heavy weight is a large man wearing a backpack, also watching the train from the footbridge. F can throw the man with the backpack onto the track in the path of the train, killing him; or can refrain from doing this, letting the five die.

To what extent do you think that it is morally permissible for F to throw the man onto the train track? (Drag the cursor to indicate your response)

Forbidden 1 2 3 4 5 6 7 Permissible

Before providing some demographic information about yourself, please think again about the two thinking styles we described at the start of this study. Try to make an image of how the people who have one and the other thinking style. Next, you will see several attributes. Please indicate to what extent you think that each of these attributes describes Global and Detailed perceivers.

Not at all 1 2 3 4 5 6 7 Very much

Global Perceivers Detailed Perceivers

1. Competent
2. Likeable
3. Friendly
4. Honest
5. Intelligent
6. Sincere
7. Warm
8. Skilled
9. Trustworthy

Demographic information

Thank you for participating in this study. Before we end, please answer a few demographic questions.

How old are you?

What is your gender? M/F

Are you a native English speaker? Y/N

Thank you for taking part in this study. Here we explain a little more about the goals of this study and what we expect to find. It also explains a few important things: Information that we gave you that was not entirely truthful. So please read this carefully.

We told you at the start of the study that our goal was to examine how people's visual perception styles can affect how they perceive social situations. In fact, our goal is slightly different: We aim to examine whether individuals' judgements of the moral actions of another individual vary depending on whether or not they belong to the same social group.

To examine this, in the first part of the study, participants performed a task and were subsequently assigned to one of the two groups: Global and Detailed Perceivers. In reality, however, assignment to the groups was unrelated to participants' performance in the task. This was done so that two groups could be experimentally created, that is, so that we could create groups about which participants know very little—which is important for an accurate test of our hypothesis. Then, you were asked some questions to see to what extent you identified with your assigned group. Finally, in the second part, participants judged the actions of either an ingroup or an outgroup target in two moral dilemmas. Our hypothesis is that people will judge the actions of members of their own group as more moral (or permissible) than actions by members of another group.

It is important that you understand that there was a little deceit involved in this study. We apologize for that. However, it is also important that you understand that this was done so that it is possible to research something which is both new and very important: How group membership can affects moral judgements. Thank you for your contribution to this scientific endeavor!

INFORMATION ON SUPPORT FOR DISTRESS

If you feel like you have been feeling upset and emotionally distressed, the following organizations offer advice and support:

- You may find emotional support services near your area through **NHS Mental Health**: (http://www.nhs.uk/conditions/stress-anxiety-depression/pages/low-mood-stress-anxiety.aspx)
- You may also find useful to contact Mental Health Foundation which offers help people in distress and need immediate help in the UK: (https://www.mentalhealth.org.uk/contact)
- You can also access the mental health charity, the Samaritans, through this website: http://www.samaritans.org/how-we-can-help-you/contact-us
- Many colleges and most universities have a free and confidential in-house counselling service, with professionally qualified counsellors and psychotherapists that you can access. You can usually find information about what they offer and how to make an appointment in the counselling service section of your university's website. This free service in universities is available to postgraduates as well as undergraduates.
- If you are from University of Exeter, it has a **Mental Health Team** which can offer free mental health advice and support: (http://www.exeter.ac.uk/wellbeing/mental_health/)
- You could also make an appointment through **Wellbeing** and have a one-to-one meeting with a counsellor. (http://www.exeter.ac.uk/wellbeing/contact/makeanappointment/)

If you have further questions or concerns regarding your participation in this study, or if you wish to receive a brief summary of the results of this study at a later point in time, please contact Prof. Barreto at m.barreto@exeter.ac.uk or Esra Dasci (PhD Candidate) at ed368@exeter.ac.uk

If your concerns are related to the ethical aspects of this research, feel free to contact the Chair of the Psychology Ethics Committee, Dr Lisa Leaver at l.a.leaver@exeter.ac.uk

Appendix B: Exploratory Analysis on Group Perceptions for Studies 1 and 2 (Chapter 3)

B1. Study 1: Do Moral Judgements of Targets Affect Evaluations of Their Group as a Whole?

Analytical strategy. We explored whether judging the morality of a target's actions affects the judgement of the morality of the group as a whole (see Figure 2). We already know that, for one of the scenarios (switch), moral judgements are themselves affected by the group membership of the target and by group identification (see main analyses in Chapter 3, Study 1). The relationship between moral judgements and perceived group morality also needs to be moderated by the group membership of the target, since we only expect perceived group morality to be affected by the moral judgement of a member of that same group. Thus, for moral judgements related to the *switch* scenario (i.e., the scenario for which moral judgements were affected by the independent variables), target group membership is expected to moderate both the relationship between identification and moral judgements of the target *and* the relationship between moral judgement of the target and group morality.

