



Universiteit  
Leiden  
The Netherlands

## **Moral character judgments and motivated cognition in legal reasoning**

Strohmaier, N.; Jong, S.W.M. de; Brigaglia, M.; Roversi, C.

### **Citation**

Strohmaier, N., & Jong, S. W. M. de. (2023). Moral character judgments and motivated cognition in legal reasoning. In M. Brigaglia & C. Roversi (Eds.), *Diritto & questioni pubbliche* (pp. 227-248). Palermo: Diritto e questioni pubbliche & Recognise. Retrieved from <https://hdl.handle.net/1887/3728330>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3728330>

**Note:** To cite this publication please use the final published version (if applicable).

# Moral Character Judgments and Motivated Cognition in Legal Reasoning

**NIEK STROHMAIER**

*Leiden University*

**SOFIA DE JONG**

*Leiden University*

## ABSTRACT

When deciding a legal case, how do judges and jurors reach their conclusions? Ideally, legal decision makers analyze the case's facts and apply the relevant laws and regulations in a formalistic fashion without being distracted by legally irrelevant information. However, research in the field of psychological science suggests this is not necessarily how legal decision making works in practice. Instead, it appears that legal decision makers can be affected by a host of unconscious biases, allowing irrelevant information to creep in. In this chapter, we discuss a specific source of unwarranted subjectivity in legal decision making that stems from moral considerations. Moral judgments and intuitions can automatically and unconsciously drive sense-making processes in such a way that a biased interpretation of the facts and legal rules allow legal decision makers to reach their desired conclusion. This so-called 'motivated reasoning' or 'motivated cognition' entails the process in which people are unconsciously motivated to reach a certain conclusion, while operating under the illusion of objectivity. This chapter (1) introduces motivated cognition, (2) reviews how this process can affect legal decision making, and (3) discusses the role played by moral character inferences in motivated legal reasoning. This chapter ends with a discussion on potential debiasing strategies and implications for legal practice.

## KEYWORDS

motivated reasoning, motivated cognition, moral character, bias, legal decision making

## CITATION

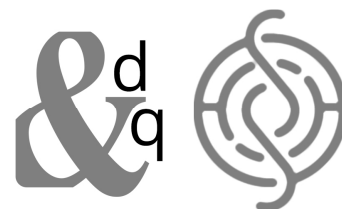
Strohmaier N., de Jong S. 2023. *Moral Character Judgments and Motivated Cognition in Legal Reasoning*, in Brigaglia M., Roversi C. (eds.), *Legal Reasoning and Cognitive Science: Topics and Perspectives*, «Diritto & Questioni pubbliche», Special Publication, August 2023.

**DIRITTO & QUESTIONI PUBBLICHE / RECOGNISE**

*Legal Reasoning and Cognitive Science: Topics and Perspectives*  
Special Publication / August, 2023, pp. 227-248

© 2023, *Diritto e questioni pubbliche*, Palermo.

ISSN 1825-0173



# Moral Character Judgments and Motivated Cognition in Legal Reasoning

NIEK STROHMAIER, SOFIA DE JONG

1. Introduction – 2. Motivated cognition in (legal) reasoning – 2.1. What is motivated cognition? – 2.2. Motivated cognition in legal reasoning – 2.2.1. Culpable Causation – 2.2.2. Cognitively cleansing tainted evidence – 2.2.3. The harm principle – 2.2.4. Motivated enforcement of phantom rules – 2.2.5. Motivated cognition and the advocacy bias in lawyers – 2.3. Motivated cognition in laypeople versus legal experts – 2.4. Reality constraints – 2.5. Summary and concluding remarks – 3. Moral judgments and the role of moral character inferences – 3.1. Introduction – 3.2. Dual-process models of moral judgment – 3.3. Biased information models of moral judgment – 3.4. Person-centered approach to moral judgment – 3.5. Empirical support for the person-centered approach to moral judgment – 3.5.1. Empirical support in the context of criminal law – 3.5.2. Empirical support in the context of civil law – 3.5.3. The side-effect effect in legal judgments – 3.6. Summary and concluding remarks – 4. Debiasing the influence of moral character inferences and motivated legal reasoning – 4.1. Requirements to reduce moral bias in judgments – 4.2. Debiasing through procedural constraints – 4.2.1. The prohibition of character evidence – 4.2.2. Linear sequential unmasking – 4.2.3. Limiting court access of litigating parties – 4.3. Individual differences in susceptibility to bias – 4.3.1. Intellectual humility – 4.3.2. Actively open-minded thinking – 4.3.3. Free will beliefs – 4.4. Summary and concluding remarks – 5. Conclusion

## 1. Introduction

The idea that legal judgments do not always follow the ideal mechanistic pattern in which a judge or jury applies the relevant laws to the facts in a particular case to render a verdict is not new. At the beginning of the 20<sup>th</sup> century, legal realism already posited that judges «decide not primarily because of law, but based (roughly speaking) on their sense of what would be “fair” on the facts of the case» (LEITER 2003, 50). Moreover, realists have argued that «legal rules and reasons figure simply as *post-hoc* rationalizations for decisions reached on the basis of non-legal considerations» (LEITER 2003, 50; see also GREEN 2005). Psychological research has also devoted ample attention to factors that can affect legal judgments and decision making in ways that are typically deemed undesirable. Consider for example the role of emotions<sup>1</sup>, political orientation (e.g., BRAMAN 2009; EPSTEIN & KNIGHT 2013; FURGESON et al. 2008), and irrelevant defendant characteristics like race and gender (e.g., LINDHOLM & CEDERWALL 2010; YOURSTONE et al. 2008). The role of bias in legal decision making is also addressed in other chapters in this book.

In this chapter, we zoom in on two particular phenomena that are directly related to biases in legal judgments. The first is *motivated reasoning* or *motivated cognition*. Through this process, certain irrelevant information or unconscious motives drive the sense-making process of legal decision makers in such a way that this allows them to reach their preferred outcome. The second phenomenon discussed in this chapter is the role of moral character inferences in legal judgments. How, when, and why do people in general and legal decision makers in particular judge the moral character of defendants and other parties involved in legal disputes? We discuss how the judgments about a defendant’s moral character drives legal decision makers to a preferred outcome, which then, through motivated cognition, drives their legal analyses in a biased fashion.

<sup>1</sup> For a review of the literature on the role of emotion in legal decision making, see FEIGENSON 2016.

The structure of this chapter is as follows: First, we introduce the relevant theories surrounding motivated cognition and how this process can affect legal decision making (sect. 2). Second, we discuss recent theorizing in moral psychology about the central role of character inferences in moral judgments (sect. 3) and the role of moral character inferences in motivated legal reasoning (sect. 4). Finally, we briefly review research on potential debiasing techniques and the possible implications for legal practice (sect. 5).

## 2. *Motivated cognition in (legal) reasoning*

### 2.1. *What is motivated cognition?*

Motivated cognition or motivated reasoning can be described as the process in which decision makers engage in «inadvertently biased processes for accessing, constructing, and evaluating beliefs» (SOOD 2013, 309; KUNDA 1990, 480) when they unconsciously have a preference regarding the outcome of an evaluative task. For example, when a judge is faced with a defendant charged with manslaughter and that judge has a strong and automatic intuition—not based on the evidence—that the defendant is guilty, this judge could be more likely to perceive and weigh the evidence in such a manner that it allows for a conviction. In contrast, if the judge has a gut feeling that the defendant is innocent, the theory of motivated cognition would suggest that the judge would perhaps be more critical of incriminating evidence, more receptive to exculpatory circumstances, or spend more time critically analyzing the forensic evidence proffered by the prosecution. In other words, the desired outcome (conviction or acquittal) will guide the sensemaking process and the weighing of the evidence by the judge or jury.

This is supported by research showing that decision makers have strong preferences for one conclusion over another. For example, preferences stemming from the evaluation of people as immoral or moral, and thus having «an affective stake in reaching a certain conclusion», can strongly alter reasoning processes in order to align the moral judgment with that desired conclusion (DITTO et al. 2009). Furthermore, these preferences can, for example, strongly influence the evidence a decision maker uses for an evaluation and the criteria for the evaluation itself (DITTO et al. 2009).

It is important to highlight the unconscious nature of motivated cognition. Decision makers are deemed to be under “the illusion of objectivity” when analyzing information and reasoning their way towards a desired conclusion (SOOD 2013, 309). Hence, legal decision makers may not be consciously aware of their own intuition regarding a specific defendant, nor that they may unconsciously be working towards a certain outcome. In a landmark paper describing the phenomenon of motivated cognition, Kunda describes the unconscious nature of motivated cognition as follows (KUNDA 1990):

«People do not realize that the process is biased by their goals, that they are accessing only a subset of their relevant knowledge, that they would probably access different beliefs and rules in the presence of different directional goals, and that they might even be capable of justifying opposite conclusions on different occasions».

