



UvA-DARE (Digital Academic Repository)

(Dis)honesty in individual and collaborative settings

A behavioral ethics approach

Leib, M.

Publication date

2021

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Leib, M. (2021). *(Dis)honesty in individual and collaborative settings: A behavioral ethics approach*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 1

Introduction

This chapter is based on Leib, M., & Shalvi, S. (In press). Justifications as a threat for honesty: a behavioral ethics approach, in: C. B. Miller and R. West (Eds.) *Integrity, Honesty, and Truth-Seeking*. Oxford University Press.

People often need to decide between being honest or profitable but dishonest. Some decisions are rather small, for example, when a customer debates whether she should let a cashier know she accidentally received too much change or keep the extra change to herself. Other dilemmas are rather big, for instance, when engineers during the 2015 Volkswagen emissions scandal had to choose between providing regulators with the correct results of emissions test or collaborate with their colleagues and dishonestly alter the results to increase their teams' and companies' profits (Goodman, 2015). Whether small or large, when accumulated, such dishonest acts have severe consequences. A store owner may lose a lot of money if all customers keep the extra change they accidentally receive. When Volkswagen's engineers did dishonestly alter the emissions test results, the damage of the heavily polluting cars released to US roads was estimated at 59 early deaths and \$450 million (Barrett et al., 2015). With such potentially extreme consequences, studying what shapes (dis)honesty and how it can be curbed is important.

Indeed, in the last decade, researchers in economics, psychology, and management have studied what drives (dis)honest behavior. Whereas some theoretical models assume an intrinsic preference for honesty (Kartik, Tercieux, & Holden, 2014; Matsushima, 2008), and work in personality psychology views honesty as a personality dimension (honesty-humility dimension in the HEXACO model; Ashton et al., 2004), prior work in behavioral ethics research, as well as the results of this dissertation, reveal honesty is rather malleable. That is, the decision to be honest or dishonest is not fixed, but affected by many external, situational factors.

This dissertation tests several key factors and focuses on two types of dishonesty – self-serving mistakes, which is a subtle type of dishonesty, and overt lies. The four empirical chapters study (dis)honesty in two settings: an *individual* setting, where people act alone and can affect their own, or others' outcomes, and a *collaborative, interactive* setting, where people are paired with a partner and can increase mutual profits by engaging in coordinated lying. Employing diverse methodologies—eye tracking, incentivized behavioral experiments, and meta-analytical tools—the four chapters address the following questions: (1) How does framing incentives as gains versus losses affect a subtle type of (dis)honesty—self-serving mistakes—and the attentional process underlying it? (2) How do prior (un)fair experiences affect lies aimed at helping and harming others? (3) In collaborative settings, where joint dishonest acts are possible, how do people select their partners? And how does the ability (vs. inability) to select a partner affect collaborative dishonesty? (4) What are the dynamics of

collaborative (dis)honesty? Namely, do partners affect each other? How do time and financial incentives affect (dis)honesty in collaborative settings? And does the prevalence of (dis)honesty differ between individual versus collaborative settings?

Studying (dis)honesty: A behavioral ethics approach

Assessing the extent to which people are honest or not is not trivial. Because people are concerned with their reputation and want to appear as good and honest, they might be reluctant to admit their wrongdoings (Krumpal, 2013). Indeed, after violating a social norm, people often shift the blame to others or downplay the consequences of their actions (Robinson & Kraatz, 1998). Further, people tend to only partially confess their lies (Peer, Acquisti, & Shalvi, 2014). Whereas asking people directly when and why they lie provides some insights into dishonest behavior (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996), such an approach might limit our understanding of how people *actually* behave.

For a more objective assessment, the behavioral ethics approach employs lab experiments and financially incentivized tasks to capture (dis)honest behavior. These tasks instruct participants to report information that is typically private and unknown to the experimenter. Participants' reports influence their own and/or others' financial outcomes. As such, participants can exploit the information asymmetry between themselves and the experimenter and lie to boost their own and/or others' pay.