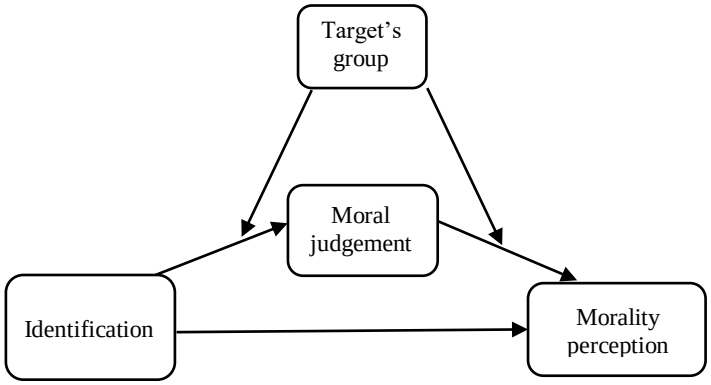

Figure 8.Model specifying the moderated mediation of the effect of group identification (X) on perception of the group's morality (Y), by moral judgements of the target (M), at different levels of target group membership (W).

To test this, for analyses involving the switch scenario, we used model 58 of the PROCESS macro (Hayes, 2016), testing the indirect effect of identification (X) on perceived (in/out)group morality (Y), through moral judgement (M), at different levels of target group

membership (W, coded as -1 = ingroup and +1 = ougtroup). This was done twice, once focusing on perceived *ingroup* morality as the outcome variable and once focusing on perceived *outgroup* morality as the outcome variable.

However, for the footbridge scenario, moral judgements of the target were not affected by target group membership, group identification, or their interaction. Therefore, in this case, we tested merely whether the relationship between moral judgements of targets in the footbridge scenario (X) and perception of group morality (Y) was moderated by the target's group membership (M) by using model 1 of PROCESS macro (Hayes, 2016), while controlling for identification.

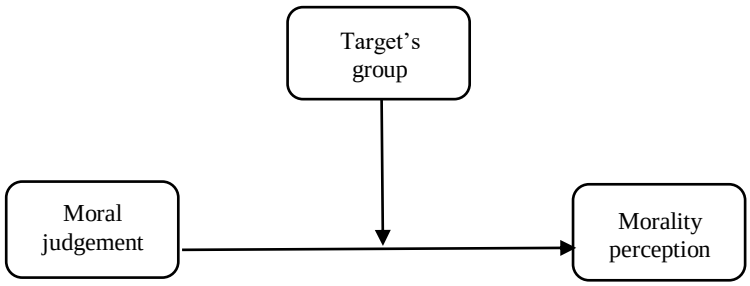

Figure 9. The moderation of the relationship between moral judgements of targets in the footbridge scenario (X) and perception of group morality (Y) by the target's group membership (M).

Perceived ingroup morality. First of all, as we already reported above, we conducted a moderation analysis to see if the effect of target's group membership (X) on moral judgements (Y) is moderated by group identification (M) in the switch scenario. From this analysis, we knew that the effect of target's group membership on moral judgements was depended on levels of group identification (see above).

New to these analyses, the overall model predicting ingroup morality (Figure 2) was also significant, $F(4,86) = 7.13$, $p < .001$, $R^2=.25$. The direct effect of moral judgement on perception of ingroup morality was not statistically significant, $b= .005$, $t(86)= .09$, $p = .93$.

The direct effect of identification on perception of ingroup morality was positive and significant, $b = .26$, $t(86) = 3.40$, $p < .01$. Also the direct effect of target's group membership was statistically significant and positive, $b = .87$, $t(86) = 2.66$, $p < .01$. Finally, there was a significant interaction between moral judgements and target's group membership on perception of ingroup morality, $b = -.18$, $t(86) = -3.02$, $p < .01$, reflecting a positive effect of moral judgement of the target on perception of the ingroup as moral when the target is an ingroup member ($b = .23$, $t(87) = 2.58$, $p = .01$), but not when it is an outgroup member ($b = -.24$, $t(87) = -2.84$, $p = .006$).

The analyses did not reveal a significant moderated mediation, 95% CI [-.05, .24]. However, when we probed the mediation at different levels of the target's group membership, we found that the indirect effect of identification on perception of ingroup morality through moral judgement was significant for those who judged an outgroup target, $b = .11$, 95% CI [.001, .27], but not (opposite to expectation) for participants who judged an ingroup target, $b = .03$, 95% CI [.-02, .15].

For the footbridge scenario, we tested merely whether the relationship between moral judgements of targets and perceptions of ingroup morality was moderated by target's group membership (see Figure 3), while controlling for identification. The overall model was statistically significant, $F(4,86) = 5.36$, $p < .001$, $R^2 = .20$. The direct effect of target's group membership was not statistically significant, $b = .21$, $t(86) = .94$, $p = .35$, and neither was the effect of moral judgements, $b = .06$, $t(86) = 1.09$, $p = .27$. The interaction between target's group membership and moral judgements was also not statistically significant, $b = -.08$, $t(86) = -1.40$, $p = .16$. The only significant effect was the direct effect of identification on perception of ingroup morality, which was positive and statistically significant, $b = .33$, $t(86) = 4.33$, $p < .001$, indicating that the higher the ingroup identification was the more moral the ingroup was perceived as a whole.