It is therefore clear that the process of motivated reasoning does not entail the conscious or strategic updating of attitudes, beliefs, or convictions in order to justify holding a position that serves the holders’ interest. Conversely, legal decision makers can typically be deemed to act in good faith and thus to strive towards analyzing the relevant facts as unbiasedly and objectively as possible. Yet, despite their best efforts, research suggests that motivated reasoning processes are frequently at play in legal decision making. The next section discusses evidence for

motivated cognition in legal reasoning. It will become clear that motivated cognition plays a role in several facets of legal reasoning, in fact it is a widespread phenomenon.

## 2.2. Motivated cognition in legal reasoning

### 2.2.1. Culpable Causation

Several studies have investigated motivated cognition in legally relevant contexts. In Alicke's classic study, although the term motivated cognition was not used *per se*, he did show how moral judgments based on legally irrelevant information can subsequently drive legally relevant judgments concerning causality and responsibility (ALICKE 1992). He presented jury-eligible participants with a case involving a traffic accident in which a speeding car T-boned another car when crossing an intersection, resulting in significant injuries for the driver in the other car. Several other contributing factors were at play, such as a stop sign that was largely covered by overhanging tree branches, an oil spill making the road slippery, and the fact the T-boned car ignored a red traffic light. Participants were asked to what extent the speeding driver was the cause of the accident and whether this driver was responsible for the injuries suffered by the other driver.

The crux of the experiment was the reason *why* the speeding driver was exceeding the speed limit while approaching the intersection. Half of the participants read that John (the driver of the speeding car) had to get home quickly to hide a vial of cocaine before his parents could see it. The other half of the participants were told that John had to get home quickly to hide an anniversary present for his parents. In other words, half of the participants believed John had a socially desirable motive for speeding whereas the other half believed he had a socially undesirable motive.

The results showed that participants in the socially undesirable motive condition of the experiment believed more strongly that John was the primary cause of the accident and also that John was judged as more responsible, compared to those in the socially desirable motive condition. This study provides strong evidence for the notion that moral considerations can drive legally relevant judgments. More specifically, if a person is presented as "morally bad" their role will be evaluated as being causally more important.

### 2.2.2. Cognitively cleansing tainted evidence

A series of studies by Sood provide more direct evidence of motivated reasoning in legal judgments. She tested whether jury-eligible participants would judge a particular piece of illegally obtained evidence as admissible if that evidence would help convict a defendant who engaged in an egregious crime, despite it formally being inadmissible based on the exclusionary rule (SOOD 2015). The exclusionary rule under US law holds that evidence obtained in an unlawful manner is inadmissible in court. An exception to this rule is when illegally obtained evidence would have inevitably come to light at a later stage through lawful means.

In the case presented to participants, police officers conducted an illegal search of a car, in which they found large quantities of drugs. Importantly, however, half of the participants were told that the police had discovered «bags of marijuana that the defendant had been selling to terminally ill cancer patients to ease their suffering», whereas the other half were told that the police had discovered «bags of heroin and needles that the defendant had been selling to high school students». Participants in the heroin version of the case gave significantly higher punishment recommendations than those in the marijuana version of the case. Moreover, the heroin group believed more strongly that the illegally obtained evidence should be admissible in court. They based their answer on their conviction that the evidence would ultimately have been discovered lawfully, and were thus using the inevitable discovery exception to justify their belief that the incriminating evidence should be admissible. This study again clearly shows that

when people have a strong moral reaction towards a person having engaged in an egregious crime, they are motivated to construe the available information in such a way that it permits punishment of the criminal act.

### 2.2.3. *The harm principle*

Motivated cognition has been also observed in the application of the harm principle. The harm principle states that «the State should use its powers to regulate individual conduct only if doing so is necessary to prevent harm to others»<sup>2</sup>. Sood tested whether people motivated to criminalize a particular act would report that the act does causes harm to others, even though the act clearly does not cause any harm.

Participants were presented with a case involving a naked man going to the supermarket. In pretests it had been established that this was perceived as not causing any harm to others, but that it deviated from social norms to the degree that people would want to criminalize the act regardless of the absence of harm (SOOD & DARLEY 2012). Half of the participants were told that in order to criminalize a certain behavior, it would need to be established that the behavior causes harm, while the other half of the participants received no instructions.

It was expected that all participants would want to criminalize this socially unacceptable behavior, but that only those participants who were informed about the harm principle would be motivated to construe the act as causing harm. The result confirmed the hypotheses, showing that all participants criminalized the act of going to the supermarket naked to the same degree, but the informed group reported higher degrees of harm than the group unaware of the harm principle. These results provide further evidence for the notion that motivated cognition can cause people to perceive elements of a case in such a way that it allows them to justify their punitive inclination, even if that justification requires construal of harm.

### 2.2.4. *Motivated enforcement of phantom rules*

Another example of motivated cognition has been found in the selective enforcement of legal rules. Legal rules should be universally applicable to all subjects within a particular jurisdiction. However, given the same transgression of a certain rule, it is not uncommon that some are punished whereas others are let off the hook. A particular subclass of legal rules where unequal treatment is frequently observed are “phantom rules”; rules that are «frequently broken and rarely enforced» (WYLIE & GANTMAN 2023, 2). Examples of these are rules on things like jaywalking and downloading music. Punishment in cases where these rules are violated is largely at the discretion of legal decision makers who can take the relevant circumstances into consideration in their judgment. The resulting ambiguity opens the door for motivated reasoning in such a way that these phantom rules are more likely to be enforced when decision makers are motivated by reasons that should not enter any legal analysis.

In a series of experiments, Wylie and Gantman tested this hypothesis. They presented participants with a case involving a man smoking marijuana on a park bench. Smoking marijuana is considered to violate a phantom rule, given its use is illegal, yet this law is infrequently and inconsistently enforced. Half of the participants learned that this man tried to provoke a stranger, thereby breaking a social norm but not a legal norm. The other half learned that the man politely asked a stranger what time it was, thus neither breaking a social nor legal norm. All participants were then asked whether it would be justified for the police to start an interaction with the man and for him to be punished.

<sup>2</sup> SOOD & DARLEY 2012, 1357; they refer to John Stuart Mill’s *On Liberty* (1859) when discussing the harm principle.

The hypothesis was that participants informed of the provocation would be more likely to enforce the phantom rule and thus indicate that it would be justified for the police to engage with the man and punish him. The results confirmed that those participants believed it was more justified to enforce the phantom rule for the social norm violating man than for the man who did not break any social norm. Collectively, the experiments show that motivated cognition can cause people to use an opening in the law to punish a person for unrelated behavior such as breaking social norms.

### 2.2.5. *Motivated cognition and the advocacy bias in lawyers*

One of the key players in the legal domain is the lawyer. Lawyers are obliged to serve their clients' needs and to act in their interest. Therefore, an inherent part of lawyers' work is the motivation to advocate for the cause that meets their clients' needs. However, lawyers are taught that they can and have to form and maintain unbiased beliefs while advocating their client's cause (MELNIKOFF & STROHMINGER 2020, 1261) so a "blind" adoption of their clients' perspective would be detrimental to an accurate assessment of their chances in court. The question therefore is whether professional lawyers can remain impartial and unbiased or whether lawyers are also susceptible to motivated cognition processes in such a way that advocating a certain position can alter their true beliefs (also called *advocacy bias*). In other words, are lawyers able to control when to stay objective and when to advocate for a certain position?

Melnikoff's and Strohminger's research shows that professional lawyers are not as "in control" as they are taught and thought to be (MELNIKOFF & STROHMINGER 2020). Across a series of experiments, participants were asked to either act as a defense or prosecuting attorney and were given a range of different cases. Ultimately, the experiments aimed to test the automaticity and controllability of the advocacy bias, by (i) using weak stimuli to test whether a minimally immersive context can already cause advocacy bias, (ii) focusing on strongly held beliefs to test whether advocacy bias can alter such strong beliefs, (iii) by instilling a motivation to be accurate and objective to test whether advocacy bias can be eliminated or reduced when participants were required to be unbiased, and (iv) by using participants trained to be impartial (i.e., professional lawyers). Collectively, the results provide robust evidence for the notion that advocating for a certain position can automatically and uncontrollably alter even strongly held beliefs, even among professional lawyers.

### 2.3. *Motivated cognition in laypeople versus legal experts*

Thus far, we have seen that motivated cognition can affect judgments concerning for example causality, responsibility, and punishment. Furthermore, it can motivate the recruitment of harm if punishment requires it, it can result in the cognitive cleansing of tainted evidence, and it can affect the enforcement of phantom rules. Above all, we have seen that motivated cognition happens automatically and is hard to control.

However, the extent to which these findings can be generalized to legal experts has to be questioned, given that the majority of the studies used jury-eligible participants (i.e., laypeople). This is particularly relevant for the many jurisdictions that do not have jury trials and thus fully rely on judges. To what extent can legal expertise be a safeguard against the unwarranted influence of legally irrelevant information on legal judgments?