One of the most commonly used tasks is the die-rolling task, first introduced by Fischbacher and Föllmi-Heusi, (2013). Participants were asked to roll a die in private and report the outcome, and were paid based on their reports. Reporting 1 yielded a payoff of 1 CHF (Swiss Franc), 2 = 2 CFH, 3 = 3 CFH, 4 = 4 CFH, 5 = 5 CFH, and 6 = 0 CFH. Because the experimenters did not know the actual die-roll outcome, participants could lie to inflate their pay. And they did. The distribution of reported die-roll outcomes was different from the uniform distribution expected if participants were honest. Interestingly, participants did not lie to the full extent. That is, not all participants reported the most profitable die-roll outcome, 5. Instead, participants reported all die-roll outcomes, but were more likely to report the most profitable outcomes—4 and 5. Different experimental manipulations (e.g., financial incentives, whether participants engaged in the task once vs. twice) affected reports. Nevertheless, none of the manipulations led participants to lie to the full extent.

The clean nature of lab experiments provides big advantages in terms of control. But does dishonesty in experimental, lab tasks generalize to everyday dishonest acts? Some recent

research suggests it does. For instance, dishonest behavior in experimental tasks is associated with riding in public transportation without paying (Dai, Galeotti, & Villeval, 2018), not returning an undeserved payment (Potters & Stoop, 2016), and milkmen in Indian dairy markets diluting of milk with water (Kröll & Rustagi, 2016). Similarly, misreporting in experimental tasks also correlates with other socially deviant behaviors, such as misbehaving in school (Cohn & Maréchal, 2017) and in a juvenile detention center (Cingl, & Korb, 2020), and skipping work without a reason (Hanna & Wang, 2017). Thus, findings obtained in the lab can, to an extent, be generalized to behaviors outside of it.

The role of justifications in shaping (dis)honesty

Considering what affects (dis)honest behavior, classical economic theory suggests people make a simple cost-benefit analysis (Becker 1968). On one side of the scale, they weigh the profits they will obtain from lying. On the other side, they weigh the probability of getting caught lying and the punishment that will follow if they are caught. If the gains outweigh the costs, people will lie; if not, they will be honest. The last decade of research, however, suggests people take into account many other considerations.

One main consideration is moral self-image. People want to see themselves as honest and moral, and are reluctant to take actions that threaten this view (Mazar, Amir, & Ariely 2008). As such, people lie, but only as long as doing so does not affect their moral self-image. As a result, people lie just a little. Indeed, in a recent meta-analysis, Abeler, Raymond, and Nosenzo (2019) summarized 90 experiments, revealing people increase their profits by lying by only 23.4% of what they could (see similar findings by Gerlach, Teodorescu, & Hertwig, 2019).

To balance between the desire to profit from dishonesty and the motivation to maintain a moral self-image, people rely on justifications to lie (Mulder, Jordan, & Rink 2015; Schweitzer & Hsee, 2002; Shalvi, Gino, Barkan, & Ayal, 2015). Put differently, people need convincing rationalizations that their acts do not reflect poorly on them in order to free themselves to lie. These justifications are internal, because they help individuals explain their behavior to themselves and not necessarily to others.

Justifications fall into two main categories—self-serving justifications and socially motivated justifications. Self-serving justifications allow individuals to process and use information in ways that benefit them and in turn free them to lie. Socially motivated justifications apply when dishonest acts benefit others as well as oneself. The prosocial

consequence of dishonesty makes people evaluate their behavior as rather ethical, thus reducing the threat to their moral self-image.