Perceived outgroup morality. When perception of outgroup morality was the ultimate outcome, the overall model was significant, $F(4,86) = 3.95$, $p < .01$, $R^2=.15$. Neither the direct effect of moral judgement, $b= .04$, $t(86)= .66$, $p = .51$, nor the direct effect of target group membership, $b= .54$, $t(86)= 1.57$, $p = .12$, on perception of outgroup morality was statistically significant. However, the interaction between moral judgement and target's group membership was significant, $b = -.14$, $t(86) = -2.25$, $p = .03$, reflecting a positive effect of moral judgement of the target on perception of the outgroup as moral when the target is an ingroup member ($b = .21$, $t(87) = 2.34$, $p = .02$), but not when it is an outgroup member ($b = -.14$, $t(87) = -1.64$, $p = .10$).

With regard to the moderated mediation, the results show that the index of moderated mediation was not significant, 95% CI [-.12, .19]. In addition, the indirect effect of identification on perception of outgroup morality through moral judgements was neither statistically significant for those who judged an ingroup, $b = .03$, 95% CI [-.03, .13], nor for those who judged an outgroup target, $b = .07$, 95% CI [-.07, .22].

For the footbridge scenario, again we tested merely whether the relationship between moral judgements of targets and perceptions of outgroup morality was moderated by target's group membership, while controlling for identification. The overall model was statistically significant, $F(4,86) = 2.69$, $p = .04$, $R^2=.11$. As for perceived ingroup morality, there were no significant main effects of target's group membership, $b = -.05$, $t(87)= -.20$, $p = .84$, or moral judgements, $b = -.02$, $t(87)= .31$, $p = .76$. There was also no significant interaction between target's group membership and moral judgements, $b = -.04$, $t(87)= -.73$, $p = .47$. The only significant effect was the direct effect of identification on perception of outgroup morality, which was positive and significant, $b = .20$, $t(86)= 2.57$, $p = .01$, indicating that the higher the ingroup identification the more moral the outgroup was perceived.

B2. Study 2: Do Moral Judgements of Targets Affect Evaluations of Their Group as a Whole?

Analytical strategy. As in the first study, here we also wanted to test if judging the morality of a target's actions affects the judgement of the morality of the group as a whole. In this study we added moral scenario to the model (see Figure 4).

To test this, we used model 68 of the PROCESS macro (Hayes, 2016), testing the indirect effect of identification (X) on perceived (in/out)group morality (Y), through moral judgement (M), at different levels of target group membership (W, coded as -1 = ingroup and +1 = outgroup) in different moral scenarios (Z). This was done separately to predict perceived *ingroup* morality and to predict perceived *outgroup* morality as the outcome variable. Because moral scenario has three levels, it was dummy coded. D1 contrasts judgements in the footbridge (0) scenario with judgements in the switch (1) scenarios (implied consent = 0). D2 contrasts the footbridge (0) with the implied consent (1) scenarios (switch = 0). Additionally, we ran each analysis twice with the relevant dummy as moderator (Z), and the other dummy as a covariate.

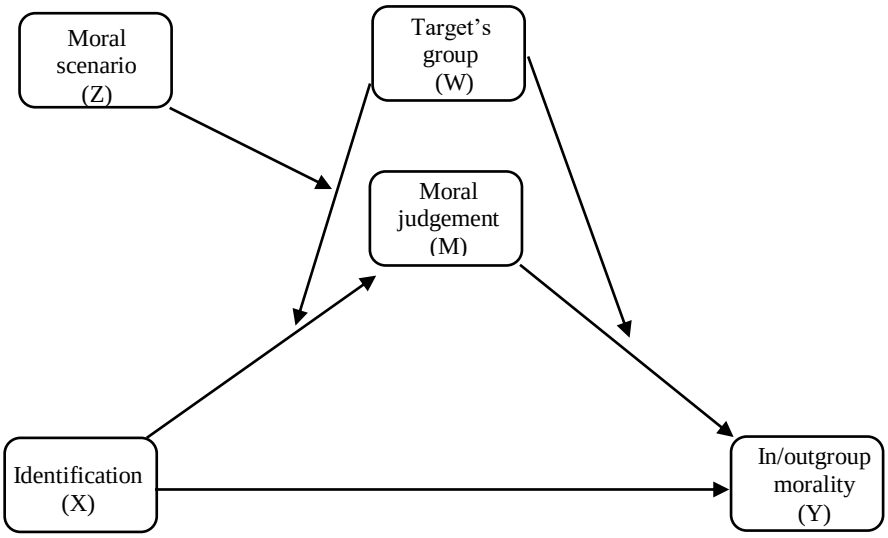

Figure 10.Model specifying the moderated mediation of the effect of group identification (X) on perception of the group's morality (Y), by moral judgements of the target (M), at different levels of target group membership (W) and moral scenario (Z).