A number of studies on biases in legal judgments suggest that judges may be less biased than laypeople. For example, Rachlinski et al. found that specialized judges (bankruptcy judges) were unaffected by omission bias, debtors' race, or apologies, all of which have been shown to affect laypeople (RACHLINSKI et al. 2007; see also RACHLINSKI et al. 2006). In a similar line of research, it was shown that even though judges do rely heavily on their intuitive faculties, thus

making them susceptible to biases, they did manage to resist hindsight bias (i.e., the overestimation of the foreseeability of past events)<sup>3</sup>.

However, despite the limited number of studies showing that judges' training and experience help them avoid certain biases, other studies show that legal experts (e.g., lawyers and judges) are affected by cognitive biases to a similar extent as non-experts. For example, the study by Melnikoff and Strohminger on the automaticity and uncontrollability of the advocacy bias found strong evidence for this bias among professional lawyers (MELNIKOFF & STROHMINGER 2020). Moreover, Guthrie and colleagues investigated whether judges are swayed by classic biases such as anchoring effects, framing effects, and hindsight bias, and found this to be the case (GUTHRIE et al. 2000). Similar results were found for professional arbitrators (HELM et al. 2016) and for insolvency law experts (STROHMAIER et al. 2021). Furthermore, research has shown that judges struggle to a similar extent as jurors with ignoring inadmissible information, even when they are reminded about the inadmissibility of certain evidence and even when they themselves ruled the evidence as inadmissible<sup>4</sup>.

In a study directly comparing laypeople with professional judges in terms of their susceptibility to bias, both groups were presented with a case in which the mayor of a beach town commissioned the construction of a new highway (KNEER & BOURGEOIS-GIRONDE 2017). However, half of each group were told that as a side effect, the surrounding environment would be harmed, whereas the other half were told the environment would actually benefit. Both groups were informed the mayor did not care that the environment would be harmed/helped and that he just wanted the new road to be built.

A common finding is that people attribute intentionality for harmful side effects, but not for beneficial side effects. Thus, the question put to the participants was whether they believed the mayor intentionally harmed/helped the environment. The authors found the same pattern for both laypeople and judges, providing evidence for the notion that judges are also susceptible to this so-called "side-effect" effect. Results of a follow-up study varying the severity of the harm caused to the environment as a result of building the new road showed that judges were affected by outcome bias. That is, judges believed the mayor intentionally harmed the environment more strongly when the environment was harmed to a more extreme extent versus when the environment was harmed relatively mildly. As a final example of motivated cognition in legal experts, a recent study investigated whether experts' legal judgments concerning foreseeability, blame, and punishment in the context of directors' liability were affected by irrelevant information about the character of company directors. Results showed that both laypeople and legal experts were equally affected by this irrelevant information, even though laypeople did appear to be more punitive (STROHMAIER et al. forthcoming).

To conclude, it is safe to assume that professional legal decision makers such as judges are not exempt from the pervasive influence of motivated reasoning processes in legal judgments. Be that as it may, future research could help delineate under what circumstances the risk of motivated cognition is highest and when legal expertise may actually provide for some kind of buffer. These insights will also help design effective interventions. In the next section, we review existing knowledge on the boundary conditions of motivated cognition.

#### 2.4. Reality constraints

From the first sections, the reader may feel that judges and jurors reach the conclusion they desire, unconsciously and automatically construing and/or altering whatever they need to

<sup>3</sup> GUTHRIE et al. 2007; for an introductory paper on hindsight bias, see ROESE & VOHS 2012; for a paper on hindsight bias in a legal context, see, e.g., STROHMAIER et al. 2021.

<sup>4</sup> WISTRICH et al. 2005; for another study on judges and inadmissible evidence, see LANDSMAN & RAKOS 1994.

justify their preferred outcome. However, research shows that motivated cognition is limited to what can still be considered reasonable and realistic. In other words, decision makers only alter their beliefs and attitudes to reach a desired outcome that can be reasonably justified to themselves and others (HSEE 1996). Sood also argues that legal decision makers do not ignore the law in order to reach their desired conclusions and that the process of motivated cognition occurs mostly if the law leaves room to do so, for example if the law is unclear or if open norms need to be interpreted and applied (SOOD 2013, 311).

Motivated cognition also appears to be “need-based”, meaning that beliefs, attitudes, and perceptions are only ‘stretched’ to the extent needed to reach the preferred conclusion (BOINEY et al. 1997). People generally strive to act as rationally and objectively as possible, at least when acting in good faith, and it is therefore unlikely that legal decision makers bend reality to the extent that it conflicts with limits set by law or their inner desire to act rationally. Or, as Sood notes, legal decision makers «engage in motivated cognition—unknowingly processing information in an outcome driven manner—to achieve their desired result, seemingly within the terms of the given legal doctrine» (SOOD 2015, 1562).

However, in a recent study, the researchers examined whether having a strong goal or motive to adopt a certain belief could perhaps break through these reality constraints (STROHMINGER & MELNIKOFF manuscript). The authors assigned participants to either the role of defense attorney or prosecuting attorney and presented them with a case involving a defendant who was unambiguously guilty. Since the experiment left no “wobble-room” regarding the defendant’s guilt, the idea of reality constraints on motivated reasoning would predict that all participants would agree that the defendant was guilty. However, if the goal of advocacy was so strong that it could override reality constraints, perhaps participants assigned to advocate the position of the defense attorney would still question the defendant’s guilt.

The results showed that advocacy, as a strong motivator, can affect perceptions regarding seemingly unambiguous facts, even when questioning the defendant’s guilt required giving at least some weight to absurd theories, such as that the defendant was abducted and cloned by aliens and it was the clone who committed the crime, or that a shape-shifting creature took on the defendant’s form and then committed the crime. Hence, even though motivated cognition generally remains within the bounds of what can be considered reasonable, when there is a strong motivation to reach a certain conclusion, it seems that reality constrains the motivated construction of certain beliefs less than once assumed.

## 2.5. Summary and concluding remarks

Thus far, we have introduced the theory of motivated cognition and given examples of studies in the legal domain showing that motivated reasoning can automatically and unconsciously alter legal decision maker’s beliefs and attitudes, which then ultimately can affect legally relevant judgments concerning causality, foreseeability, responsibility, and blame. We also discussed whether there is a meaningful difference between laypeople and experts when it comes to their susceptibility to motivated cognition, and concluded that it is safe to assume that legal expertise and training provide insufficient protection. Finally, we described the constraints typically put on motivated cognition, as decision makers tend to stay within the limits of the law and also within the boundaries of what can be considered reasonable. At the same time, though, we noted recent research that shows that these reality constraints can be “overruled” under certain circumstances, further highlighting the pervasive impact of motivated cognition.

Further investigations into what exactly causes that initial desire to reach a certain conclusion in the first place are needed, answering questions like, what exactly is the instigator of motivated cognition? As we have seen thus far, the starting point for motivated reasoning processes is often a strong moral reaction in response to misconduct or other perceived moral transgressions. It appears

that it is primarily our need to blame and punish wrongdoers and thus our initial moral judgment regarding someone's character that drives these processes. That is, does someone have a good or bad moral character and is this person therefore a worthy recipient of praise or blame? To better understand how these moral character inferences can motivate legal reasoning, in the next section we discuss how moral judgments take shape, firstly exploring moral judgments in a general sense and then by focusing on the role of moral character inferences.

### 3. *Moral judgments and the role of moral character inferences*

#### 3.1. *Introduction*

When we judge the blameworthiness of people's actions, we typically take several factors into account that we consider relevant. Take the example of Pete who harmed Henry, a poodle puppy, while playing fetch. One relevant factor in determining Pete's blameworthiness would be whether Pete intentionally harmed Henry, or whether it was a mere accident. Another relevant factor would be whether it was actually the stick that Pete threw to Henry that hit the puppy in the head, or whether it was the stick thrown by another dog owner playing with their dog nearby. Yet another relevant factor would be whether it was reasonably foreseeable that playing fetch with a stick could result in Henry being hit in the head. As a final example, a relevant factor would be whether it was within Pete's power to avoid the harm from occurring, for example by choosing a different object to play fetch with or using a smaller stick. In other words, when determining a person's blameworthiness, ideally we carefully analyze relevant factors such as causality, intentionality, foreseeability, and controllability.

Early models of moral judgment adhered to a prescriptive approach as they described how moral judgments are ideally made (SHAVER 1985). Since then, however, theories of moral judgment have sought to better understand how moral judgments are actually made in real life and key factors such as causality, mental state (e.g., was the harm intentional?), and preventability were identified (see, e.g., CUSHMAN 2008).

In the next sections, we introduce models of moral judgment that incorporate both automatic, intuitive processes, as well as more deliberate and cognitively intense processes (so-called *dual-process* models). We then review models that focus primarily on the quick, affective, and intuitive reactions towards moral transgressions, which generally attribute less significance to more conscious deliberations. Finally, we discuss recent developments in research on the role of moral character inferences in moral judgment and highlight the centrality and importance of such inferences. We close with a discussion on the empirical support for the prominent role of moral character inferences in moral judgment.