Perceived ingroup morality. First of all, as we already reported above, we conducted a moderation analysis to see if the effect of target's group membership on moral judgements is moderated by group identification and moral scenario. From this analysis, we obtained a statistically significant overall model, that if the effect of identification on moral judgement was moderated by target's group membership and the type of moral scenario (D1: footbridge vs. switch), $F(8,269) = 24.41$, $p < .001$, R^2=.45. The direct effect of identification on moral judgement was positive and statistically significant, $b = .24$, $t(269) = 2.53$, $p = .01$, indicating that the higher the identification was the more moral the target was judged. The direct effect of target's group membership on moral judgement was not significant, $b = .47$, $t(269) = 1.07$, $p = .28$. Further, the interaction between identification and target's group membership was not statistically significant, $b = -.07$, $t(269) = -.75$, $p = .45$. The effect of D_1 (footbridge vs. switch) on moral judgement was positive and statistically significant, $b = 3.12$, $t(269) = 4.48$, $p < .001$, indicating that targets in the switch scenario was judged as more moral than those who were in the footbridge. However, the interaction between identification and D_1 was not statistically significant, $b = -.24$, $t(269) = -1.62$, $p = .11$. The interaction between target's group membership and D_1 was negative and marginally significant, $b = -1.34$, $t(269) = -1.94$, $p = .05$, indicating that the outgroup targets were judged to be more moral than the ingroup targets regarding footbridge vs. switch scenarios. Also, the three-way interaction between identification, target's group membership and D_1 was positive and marginally significant, $b = .25$, $t(269) = 1.67$, $p = .09$.

The model for perceived ingroup morality, we obtained an overall significant model, $F(6,272) = 9.15$, $p < .001$, R^2=.17. The direct effect of moral judgement on perceived ingroup morality was not significant, $b = .01$, $t(272) = .19$, $p = .85$. However, the direct effect of identification on perceived ingroup morality was positive and statistically significant, $b = .33$, $t(272) = 6.77$, $p < .001$, indicating that the higher the identification was the more moral the

ingroup was perceived as a whole. The direct effect of target's group membership was not statistically significant, b = -.10, t(272) = -.68, p = .50. Finally the interaction between moral judgement and target's group membership on perceived ingroup morality was not statistically significant, b = .01, t(272) = .33, p = .74.

As we already mentioned above, the direct effect of identification on perceived ingroup morality was statistically significant. However, there were no conditional indirect effects (all 95% CIs included 0) of identification on perceived ingroup morality through moral judgement at different levels of target's group membership and D_1, suggesting that the effect of identification was not transmitted through moral judgement to perceived ingroup morality.

Next, we entered D_2 (footbridge vs. implied consent) in to analysis as our moderator and D_1 as a covariate. When moral judgement was a DV, the overall model was again statistically significant, F(9,269) = 23.29, p < .001, R^2=.44. Neither the main effect of identification, b = .10, t(269) = 1.18, p = .24, nor the main effect of target's group membership, b = .01, t(269) = .03, p = .98, on moral judgement was statistically significant. Also the interaction between identification and target's group membership was not statistically significant, b = .02, t(269) = .29, p = .77. However, the main effect of D_2 on moral judgement was positive and statistically significant, b = 2.58, t(269) = 3.22, p = .001, indicating that targets in the implied consent scenario was judged as more moral than those who were in the footbridge. However, two way interactions (identification*D_2, b = .13, t(269) = .79, p = .43, target's group membership*D_2, b = -.41, t(269) = -.52, p = .61) and the three way interaction (identification*target's group membership*D_2, b = .03, t(269) = .19, p = .85) were not significant.

The model for ingroup morality, we also obtained an overall significant model, F(6,272) = 8.86, p < .001, R^2=.16. However, only the direct effect of identification was

positive and statistically significant, $b = .32$, $t(272) = 6.71$, $p < .001$. The main effects of moral judgement ($b = .03$, $t(272) = .85$, $p = .40$), target's group membership ($b = -.11$, $t(272) = -.71$, $p = .48$) and the interaction between moral judgement and target's group membership ($b = .01$, $t(272) = .34$, $p = .73$) were not significant. The conditional indirect effects of identification on perceived ingroup morality was not statistically significant at any levels of target's group membership and moral scenario (footbridge vs. implied consent).

Perceived outgroup morality. Similarly, we ran moderated mediation analysis twice with the relevant dummy as an IV and the other as covariate to test the moderated mediation effects on perceived outgroup morality. The results was the same with the first part of analysis in perceived ingroup morality (see above).

The model for outgroup morality, we again obtained an overall significant model, $F(6,272) = 7.31$, $p < .001$, $R^2=.14$. The direct effect of moral judgement on perceived outgroup morality was positive and statistically significant, $b = .08$, $t(272) = 2.37$, $p = .02$, indicating that the higher the moral judgement was the more moral the outgroup was perceived. Also the direct effect of identification on perceived outgroup morality was positive and statistically significant, $b = .26$, $t(272) = 5.21$, $p < .001$, indicating that the higher the identification the more moral the outgroup was perceived. However the direct effect of target's group membership on perceived outgroup morality was not statistically significant, $b = -.002$, $t(272) = -.01$, $p = .99$. Finally, the interaction between moral judgement and target's group membership was not statistically significant, $b = .005$, $t(272) = .15$, $p = .88$.

With regard to the moderated mediation analysis, we did not obtain any statistically significant indirect effects of identification on perceived outgroup morality through moral judgement at the different levels of target's group membership and moral scenario.

Next, when we entered D_2 (footbridge vs. implied consent) as our IV and the other dummy as a covariate. Again the overall model was statistically significant, $F(9,269) =$

23.29, $p < .001$, R^2=.44 when moral judgment was a DV. However, there were no significant main or interaction effects on moral judgement (same as results with D_2 in perceived ingroup morality).

When we added perceived outgroup morality into model as a DV, the model was overall significant, $F(6,272) = 7.26$, $p < .001$, R^2=.14. The direct effect of moral judgement on perceived outgroup morality was positive and statistically significant, $b = .11$, $t(272) = 3.50$, $p < .001$, indicating that the higher the moral judgement was the more moral the outgroup was perceived. Also the direct effect of identification was positive and statistically significant, $b = .26$, $t(272) = 5.16$, $p < .001$, indicating that the higher the identification the more moral the outgroup was perceived. However the direct effect of target's group membership on perceived outgroup morality was not statistically significant, $b = .002$, $t(272) = .01$, $p = .98$. Finally, the interaction between moral judgement and target's group membership was not significant, $b = .004$, $t(272) = .12$, $p = .91$.

With regard to the moderated mediation analysis, although there was a statistically significant and positive direct effect of identification on perceived outgroup morality, there were no conditional indirect effects of identification on perceived outgroup morality at any levels of target's group membership and moral scenario (D_2: footbridge vs implied consent).

Appendix C: Implied Consent Scenario and the response scale used in Study 2

L is 27 years old and likes walking in their leisure time. L, who is a global [detailed] perceiver, is taking his daily walk near the train tracks when he notices that the train that is approaching is out of control. L sees what has happened: the driver of the train saw a man walking across the tracks and slammed on the brakes, but the brakes failed and the driver fainted. The train is now rushing toward the man. It is moving so fast that he will not be able to get off the track in time. L is standing next to the man, whom they can throw off the track out of the path of the train, thereby preventing it from killing the man. The man is frail and standing with his back turned. L can throw the man, injuring him; or can refrain from doing this, letting the man die.

To what extent do you think that it is morally permissible for L to throw the man? (Drag the cursor to indicate your response)

Forbidden 1 2 3 4(Permissible) 5 6 7 Praiseworthy

Appendix D: Full Wording of the Scenarios and Measures in Pilot Study

Stealing Medication Scenarios

Ambiguously immoral condition:
M is terminally ill. The doctor has told them that there is one drug that could cure them, but this drug is not freely available and is very expensive, so M cannot afford it. M tried to borrow money in various ways, but they could not get the cost of the medication together. At this point, M decided to break into a pharmacy and steal the medication.

Clearly immoral condition:
M has a non-life threatening illness and has been using medication to control the illness and the associated pain. M has become addicted to some of this medication and cannot afford to buy as much as the addiction requires. M tried to borrow money in various ways, but they could not get the cost of the medication together. At this point, M decided to break into a pharmacy and steal the medication.

Provoking an Accident Scenarios

Ambiguously immoral condition:
When M was passing by, they saw a stranger having a heart attack. M tried to help this person and decided to drive them to the hospital. When rushing to the hospital, M accidentally hit someone who ended up with a broken leg as a result of the accident.

Clearly immoral condition:
M was on their way to an appointment when they realised that they didn't have their glasses. M was late and could not find their glasses quickly enough, so they decided to drive without them. When rushing to their appointment, M accidentally hit someone who ended up with a broken leg as a result of the accident.

To what extent do you think that M's action is morally acceptable?

Unacceptable 1 2 3 4 5 6 7 Acceptable

How sure are you about your judgement of the (un)acceptability of M's actions?

Not at all sure 1 2 3 4 5 6 7 Very sure

To what extent do you think that M is a moral person?

Not at all 1 2 3 4 5 6 7 Very much

Do you think that M is:

Strongly disagree 1 2 3 4 5 6 7 Strongly agree
More likely a man
More likely a woman
Equally likely to be either a man or a woman

Appendix E: Materials Used in Studies 5 and 6 (Chapter 4)

E.1. Scenarios Used in Studies 5 and 6

Ambiguously immoral condition:
When Mary was passing by, she saw a stranger having a heart attack. Mary tried to help this person and decided to drive them to the hospital. When rushing to the hospital, Mary accidentally hit someone who ended up with a broken leg as a result of the accident.

Clearly immoral condition:
Mary was on their way to an appointment when she realised that she didn't have her glasses. Mary was late and could not find her glasses quickly enough, so she decided to drive without them. When rushing to her appointment, Mary accidentally hit someone who ended up with a broken leg as a result of the accident.