#### 3.2. *Dual-process models of moral judgment*

Prominent models of moral judgment attribute qualities to both reason and emotion. Specifically, dual-process models stipulate that moral judgments can be the result of two distinct modes of operation. On the one hand there is the quick, automatic, affective mode, and on the other there is the more deliberate, slow, cognitive mode. A well-known dual-process model is one put forward by Greene, which is based on an extensive line of research involving moral dilemmas (GREENE et al. 2001; see also GREENE & HAIDT 2002).

Greene presented participants with classic moral dilemmas such as being able to save five lives by having to push another person onto a train track, thereby stopping the train from running over the five people laying on the track. Both behavioral and neurological data show that more affective modes of processing typically result in deontological judgments (i.e.,

following clear moral rules such as “killing is wrong”) regarding the permissibility of killing one person to save five. In contrast, more cognitive, deliberate modes of processing tend to result in consequentialist judgments, meaning that the permissibility of an action is judged based on the consequences, rendering the killing of one person to save five morally permissible.

However, it has been argued that Greene’s model is better equipped to predict judgments of moral permissibility and less so for attributing (legal) blame (MONROE & MALLE 2017, 123). Moreover, recent insights in neuroscience suggest that there are more than two processes in moral judgment, meaning that dual-process models focusing on the distinct tracks of intuition and emotion and on deliberate reasoning are an oversimplification of how moral judgments are actually formed in real life (VAN BAVEL et al. 2015). Rather than being separate streams, it seems more plausible that rational deliberation and emotion interact in a variety of different ways (HELION & PIZARRO 2015).

The Path Model of Blame, coined by Malle, Guglielmo, and Monroe, permits a more complex interplay between emotion and reason, with moral judgments at times being fast and automatic, and at other times more effortful and deliberate<sup>5</sup>. The model’s main aim is to describe *what* exactly is being processed, rather than through which *mode* (i.e., automatic vs. deliberate). The model posits that upon the detection of a norm violating act, the first factor to be analyzed is whether the agent is the cause of a certain harm. If not, no blame is assigned. If causality is established, the next step is to determine the degree of intentionality. If an agent intentionally caused a certain harm, the final degree of blame assigned to the agent is dependent on the reasons for this intention. If the agent did not intend to cause the harm, the degree of blame is dependent on the extent to which the agent had the duty to prevent the harm and whether the agent was capable of doing so.

A limitation of this model is that it does not account for the severity of a transgression (GOODWIN 2014). Ample research on outcome bias has shown that with all else being equal, moral transgressions are judged harsher when the outcome is more severe (see, e.g., KNEER & MACHERY 2019; KNEER & SKOCZEŃ 2023; KNEER 2022). Moreover, in the model, the degree of blame assigned to a person is the final “product” of processing the factors included in the model. Alternative theories, that will be discussed next, claim that the attribution of blame actually happens relatively fast, automatically, and intuitively, and subsequently affects perceptions of causality, intent, and other factors deemed relevant by the Path Model of Blame (GOODWIN 2014, 217 f.). A key criticism of the Path Model of Blame, therefore, is that it assigns insufficient weight to the role of motivated reasoning processes initiated by quick and intuitive blame judgments (NADLER 2014). We now focus on several theories of moral judgment that give more primacy to quick and automatic moral intuitions and subsequent motivated reasoning processes.

### 3.3. *Biased information models of moral judgment*

Several theories of moral judgment assign more weight to the role of emotion and intuition than the dual-process models discussed thus far. These “emotion heavy” theories do not go as far to say that conscious reasoning plays no role in moral judgment, but, in the more “extreme” versions of these theories, conscious deliberation mostly serves to *post-hoc* rationalize the judgment that was instantly formed through affective and intuitive processes. A prominent theory in this domain is that of Jonathan Haidt, called “social intuitionism” (HAIDT 2001).

In its essence, the social intuitionist model holds that moral judgment is mostly based on intuitions instead of on moral reasoning (HAIDT 2001, 1024). The idea is that when learning of a moral transgression, people experience a quick and automatic moral intuition, which is then

<sup>5</sup> MONROE & MALLE 2017, 124; for their first paper positing the model see, MALLE et al. 2014.

potentially, but not necessarily, followed by motivated reasoning in order to rationalize the initial intuition (HAIDT 2001, 817). Hence, according to Haidt's theory, there is a direct link between a person's moral intuition and their moral judgment, without an intermediary phase of conscious reasoning. Haidt states that, in contrast to the rationalist models, «judgment comes first, based on educated intuition; justification is undertaken next» (HAIDT 2012, 868). Consistent with this theory, studies on “moral dumbfounding” show that people can have a strong moral reaction towards certain acts and also condemn these acts, without being able to give a good explanation for their moral judgment (see e.g., HAIDT & HERSH 2001). Evidence for Haidt's theory comes largely from the studies on motivated cognition discussed in Section 2.

Alicke's culpable control model, which we have touched on earlier, is largely compatible with Haidt's theory. Alicke suggests that social groups have a need to hold transgressors accountable in order to maintain social order (ALICKE 2000). This need is reflected in humans' natural inclination to blame. Alicke found that this inclination to blame affects the evaluation of culpable control. When people are motivated to blame an individual, they are more likely to conclude that that individual had causal control over a certain outcome. This is why the speeding driver who wanted to hide a vial of cocaine being judged to be more causally responsible for the traffic accident than the speeding driver trying to get home to hide his parent's anniversary present. In Alicke's theory, it is the affective reaction and spontaneous evaluations of an agent and their behavior that directly affects blame judgments, while at the same time allowing for an indirect link between affect and blame through the evaluation of mental state and causal control.

Thus far, we show that moral judgments can, for an important part, be the result of automatic, affective reactions triggered by witnessing a harmful event. The theories differ in terms of the role of conscious deliberations being either limited to mere *post-hoc* justification of the initial blame judgment, or to rational deliberation that is influenced by affective processes through a complex interplay between the two. Hence, moral judgments are at the very least influenced by quick and automatic affective reactions, and are potentially their sole determinant. But what is the source and focus of those initial affective reactions and moral intuitions? What precisely triggers people to have a certain moral reaction? Recent theorizing in moral psychology suggests that moral character inferences take center stage, to which we turn next.

### 3.4. *Person-centered approach to moral judgment*

The theories discussed thus far focus mainly on moral judgments in response to a certain act. Hence, the blameworthiness of an agent is derived from elements of the act, such as whether someone acted intentionally, whether the harm caused by the act was foreseeable, whether the act was in fact the cause of a certain harm, etc. Such act-based theories of moral judgment, however, cannot account for a range of studies, some of which have already been discussed, that show that the same act is judged differently based on factors independent from the actual norm-violating act. For instance, Alicke's study, in which he shows that the same act (speeding whilst approaching an intersection, resulting in a collision) is judged differently depending on the social acceptability of the driver's motive for speeding.

As an alternative to the act-based theories of moral judgment, there are the more recent person-centered theories, where people primarily ask themselves whether a person is good or bad, rather than whether a certain act is right or wrong (UHLMANN et al. 2015). From an evolutionary perspective, it makes sense that we are quick to judge whether someone is a friend or a foe, as it is adaptive to quickly know whether someone will either be an ally in achieving our goals or a threat to our survival. Indeed, evaluating a person's moral character can be crucial and it is therefore important to judge correctly. Think about the horrific opening scene of the movie *Silence of the lambs* in which a girl kindly accepts to help a man who turns out to be a serial killer. Albeit extreme, this example shows the practical importance of getting our moral

character evaluations right. Given the evolutionary advantage of assessing people's character swiftly and accurately, we have evolved to do so rather aptly.

This idea of moral character inferences taking center stage in our moral judgments is consistent with findings from research on impression formation. Studies on impression formation show that evaluating a person's moral character is the first thing we do when forming impressions of others, and that this process occurs automatically and can already be observed at an early age and across cultures<sup>6</sup>. Recent research has even shown that inferences of a person's moral character are a more important driver in impression formation than attributes such as a person's warmth and competence, which until recently were believed to be the two key dimensions in person evaluations (BRAMBILLA et al. 2021). Based on the primacy of moral character inferences, person-centered theories argue that descriptive models of moral judgment should include moral character evaluations as a central feature (see GOODWIN et al. 2014; GOODWIN 2015).

In short, based on recent theorizing, it appears that people form quick and automatic judgments of a person's moral character, and these judgments are expected to then guide their sensemaking process. In this way, a person with a bad moral character is expected to have acted more intentionally when causing a certain harm than one with a good moral character. Hence, rather than focusing on the blameworthiness of the act, the act is used to infer a person's moral character, which then informs our moral judgments.

### 3.5. *Empirical support for the person-centered approach to moral judgment*

Empirical support for the centrality of moral character inferences in moral judgments is plentiful. We have repeatedly referred to Alicke, who demonstrates that judgments about blame and causality are strongly influenced by moral attributes of the speeding driver, even though all the factors considered relevant in prescriptive models of moral judgment were kept constant across the different experimental conditions. It appears that participants judged the moral character of the speeding driver with the socially undesirable motive to be worse than that of the other with the socially desirable motive. This evaluation of the speeding driver's moral character is expected to then have influenced participants' perceptions of the causal role of the driver in the accident.

#### 3.5.1. *Empirical support in the context of criminal law*

In the context of criminal law, research has been conducted on whether jurors follow the ideal model of legal decision making in which relevant factors such as causality, intentionality, and foreseeability are carefully weighed, or whether jurors are affected by irrelevant information about a defendant's moral character, as predicted by the person-centered theories of moral judgment (NADLER & MCDONNELL 2011). In a series of experiments, participants were presented with a case in which all legally relevant factors were constant, such as the (severity of the) harm that occurred and the mental state of the defendant, but in which the moral character of the defendant varied.

For example, in one of the experiments, participants read a case about a woman living with her two dogs that sometimes escaped from her fenced yard and behaved aggressively towards children. At a certain point, the dogs escape again, reap havoc in the neighborhood, and ultimately attack two young boys, one of whom dies from his injuries. Those in the good moral character condition read that the owner of the dogs is a very social person who maintains a healthy lifestyle and has many close friends and adores her two young nieces. The other half read that the owner does not socialize much and spends much of her time watching trash tv

<sup>6</sup> For further relevant citations on the topic, see UHLMANN et al. 2015, 74.

while smoking and eating junk food, and does not like to spend time with her two nieces.

After reading the case, participants were asked to what extent the dogs' owner was responsible for the death of the young boy, to what extent she was the cause of the death, how much blame she deserved, and how foreseeable the death of the boy was from her perspective. It is clear that the participants who read the good moral character version of the case will have a more favorable impression of the woman's character<sup>7</sup>, and that the provided character information should have no bearing on any legally relevant judgment. Still, the results showed that the overall responsibility for the boy's death was rated higher by those who read the bad moral character version of the case and the causal role of the dogs' owner was believed to be higher.

In a follow-up study, participants were presented not with just one moral character version of the case (i.e., either the good character version or the bad character version), but with both (NADLER 2012). This was done to test whether jurors believe that moral character inferences *should* in fact influence their legal judgments, or whether they believe that normatively speaking, character information should be disregarded. If for example a participant first read the aggressive dog case in which the owner has a good moral character and, directly after, read the same case but then with the owner having a bad moral character, and the participant still believes the morally bad dog owner deserves more blame, it would then appear that jurors believe moral character information is somehow relevant for their legal judgment. If, however, the effect of moral character on legally relevant judgment disappears when jurors read both versions, this would suggest that jurors agree that irrelevant character information is irrelevant for legal judgments. The results of this study consistently showed that the biasing effect of moral character inferences disappears when participants are presented with both the good and the bad moral character version of the case, suggesting that jurors can be affected by irrelevant character information, but rightly believe they ought not to (see also Alicke & Zell 2009).

### 3.5.2. Empirical support in the context of civil law

A recent study was set up to extend the findings discussed thus far to the context of civil law instead of criminal law, and to legal experts instead of laypeople (STROHMAIER et al. forthcoming). This study also found evidence for irrelevant moral character inferences biasing legal judgments (STROHMAIER et al. forthcoming). Specifically, legal professionals were presented with a case concerning a company in financial distress where the director was faced with a range of difficult decisions whilst trying to save his company from bankruptcy. In the end, the company fails, and questions arise about the director's liability for damages incurred by creditors.

In this specific legal context, the foreseeability of the company's bankruptcy is important for determining liability, as well as the director's actual awareness of the likelihood the company was going to fail. Importantly, as in the studies discussed thus far, participants either read that the director had a bad or good moral character, for example he was a bad husband and absent father, or in contrast a very loving husband very much involved in his children's upbringing. The information provided about the director's character was clearly irrelevant from a legal point of view.

The results showed that, relative to the legal professionals who read the good moral character version of the case, those who read the bad moral character version believed that the company's bankruptcy was more foreseeable, that the director was aware of the likelihood of bankruptcy, and also believed the director deserved more blame for the damages suffered by creditors. It thus seems that the influence of moral character inferences in legal judgments stretches beyond the domain of criminal law and also affects legal experts.

<sup>7</sup> The data confirmed that participants in the good moral character condition indeed viewed the woman's character more positively than participants in the bad moral character condition, see NADLER & MCDONNELL 2011, 286.

### 3.5.3. *The side-effect effect in legal judgments*

Another line of research in which judgments of an agent's moral character have been shown to affect legally relevant judgments, such as those of intentionality and blame, concerns the so-called "Knobe effect", also termed the "side-effect effect". In the original study, participants read a short case about a chairman of a company's board being approached by an executive about a new program designed to significantly increase profits, whilst also affecting the environment (for the original study, see KNOBE 2003). The chairman then expresses a total disregard for the environment and states to be solely interested in making as much profit as possible. The new program is launched and the environment is affected as predicted. Crucially, however, half of the participants were told that as a side effect of running the new program, the environment would be harmed. The other half were told that the environment would actually benefit from the program. Either way, the chairman did not care. Depending on the version of the case given to them, participants were then asked to indicate whether they believed the chairman intentionally harmed or helped the environment. Respondents generally agreed that the chairman intentionally harmed the environment, but did not believe the chairman intentionally helped the environment. This asymmetry is puzzling, given that it is clearly stated that the chairman is only interested in increasing profits and does not care about how the new program affects the environment. Explaining the results through the lens of moral character inferences, it seems reasonable to assume that people generally formed a negative impression of the chairman's moral character, given we typically believe people ought to at least have some regard for the environment and put environmental concerns in the scale when making business decisions. The fact that the participants were probably motivated to blame the chairman for harming the environment (albeit as a side effect of his true intentions) and reluctant to give praise for helping the environment, it makes sense they indicated the chairman did intentionally harm the environment, but were reluctant to say the chairman intentionally helped the environment. Hence, rather than intentionality judgments serving as key input for blame judgments, it appears that evaluations of an agent's moral character trigger the need to blame the agent, which then drives perceptions of intentionality<sup>8</sup>.

### 3.6. *Summary and concluding remarks*

In this section, we reviewed the workings of moral judgments in order to better understand how legal decision making can be affected by motivated reasoning processes. We introduced several models of moral judgments, ranging from those in which conscious and deliberate processes are given significant weight, to those in which emotions and automatic moral intuitions take center stage and in which conscious deliberations are primarily for *post-hoc* rationalizations. Furthermore, we discussed the dominant role of moral character inferences in our moral intuitions. We argue that it is typically these quick and automatic moral character inferences that trigger the need to blame morally bad people and that it is this need to blame that subsequently drives important constituents of blame (e.g., intentionality, causality, control, etc.). Thus, instead of moral and legal judgments being the final product of a careful weighing of relevant factors, it is more likely that automatic inferences about people's moral character bias our sensemaking processes in such a way that it allows us to blame and hold liable morally bad people, and to go easy on people we perceive to be morally good.

Having explained how motivated cognition can unconsciously drive legal reasoning, and having discussed how moral character inferences are often the key drivers behind motivated

<sup>8</sup> For empirical support for the claim that moral character inferences are an important driver of the side-effect effect, see, e.g., SRIPADA & KONRATH 2011. See also SRIPADA 2012.

legal reasoning, an important question arises: how can we improve legal decision making in order to limit the unwanted influence of moral biases and motivated cognition? The final part of this chapter addresses this question.

#### 4. *Debiasing the influence of moral character inferences and motivated legal reasoning*

The question of what can be done to limit or even prevent the unwanted influence that moral character information can have on legal reasoning is difficult to answer. Current research on potential debiasing techniques has not yet provided direction, however, the problems outlined in this chapter beg for at least some consideration of the current state of knowledge on these techniques.

##### 4.1. *Requirements to reduce moral bias in judgments*

In order for people to reduce the influence of moral biases in moral and legal judgments, at least four requirements need to be met (NADELHOFFER 2006, 212). First, awareness must be created about a particular psychological process a person needs to protect themselves against. Second, a person needs to be motivated to counter the unwanted effects. Third, in order to effectively correct for the unwanted effect, people need to be aware of «the direction and magnitude of the bias» (WILSON & BREKKE 1994, 118). And finally, people must have sufficient cognitive control to effectively prevent a particular bias from exerting an effect.

For legal practitioners, the first requirement is often already absent, as they are generally unfamiliar with the body of literature on unconscious bias in legal reasoning and decision making. Furthermore, legal professionals commonly believe that their expertise and experience protects them from being affected by biases. This first problem could in theory be relatively easily overcome through training and education; when it comes to decision making by judges or jurors, the first three steps could be achieved through proper training in combination with clear (jury) instructions, even though the effects of the latter have been mixed<sup>9</sup>. The major hurdle to overcoming bias is the fourth requirement, as people generally lack the cognitive control to consciously counter effects of bias (NADELHOFFER 2006, 212).