Note. Target's name and pronouns in the scenarios are changed to *John (he/his/him)* for male target condition to manipulate target gender in Study 6.

To what extent do you think that Mary's action is acceptable?

Unacceptable 1 2 3 4 5 6 7 Acceptable

To what extent do you think that Mary's action is moral?

Not all 1 2 3 4 5 6 7 Very much

To what extent do you think that Mary is a trustworthy person?

Not all 1 2 3 4 5 6 7 Very much

To what extent do you think that Mary is a reliable person?

Not all 1 2 3 4 5 6 7 Very much

To what extent do you think that Mary is a moral person?

Not all 1 2 3 4 5 6 7 Very much

What is your overall impression of Mary?

Very negative 1 2 3 4 5 6 7 Very positive

E.2. Materials Used in Study 7

Mary [John], who is a mother [father] of two children, has been a full-time employee in the finance department of an energy company, where she [he] has been working for five years. Mary [John] has recently been promoted to a management role, which resulted in a very hectic schedule for her [him]. Therefore, she [he] is not able to spare as much time for her [his] family as she [he] used to. For instance, she [he] missed her [his] kid's school play last week since she [he] had to travel to another city for an important meeting with new business clients on the same day.

To what extent do you think that Mary's action is acceptable?

Unacceptable 1 2 3 4 5 6 7 Acceptable

To what extent do you think that Mary is moral person?

Not all 1 2 3 4 5 6 7 Very much

To what extent do you think that Mary is competent person?

Not all 1 2 3 4 5 6 7 Very much

What is your overall impression of Mary?

Not all 1 2 3 4 5 6 7 Very much

E.3. Ambivalent Sexism Inventory (Glick & Fiske, 1996)

Below is a series of statements concerning men and women and their relationships in contemporary society. Please indicate the degree to which you agree or disagree with each statement using the following scale:

(0) Strongly disagree; (1) disagree somewhat; (2) disagree slightly; (3) agree slightly; (4) agree somewhat; (5) agree strongly.

1. No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman.
2. Many women are actually seeking special favors, such as hiring policies that favour them over men, under the guise of asking for "equality."
3. In a disaster, women ought not necessarily to be rescued before men.
4. Most women interpret innocent remarks or acts as being sexist.
5. Women are too easily offended.
6. People are often truly happy in life without being romantically involved with a member of the other sex.
7. Feminists are not seeking for women to have more power than men.
8. Many women have a quality of purity that few men possess.
9. Women should be cherished and protected by men.
10. Most women fail to appreciate fully all that men do for them.
11. Women seek to gain power by getting control over men.
12. Every man ought to have a woman whom he adores.
13. Men are complete without women.
14. Women exaggerate problems they have at work.

15. Once a woman gets a man to commit to her, she usually tries to put him on a tight leash.
16. When women lose to men in a fair competition, they typically complain about being discriminated against.
17. A good woman should be set on a pedestal by her man.
18. There are actually very few women who get a kick out of teasing men by seeming sexually available and then refusing male advances.
19. Women, compared to men, tend to have a superior moral sensibility.
20. Men should be willing to sacrifice their own wellbeing in order to provide financially for the women in their lives.
21. Feminists are making entirely reasonable demands of men.
22. Women, as compared to men, tend to have a more refined sense of culture and good taste.

Appendix F: Exploratory Analyses for Chapter 4

F.1. Regression analysis for Study 5

Table 17. Regression analysis predicting moral judgements when HS is a moderator in the model in Study 5.

Predictor	*b*	*b* 95% CI [LL, UL]	*beta*	Fit	Difference
(Intercept)	3.00	[2.83, 3.17]			
Hostile Sexism	.001	[-0.09, 0.10]	.00		
Moral Extremity	2.13**	[1.92, 2.33]	.71		
Participant Gender	-.21*	[-0.42, -.01]	-.07		
				R^2 = .496** 90% CI [0.44, 0.55]	
(Intercept)	2.98	[2.79, 3.18]			
Hostile Sexism	-.03	[-0.20, 0.13]	-.02		
Moral Extremity	2.16**	[1.87,2.46]	.72		
Participant Gender	-.17	[-0.46, 0.11]	-.06		
Hostile Sexism x Moral Extremity	.12	[-0.07, 0.32]	.06		
Hostile Sexism x Participant Gender	-.05	[-0.25, 0.14]	-.03		
Moral Extremity x Participant Gender	-.07	[-0.48, 0.34]	.02		
				R^2 = .498** 90% CI [0.44, 0.55]	ΔR^2 = .002 90% CI [-0.005, 0.01]
(Intercept)	2.98	[2.79, 3.18]			
Hostile Sexism	-.03	[-0.23, 0.16]	-.02		
Moral Extremity	2.16**	[1.87,2.46]	.72		
Participant Gender	-.17	[-0.46, 0.12]	-.06		
Hostile Sexism x Moral Extremity	.12	[-0.17, 0.41]	.06		
Hostile Sexism x Participant Gender	-.06	[-0.33, 0.22]	-.03		
Moral Extremity x Participant Gender	-.07	[-0.48, 0.34]	-.02		
Hostile Sexism x Moral Extremity x Participant Gender	.003	[-0.38, 0.39]	.001		
				R^2 = .498** 90% CI [0.44, 0.55]	ΔR^2 = .000 90% CI [-0.001, 0.002]

Note. A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

F.2. Regression analysis for Study 6

Table 18. Regression analysis predicting moral judgments when HS is a moderator in the model in Study 6

Predictor	*b*	*b* 95%CI [LL, UL]	*beta*	*r*	Fit	Difference
(Intercept)	3.00**	[2.86, 3.15]				
Hostile Sexism	.05	[-0.02, 0.12]	.04	.06*		
Moral Ambiguity	2.34**	[2.20, 2.50]	.74	.74**		
Participant Gender	-.16*	[-0.31, -0.02]	-.05	-.05		
Target Gender	-.04	[-0.18, 0.11]	-.01	-.01		
					R^2 = .552** 90% CI [0.51, 0.59]	
(Intercept)	2.98	[2.80, 3.17]				
Hostile Sexism	.06	[-0.07, 0.19]	.04	.06*		
Moral Ambiguity	2.33**	[2.08,2.59]	.74	.74**		
Participant Gender	-.19	[-0.42, 0.05]	-.06	-.05		
Target Gender	.09	[-0.16, 0.34]	.03	-.01		
Hostile Sexism x Moral Ambiguity	-.02	[-0.16, 0.12]	-.01	.03		
Hostile Sexism x Participant Gender	.05	[-0.09, 0.18]	.02	.07*		
Hostile Sexism x Target Gender	-.05	[-0.18, 0.09]	-.02	.03		
Moral ambiguity x Target Gender	-.17	[-0.46, 0.12]	-.04	.40**		
Participant Gender x Target Gender	-.09	[-0.38, 0.20]	-.03	-.03		
Moral Ambiguity x Participant Gender	.16	[-0.13, 0.45]	.04	.46**		
					R^2 = .554** 90% CI [0.52, 0.60]	ΔR^2 = .002 90% CI [-0.003, 0.01]
(Intercept)	3.02	[2.82, 3.22]				
Hostile Sexism	.04	[-0.13, 0.22]	.03	.06*		
Moral Ambiguity	2.25**	[1.96, 2.55]	.71	.74**		
Participant Gender	-.25	[-0.51, 0.01]	-.08	-.05		
Target Gender	.01	[-0.28, 0.30]	.004	-.01		
Hostile Sexism x Moral Ambiguity	-.04	[-0.29, 0.21]	-.02	.03		
Hostile Sexism x Participant Gender	.03	[-0.19, 0.25]	.02	.07*		
Hostile Sexism x Target Gender	.03	[-0.21, 0.26]	.01	.03		
Moral ambiguity x Target Gender	.002	[-0.43, 0.44]	.001	.40**		
Participant Gender x Target Gender	.05	[-0.34, 0.44]	.01	-.03		
Moral Ambiguity x Participant Gender	.30	[-0.10, 0.70]	.08	.46**		
Hostile Sexism x Moral Ambiguity x Participant Gender	0.10	[-0.18, 0.37]	.03	-.06		
Hostile Sexism x Moral Ambiguity x Target Gender	-.09	[-0.36, 0.19]	-.03	.02		
Moral Ambiguity x Target Gender x Participant Gender	-.31	[-0.90, 0.28]	-.06	.26**		
Hostile Sexism x Participant Gender x Target Gender	-.06	[-0.34, 0.22]	-.02	.04		
					R^2 = .555** 90% CI [0.52, 0.60]	ΔR^2 = .001 90% CI [-0.003, 0.005]

(Table 18 cont.)						
(Intercept)	3.02	[2.82, 3.22]				
Hostile Sexism	.04	[-0.14, 0.22]	.03	.06*		
Moral Ambiguity	2.25**	[1.96, 2.55]	.71	.74**		
Participant Gender	-.25	[-0.51, 0.01]	-.08	-.05		
Target Gender	.01	[-0.28, 0.30]	.004	-.01		
Hostile Sexism x Moral Ambiguity	-.03	[-0.32, 0.26]	-.01	.03		
Hostile Sexism x Participant Gender	.04	[-0.21, 0.28]	.02	.07*		
Hostile Sexism x Target Gender	.03	[-0.23, 0.30]	.01	.03		
Moral ambiguity x Target Gender	.002	[-0.43, 0.44]	.001	.40**		
Participant Gender x Target Gender	.05	[-0.34, 0.44]	.01	-.03		
Moral Ambiguity x Participant Gender	.30	[-0.10, 0.70]	.08	.46**		
Hostile Sexism x Moral Ambiguity x Participant Gender	0.08	[-0.29, 0.46]	.03	-.06		
Hostile Sexism x Moral Ambiguity x Target Gender	-.10	[-0.51, 0.31]	-.03	.02		
Moral Ambiguity x Target Gender x Participant Gender	-.31	[-0.90, 0.28]	-.06	.26**		
Hostile Sexism x Participant Gender x Target Gender	-.07	[-0.44, 0.30]	-.02	.04		
Hostile Sexism x Moral Ambiguity x Participant Gender x Target Gender	.03	[-0.53, 0.58]	.01	-.05		
					R^2 = .555**	ΔR^2 < .001
					90% CI [0.52, 0.60]	90% CI [-0.001, 0.001]

Note. A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. R^2 requires a confidence coefficient of (1 - 2α) if we are to infer statistical significance ($p < .05$) from an interval that does not contain zero – i.e., 90% (not 95%) confidence intervals for R^2 correspond to the traditional .05 criterion of statistical significance. * indicates $p < .05$. ** indicates $p < .01$.