##### 4.2. *Debiasing through procedural constraints*

Therefore, rather than relying on the motivation or willpower of people to overcome their own biases, it may be worthwhile to search for solutions in the procedural domain.

###### 4.2.1. *The prohibition of character evidence*

An example of constraints in legal procedures that are (in part) meant to prevent irrelevant information from biasing legal decision making is the prohibition of character evidence in legal proceedings. In many common law jurisdictions (e.g., US, UK, Australia), there are restrictive rules for the admissibility of character evidence, based on the assumption that such evidence may be given disproportionate weights by juries<sup>10</sup>. Unfortunately, however, these regulations have many loopholes, meaning that in practice character evidence is often deemed admissible (CULBERG 2009; MELILI 1998). And even when character evidence is ruled inadmissible, research

<sup>9</sup> For studies in the effectiveness of jury instructions, see, e.g., HALVERSON et al. 1997; CONKLIN 2021; INGRISELLI 2014; DAFTARY-KAPUR et al. 2010; SAUERLAND et al. 2020; SOOD 2015; GREENE & DODGE 1995.

<sup>10</sup> For a comprehensive historical account, see SEVIER 2019.

has shown that it is difficult to ignore this inadmissible evidence once it has been registered (DAFTARY-KAPUR et al. 2010; STEBLAY et al. 2006). Moreover, and as discussed, people cannot help but form quick and intuitive judgments about people's moral character, meaning that even in the absence of character evidence as such, it is likely that legal decision making will form an opinion of a defendant's character as either good or bad.

#### 4.2.2. *Linear sequential unmasking*

A more promising avenue of procedural constraints can be found in the field of forensic science where a procedure has been introduced called "linear sequential unmasking" (DROR et al. 2015; STOEL et al. 2014). In short, a first investigator (the case manager) has access to all information available in a particular case, part of which is irrelevant and has biasing potential. This first investigator then filters that information and only provides the relevant information to a second investigator needed to conduct a particular part of an investigation. For example, when analyzing a fingerprint found on a crime scene, information that can unconsciously motivate an investigator to find matching elements with the fingerprint of a particular suspect ought to be filtered out and thus not passed on the investigator conducting that analysis. Applied to the context of legal decision making, a form of trial bifurcation in which an initial judge who is given access to all case files and other sources of (irrelevant) information filters the irrelevant information for a second judge may prove promising. After all, information with biasing properties that is never registered cannot exert any effect. A clear downside is that any form of trial bifurcation puts significant demands on courts that typically have limited resources and a significant backlog of cases.

#### 4.2.3. *Limiting court access of litigating parties*

An extreme example of procedural constraints would be to limit judges' and jurors' access to information about the litigating parties by, for example, not allowing parties to appear in court. In some (parts of) legal proceedings, parties do not appear in court and judges/jurors only base their legal decisions on written documents. This mostly applies to civil cases<sup>11</sup>. However, typically, both parties do appear in court, meaning the judge and jury can see what they look like, how they dress, how they talk, how they conduct themselves, etc. All of these factors evidently affect decision makers' perceptions and judgments. Seen from the goal of accuracy and objectivity, it would be better if judges and jurors would have no access to this type of information that might bias them in a certain direction.

From a procedural fairness point of view, however, it is typically considered to be important and desirable for parties to be present in court and to have their say; this is considered a fundamental principle of procedural law. Moreover, it can be desirable and more efficient for judges/jurors to be able to ask questions directly instead of having to request a reaction in writing. It is therefore unlikely that the right (or obligation) to appear in court will ever be revoked. However, in line with the notion of linear sequential unmasking, perhaps the veil of ignorance could be lifted later in the process to allow the majority of a legal proceeding to take place with as little biasing information as possible. Still, as explained above, it would also be necessary to filter some information from the written documents, as written information also has biasing potential.

<sup>11</sup> For example, the Dutch Supreme Court typically does not see parties in person, but only decides on their written statements.

### 4.3. *Individual differences in susceptibility to bias*

In addition to looking at procedural interventions as a way to debias, to what extent can we derive valuable insights from studies looking at individual differences in susceptibility to biases? Why are some people more affected by certain biases than others? What can we learn from these individual differences?

#### 4.3.1. *Intellectual humility*

Recently, increased attention has been devoted to the notion of “intellectual humility”. This «involves recognizing that there are gaps in one’s knowledge and that one’s current beliefs might be incorrect» (PORTER et al. 2022). What makes intellectual humility (IH) conceptually distinct from, for example, perspective-taking and general modesty is that IH focuses on recognizing a person’s own fallibility and ignorance. In simple terms, people high in IH typically acknowledge that there are many things they do not know and that they may even be wrong about the things they believe they do know.

The past decade has seen a steep increase in research on this topic and, for the present purposes, an important consequence of IH is thought to be a reduced susceptibility to certain biases (PORTER et al. 2022; BOWES et al. 2022). This makes sense, given that IH allows a person to be intellectually flexible and not to hold strictly to certain beliefs or attitudes. It follows that people high in IH will experience a reduced motivation to reach a certain outcome, relative to those low in IH. Therefore, it might be a solution to select people high in IH or train people in IH. However, a possible drawback is that it is likely to be hard for someone low in IH to become high in IH, if only because it would require a significant time investment<sup>12</sup>.

#### 4.3.2. *Actively open-minded thinking*

A construct similar to IH, and one that can serve as a prescriptive model for rational thinking that can be consciously adopted, is that of “actively open-minded thinking” (AOT). AOT, as a style of thinking, «includes the tendency to weigh new evidence against a favored belief, to spend sufficient time on a problem before giving up, and to carefully consider the opinions of others in forming one’s own» (HARAN et al. 2013, 189). Or as others have defined it:

«a thinking disposition encompassing the cultivation of reflectiveness rather than impulsivity; the desire to act for good reasons; tolerance for ambiguity combined with a willingness to postpone closure; and the seeking and processing of information that disconfirms one’s beliefs» (STANOVICH & TOPLAK 2019, 156).

Importantly, AOT has been associated with a reduced susceptibility to (among others) belief bias (see e.g., MACPHERSON & STANOVICH 2007), which is roughly similar to confirmation bias and myside bias, both of which entail a tendency to evaluate evidence in such a way that it favors a person’s existing opinions and beliefs.

For an important part AOT is considered to be a predisposition, meaning that is a relatively stable trait and thus that some are more likely to adopt this style of thinking and reasoning than others. However, it has been argued that AOT can also be consciously adopted by adhering to a set of principles. These include, for example, principles such as “correct conclusions are more likely when more than one possible conclusion is evaluated”, “evaluations of possible

<sup>12</sup> For a discussion of ways to improve one’s intellectual humility, see PORTER et al. 2022, 531 f.

conclusions will be more accurate when both positive and negative evidence is sought in a balanced way”, and “evidence, once found and evaluated, should be used in a balanced way, that is, independent of whether it favors or opposes the front runner”. As such, descriptive elements of AOT can be used in a prescriptive way and thus serve as a good standard for evaluating rational thinking in oneself and others.

#### 4.3.3. *Free will beliefs*

In a recent paper on another, related example of individual differences in susceptibility to bias, it was shown that legal professionals who believed more strongly in the notion of free will also showed an increased hindsight bias relative to those who were more skeptical of the idea that humans have free will (STROHMAIER et al. 2021). The proposed mechanism underlying this finding was thought to be that believing in free will correlates with punitive inclinations and reduced IH. Thus, those scoring high on the belief in free will scale were believed to have a stronger need to condemn a potential wrongdoer and would be less comfortable with the uncertainty and ambiguity resulting from withholding judgment.

The idea of withholding judgment and engaging in further exploration has recently also found support in popular media, as evidenced by Julia Galef’s book *The Scout Mindset* (GALEF 2021). The author argues that we can prevent ourselves from falling into the trap of bias and in particular from motivated reasoning through adopting a scout mindset. A scout mindset is loosely defined as deriving pleasure from seeking new information, from challenging even our deeply held beliefs, and as being comfortable with being wrong as our self-worth is then not tied to our being right or wrong. These ideas are largely based on the literature discussed in this section.

Hence, it appears that in order to reduce susceptibility to (moral) biases and motivated cognition, it helps to withhold moral judgment, gather additional information, stress test our convictions, and only be confident in our judgment once we have actively and thoroughly engaged in a fact-finding mission. In other words, the scout mindset, intellectual humility, and actively open-minded thinking (as overlapping concepts) can help protect against a range of moral biases and prevent engaging in motivated (legal) reasoning.

#### 4.4. *Summary and concluding remarks*

To summarize, due to our limited cognitive control, it is difficult for us to overcome the automatic and unconscious processes of motivated moral cognition. Thus, even when we are made aware of the risk of unconscious bias and are motivated to correct for them, actually doing so may prove difficult. Solutions are therefore more likely to be found in changing legal procedures, particularly by adopting those that prevent biasing information from being observed in the first place. Additionally, striving towards becoming more intellectually humble and adopting an actively open-minded thinking style can help reduce the tendency of holding on to previously held beliefs. This may, in turn, stimulate us to actively engage in a critical review of a range of different conclusions, ultimately reducing the chance of falling prey to motivated reasoning.