F.3. Decomposing the three-way interaction in Study 5 (Chapter 4)

We decomposed the marginally significant three way interaction further for exploratory purposes. To do so, we ran a moderation analysis to test whether the effect of moral ambiguity (X) on moral judgements (Y) was moderated by benevolent sexism (M) for each gender group separately. We used Model 1 in the PROCESS macro created by Hayes (2016) entering benevolent sexism as a centred continuous variable.

For female participants, the overall model was significant, $F(3,230)=73.930$, $R^2 = .491$, $p < .001$. The main effect of benevolent sexism score on moral judgement was not significant, $b=.045$, $t(230)=.375$, $p = .708$. There was a positive and significant main effect of moral ambiguity on moral judgements, $b=2.101$, $t(230)=14.442$, $p < .001$, but the predicted two-way interaction between moral ambiguity and benevolent sexism was not significant, $b=-.064$, $t(230)=-.406$, $p = .685$.

For male participants, the overall model was significant, $F(3,200)=80.546$, $R^2 = .547$, $p < .001$. The main effect of benevolent sexism on moral judgements was positive and significant, $b=.456$, $t(200)=4.433$, $p < .001$. There was also a positive and significant main effect of moral ambiguity on moral judgements, $b=2.133$, $t(200)=14.685$, $p < .001$. Finally, the predicted interaction between moral ambiguity and benevolent sexism was negative and significant, $b=-.428$, $t(200)=-2.845$, $p = .005$. However, the interaction pattern was not as predicted, as it indicated that the higher the benevolent sexism was the higher the moral judgements of female targets in non-ambiguous condition (see Figure 10).

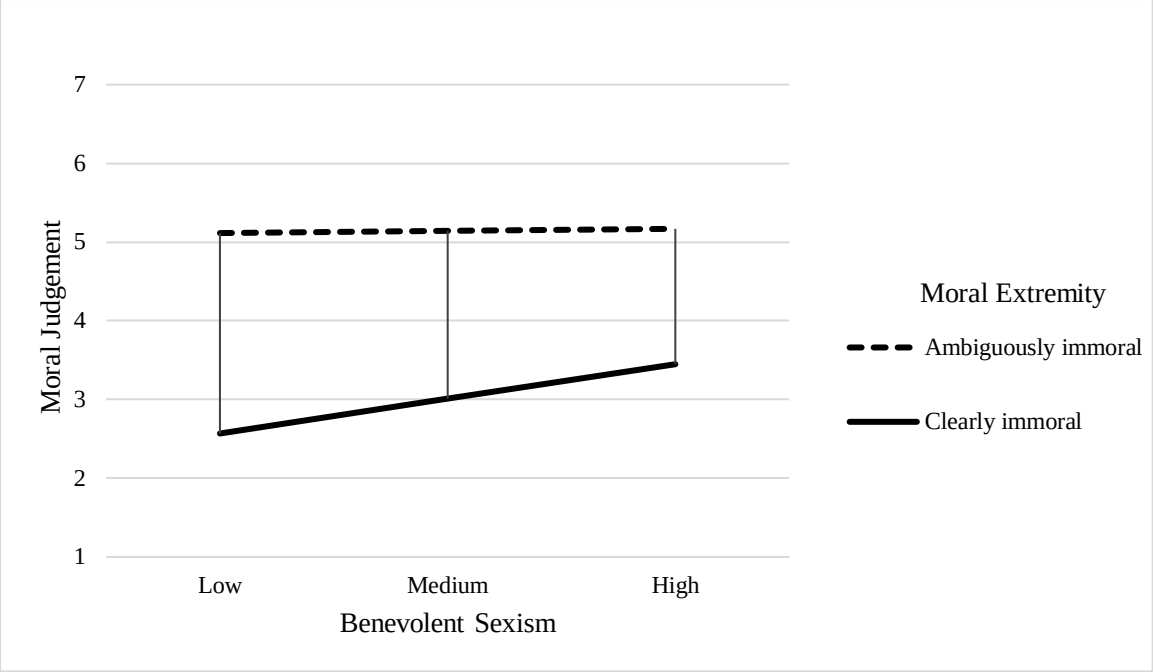

Figure 11. Moderating effect of benevolent sexism on moral judgements at the levels of moral ambiguity for male participants in Study 6.