### 5. *Conclusion*

In this final part of this chapter, we conclude by highlighting the important process of motivated reasoning in legal judgments and the crucial role of moral character inferences in motivated legal reasoning. The key takeaways are the following.

First, legal decision making is unlikely to follow the mechanistic ideal as much as is often believed or hoped. Instead, it seems more likely that motivated reasoning plays an important part in how judges and jurors reach their verdicts.

Second, an important starting point for these motivated reasoning processes is the automatic and instant evaluations of defendants' moral characters. When being presented with information about a defendant's character (e.g., through seeing them in court, pre-trial publicity, character evidence, etc.), legal decision makers cannot help but form an opinion of a defendant's moral character automatically and swiftly. This character evaluation then biases sensemaking processes in such a way that they construe a case in a way that allows them to allot blame to a morally bad defendant and to go easy on a morally good defendant, for example through their attributions of intentionality, causality, control, and foreseeability (i.e., through motivated reasoning).

A third takeaway is that legal expertise is unlikely to prove effective in providing any protection against the automatic and unconscious influence that moral character inferences and blame intuitions have on legal judgments (through motivated reasoning). We reviewed the literature focusing on finding differences between laypeople and legal experts in terms of their susceptibility to a range of different biases, and we can safely conclude that legal experts are equally affected by unwanted psychological influences in their reasoning and judgments. This is an important point as legal professionals still brush off the risk of bias in their decision making and hide behind the argument that they are professionals and therefore not a likely to fall victim to processes affecting the general public. This position is thus not supported by insights derived from scientific research.

Finally, it should have become clear that debiasing the automatic and pervasive influence of motivated reasoning stemming from moral character inferences is a great challenge. It is unlikely we can rely on people's willpower and cognitive control to simply will their way out of biases affecting judgments. Instead, it is probably better to search for solutions in the domain of procedural adjustments that limit the amount of biasing information available to legal decision makers. However, procedural adjustments are costly and will put extra pressure on the heavy workload of courts.

As a more explorative avenue for debiasing potential, we discussed the potential of exploring individual differences in thinking styles and intellectual humility. Limiting motivated cognition is a major challenge as it requires a change within the decision makers themselves. Increasing their intellectual humility and adopting an actively open-minded thinking style may significantly limit the degree to which they engage in motivated reasoning. However, given that change is difficult for anyone as well as being time consuming, it remains to be seen how feasible this strategy really is. Awareness of motivated reasoning processes surely is the first step, and we hope that this chapter will inspire a future generation of legal scholars and practitioners to take up the challenge of improving legal judgments by incorporating insights from behavioral sciences.

## References

- ALICKE M.D. 1992. *Culpable Causation*, in «Journal of Personality and Social Psychology», 63, 3, 368 ff.
- ALICKE M.D. 2000. *Culpable Control and the Psychology of Blame*, in «Psychological Bulletin», 126, 4, 556 ff.
- ALICKE M.D., ZELL E. 2009. *Social Attractiveness and Blame*, in «Journal of Applied Social Psychology», 39, 9, 2089 ff.
- BOINEY L.G., KENNEDY J., NYE P. 1997. *Instrumental Bias in Motivated Reasoning: More When More Is Needed*, in «Organizational Behavior and Human Decision Processes», 72, 1, 1 ff.
- BOWES S. M., COSTELLO T.H., LEE C., MCELROY-HELTZEL S., DAVIS D. E., LILIENFELD S.O. 2022. *Stepping Outside the Echo Chamber: Is Intellectual Humility Associated with Less Political Myside Bias?*, in «Personality and Social Psychology Bulletin», 48, 1, 150 ff.
- BRAMAN E. 2009. *Law, Politics, and Perception: How Policy Preferences Influence Legal Reasoning*, University of Virginia Press.
- BRAMBILLA M., SACCHI S., RUSCONI P., GOODWIN G.P. 2021. *The Primacy of Morality in Impression Development: Theory, Research, and Future Directions*, in «Advances in experimental social psychology», 67, 187 ff.
- CONKLIN M. 2021. *I Knew It All Along: The Promising Effectiveness of a Pre-Jury Instruction at Mitigating Hindsight Bias*, in «Baylor Law Review», 74, 307.
- CULBERG D. 2009. *The Accused's Bad Character: Theory and Practice*, «Notre Dame Law Review», 84, 3, 1343 ff.
- CUSHMAN F. 2008. *Crime and Punishment: Distinguishing the Roles of Causal and Intentional Analyses in Moral Judgment*, in «Cognition», 108, 2, 353 ff.
- DAFTARY-KAPUR T., DUMAS R., PENROD S.D. 2010. *Jury Decision-Making Biases and Methods to Counter Them*, in «Legal and Criminological Psychology», 15, 1, 133 ff.
- DITTO P.H., PIZARRO D.A., TANNENBAUM D. 2009. *Motivated Moral Reasoning*, in «Psychology of Learning and Motivation», 50, 307 ff.
- DROR I.E., THOMPSON W.C., MEISSNER C.A., KORNFELD I., KRANE D., SAKS M., RISINGER M. 2015. *Context Management Toolbox: A Linear Sequential Unmasking (LSU) Approach for Minimizing Cognitive Bias in Forensic Decision Making*, in «Journal of Forensic Sciences», 60, 4, 1111 ff.
- EPSTEIN L., KNIGHT J. 2013. *Reconsidering Judicial Preferences*, in «Annual Review of Political Science», 16, 11 ff.
- FEIGENSON N. 2016. *Jurors' Emotions and Judgments of Legal Responsibility and Blame: What Does the Experimental Research Tell Us?*, in «Emotion Review», 8, 1, 26 ff.
- FURGESON J.R., BABCOCK L., SHANE P.M. 2008. *Do a Law's Policy Implications Affect Beliefs about Its Constitutionality? An Experimental Test*, in «Law and Human Behavior», 32, 219 ff.
- GALEF J. 2021. *The Scout Mindset: Why Some People See Things Clearly and Others Don't*, Penguin.
- GOODWIN G.P. 2014. *How Complete Is the Path Model of Blame?*, in «Psychological Inquiry», 25, 2, 215 ff.
- GOODWIN G. P. 2015. *Moral Character in Person Perception*, in «Current Directions in Psychological Science», 24, 1, 38 ff.
- GOODWIN G.P., PIAZZA J., ROZIN P. 2014. *Moral Character Predominates in Person Perception and Evaluation*, in «Journal of Personality and Social Psychology», 106, 1, 148 ff.

- GREEN M.S. 2005. *Legal Realism as Theory of Law*, in «William & Mary Law Review», 46, 6, 1915 ff.
- GREENE E., DODGE M. 1995. *The Influence of Prior Record Evidence on Juror Decision Making*, in «Law and Human Behavior», 19, 1, 67 ff.
- GREENE J.D., HAIDT J. 2002. *How (and Where) Does Moral Judgment Work?*, in «Trends in Cognitive Sciences», 6, 12, 517 ff.
- GREENE J.D., SOMMERVILLE R.B., NYSTROM L.E., DARLEY J.M., COHEN J.D. 2001. *An fMRI Investigation of Emotional Engagement in Moral Judgment*, in «Science», 293, 5537, 2105 ff.
- GUTHRIE C., RACHLINSKI J.J., WISTRICH A.J. 2000. *Inside the Judicial Mind*, in «Cornell Law Review», 86, 777 ff.
- GUTHRIE C., RACHLINSKI J.J., WISTRICH A.J. 2007. *Blinking on the Bench: How Judges Decide Cases*, «Cornell Law Review», 93, 1 ff.
- HAIDT J. 2001. *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, in «Psychological review», 108, 4, 814 ff.
- HAIDT J. 2012. *Moral Psychology and the Law: How Intuitions Drive Reasoning, Judgment, and the Search for Evidence*, in «Alabama Law Review», 64, 867 ff.
- HAIDT J., HERSH M.A. 2001. *Sexual Morality: The Cultures and Emotions of Conservatives and Liberals*, in «Journal of Applied Social Psychology», 31, 1, 191 ff.
- HALVERSON A.M., HALLAHAN M., HART A.J., ROSENTHAL R. 1997. *Reducing the Biasing Effects of Judges' Nonverbal Behavior with Simplified Jury Instruction*, in «Journal of Applied Psychology», 82, 4, 590 ff.
- HARAN U., RITOV I., MELLERS B.A. 2013. *The Role of Actively Open-Minded Thinking in Information Acquisition, Accuracy, and Calibration*, in «Judgment and Decision Making», 8, 3, 188 ff.
- HELION C., PIZARRO D.A. 2015. *Beyond Dual-Processes: The Interplay of Reason and Emotion in Moral Judgment*, in CLAUSEN J., LEVY N. (eds.), *Handbook of Neuroethics*, Springer, 109 ff.
- HELM R.K., WISTRICH A.J., RACHLINSKI J.J. 2016. *Are Arbitrators Human?*, in «Journal of Empirical Legal Studies», 13, 4, 666 ff.
- HSEE C.K. 1996. *Elastic Justification: How Unjustifiable Factors Influence Judgments*, in «Organizational Behavior and Human Decision Processes», 66, 1, 122 ff.
- INGRISSELLI E. 2014. *Mitigating Jurors' Racial Biases: The Effects of Content and Timing of Jury Instructions*, in «Yale Law Journal», 124, 1690 ff.
- KNEER M. 2022. *Reasonableness on the Clapham Omnibus: Exploring the Outcome-Sensitive Folk Concept of Reasonable*, in BYSTRANOWSKI P., JANIK B., PROCHNICKI M. (eds.), *Judicial Decision-Making: Integrating Empirical and Theoretical Perspectives*, Springer Nature, 25 ff.
- KNEER M., BOURGEOIS-GIRONDE S. 2017. *Mens Rea Ascription, Expertise and Outcome Effects: Professional Judges Surveyed*, in «Cognition», 169, 139 ff.
- KNEER M., MACHERY E. 2019. *No Luck for Moral Luck*, in «Cognition», 182, 331 ff.
- KNEER M., SKOCZEŃ I. 2023. *Outcome Effects, Moral Luck and the Hindsight Bias*, in «Cognition», 232, 105258.
- KNOBE J. 2003. *Intentional Action and Side Effects in Ordinary Language*, in «Analysis», 63, 3, 190 ff.
- KUNDA Z. 1990. *The Case for Motivated Reasoning*, in «Psychological Bulletin», 108, 3, 480 ff.
- LANDSMAN S., RAKOS R.F. 1994. *A Preliminary Inquiry into the Effect of Potentially Biasing Information on Judges and Jurors in Civil Litigation*, in «Behavioral Sciences & the Law», 12, 2, 113 ff.
- LEITER B. 2003. *American Legal Realism*, In EDMUNDSON W., GOLDING M. (eds.), *The Blackwell Guide to Philosophy of Law and Legal Theory* (2<sup>nd</sup> Ed.), Blackwell.

- LINDHOLM T., CEDERWALL J.Y. 2010. *Ethnicity and Gender Biases in the Courtroom*, in GRANHAG P.A. (ed.), *Forensic Psychology in Context: Nordic and International Approaches*, Routledge, 228 ff.
- MACPHERSON R., STANOVICH K.E. 2007. *Cognitive Ability, Thinking Dispositions, and Instructional Set as Predictors of Critical Thinking*, in «Learning and Individual Differences», 17, 2, 115 ff.
- MALLE B.F., GUGLIELMO S., MONROE A.E. 2014. *A Theory of Blame*, in «Psychological Inquiry», 25, 2, 147 ff.
- MELILLI K.J. 1998. *The Character Evidence Rule Revisited*, «Brigham Young University Law Review», 4, 1547 ff.
- MELNIKOFF D.E., STROHMINGER N. 2020. *The Automatic Influence of Advocacy on Lawyers and Novices*, in «Nature Human Behaviour», 4, 12, 1258 ff.
- MONROE A.E., MALLE B.F. 2017. *Two Paths to Blame: Intentionality Directs Moral Information Processing along Two Distinct Tracks*, in «Journal of Experimental Psychology: General», 146, 1, 123 ff.
- NADELHOFFER T. 2006. *Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Juror Impartiality*, in «Philosophical explorations», 9, 2, 203 ff.
- NADLER J. 2012. *Blaming as a Social Process: The Influence of Character and Moral Emotion on Blame*, in «Law and Contemporary Problems», 75, 2, 1 ff.
- NADLER J. 2014. *The Path of Motivated Blame and the Complexities of Intent*, in «Psychological Inquiry», 25, 2, 222 ff.
- NADLER J., MCDONNELL M.H. 2011. *Moral Character, Motive, and the Psychology of Blame*, in «Cornell Law Review», 97, 255 ff.
- PORTER T., ELNAKOURI A., MEYERS E.A., SHIBAYAMA T., JAYAWICKREME E., GROSSMANN I. 2022. *Predictors and Consequences of Intellectual Humility*, in «Nature Reviews Psychology», 1, 9, 524 ff.
- RACHLINSKI J. J., GUTHRIE C., WISTRICH A. J. 2006. *Inside the Bankruptcy Judge's Mind*, in «Boston University Law Review», 86, 1227 ff.
- RACHLINSKI J.J., GUTHRIE C., WISTRICH A. J. 2007. *Heuristics and Biases in Bankruptcy Judges*, in «Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft», 163, 167 ff.
- ROESE N.J., VOHS K.D. 2012. *Hindsight Bias*, «Perspectives on Psychological Science», 7, 5, 411 ff.
- SAUERLAND M., OTGAAR H., MAEGHERMAN E., SAGANA A. 2020. *Allegiance Bias in Statement Reliability Evaluations Is Not Eliminated by Falsification Instructions*, in «Zeitschrift für Psychologie», 228, 3, 210 ff.
- SEVIER J. 2019. *Legitimizing Character Evidence*, in «Emory Law Journal», 68, 3, 441 ff.
- SHAVER K.G. 1985. *The Attribution of Blame: Causality, Responsibility, and Blameworthiness*, Springer.
- SOOD A.M., DARLEY J.M. 2012. *The Plasticity of Harm in the Service of Criminalization Goals*, in «California Law Review», 100, 5, 1313 ff.
- SOOD A.M. 2013. *Motivated Cognition in Legal Judgments—An Analytic Review*, in «Annual Review of Law and Social Science», 9, 307 ff.
- SOOD A.M. 2015. *Cognitive Cleansing: Experimental Psychology and the Exclusionary Rule*, in «Georgia Law Journal», 103, 1543 ff.
- SRIPADA C.S. 2012. *Mental State Attributions and the Side-Effect Effect*, in «Journal of Experimental Social Psychology», 48, 1, 232 ff.

- SRIPADA C.S., KONRATH S. 2011. *Telling More Than We Can Know about Intentional Action*, in «Mind & Language», 26, 3, 353 ff.
- STANOVICH K.E., TOPLAK M.E. 2019. *The Need for Intellectual Diversity in Psychological Science: Our Own Studies of Actively Open-Minded Thinking as a Case Study*, in «Cognition», 187, 156 ff.
- STEBLAY N., HOSCH H.M., CULHANE S.E., MCWETHY A. 2006. *The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis*, in «Law and Human Behavior», 30, 4, 469 ff.
- STOEL R., BERGER C.E.H., KERKHOFF W., DROR I. 2014. *Minimizing Contextual Bias in Forensic Casework*, in HICKMAN M.J., STROM K.J. (eds.), *Forensic Science and the Administration of Justice. Critical Issues and Directions*, SAGE Publications, 67 ff.
- STROHMAIER N., PLUUT H., VAN DEN BOS K., ADRIAANSE J., VRISENDORP R. 2021. *Hindsight Bias and Outcome Bias in Judging Directors' Liability and the Role of Free Will Beliefs*, in «Journal of Applied Social Psychology», 51, 3, 141 ff.
- STROHMAIER N., ZEHNDER M.A., KNEER M. forthcoming. *Moral Character Inferences Bias Legal Judgments of Jurors and Experts Across Jurisdictions*.
- STROHMINGER N., MELNIKOFF D. manuscript. *Breaking Reality's Constraints on Motivated Cognition*. Online pre-print. Available on: <https://psyarxiv.com/qnda3/download?format=pdf>.
- UHLMANN E.L., PIZARRO D.A., DIERMEIER D. 2015. *A Person-Centered Approach to Moral Judgment*, in «Perspectives on Psychological Science», 10, 1, 72 ff.
- VAN BAVEL J.J., FELDMAN HALL O., MENDE-SIEDLECKI P. 2015. *The Neuroscience of Moral Cognition: From Dual Processes to Dynamic Systems*, in «Current Opinion in Psychology», 6, 167 ff.
- WILSON T.D., BREKKE, N. 1994. *Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations*, in «Psychological Bulletin», 116, 1, 117 ff.
- WISTRICH A.J., GUTHRIE C., RACHLINSKI J.J. 2005. *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, in «University of Pennsylvania Law Review», 153, 1251 ff.
- WYLIE J., GANTMAN A. 2023. *Doesn't Everybody Jaywalk? On Codified Rules That Are Seldom Followed and Selectively Punished*, in «Cognition», 231, 105323.
- YOURSTONE J., LINDHOLM T., GRANN M., SVENSON O. 2008. *Evidence of Gender Bias in Legal Insanity Evaluations: A Case Vignette Study of Clinicians, Judges, and Students*, in «Nordic Journal of Psychiatry», 62, 4, 273 ff.