Self-serving justifications: Using tempting yet irrelevant information

People often use tempting yet irrelevant information to justify their lies. For instance, seeing die-roll outcomes that are more profitable but irrelevant for pay increases dishonesty in die rolling tasks. Specifically, instead of following the instructions and reporting the relevant-for-pay outcome, participants report the highest outcome they see (Shalvi, Dana, Handgraaf, & De Dreu, 2011; Bassarak et al., 2017; Gächter & Schulz, 2016). Additionally, information about peers' behavior affects lies. Learning that others lied pushes people to lie as well (e.g., Gino, Ayal, & Ariely, 2009; O'Fallon & Butterfield, 2012; Leib & Schweitzer, 2020).

Pittarello, Leib, Gordon-Hecker, and Shalvi (2015) set out to test how people use irrelevant yet tempting information in ambiguous situations. Ambiguity allows individuals to interpret information in ways that align with their motivation (Balcetis & Dunning, 2006; 2007; 2009). In ambiguous situations, telling yourself, as well as others if suspicion arises, that you merely misunderstood or made a mistake and that you did not intend to do something wrong is easy. Devising the ambiguous-die paradigm, across multiple trials, Pittarello and colleagues (2015) had participants observe an "X" on a computer screen for one second, followed by an image of six die-roll outcomes for two seconds. Then, participants had to report the die-roll outcome that appeared closest to the "X" (the "target die"). The higher the outcome participants reported, the more they were paid; see Figure 1.1. In half of the trials, the die appearing second closest to the "X" (the "die next to the target") was higher than the target die, and thus tempting to report. In the other half of the trials, the die next to the target was lower than the target, and thus not tempting to report.

Between trials, the authors further altered the ambiguity of the trial to be either low, medium, or high by altering the location of the "X." The higher the ambiguity, the further the "X" was from the target die, making it harder to identify. Crucially, the "X" always appeared closer to the target die than to the die next to the target. Will participants use tempting yet irrelevant information to boost their profits? If so, how? Will they report the most profitable outcome, regardless of its location? Alternatively, will participants report tempting yet irrelevant outcomes when doing so is justifiable? Finally, how will ambiguity shape reports?

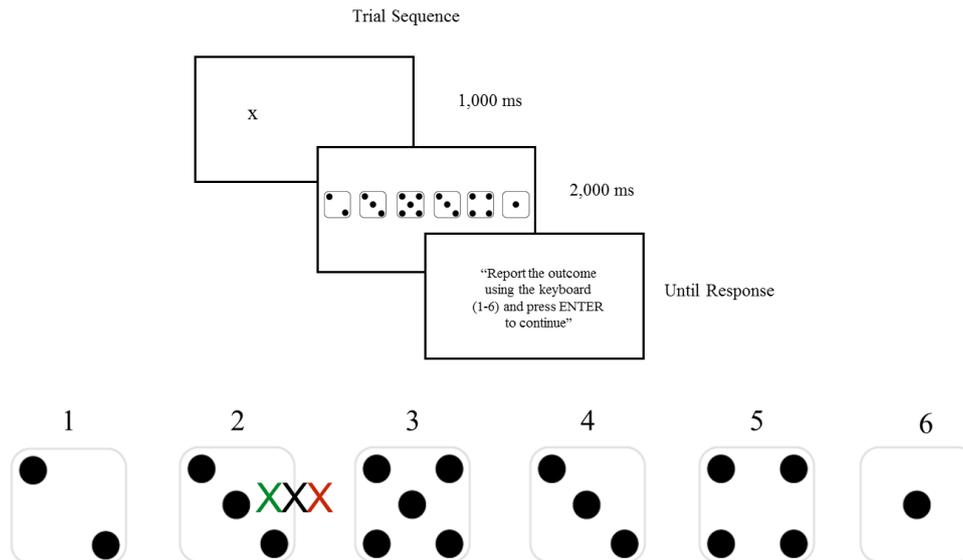


Figure 1.1. The procedure and an example trial of the ambiguous die paradigm. In the example, the target die is 3; the die next to the target is 5. In each trial, ambiguity was low (the leftmost “X” in green), medium (the center “X” in black), or high (the rightmost “X”, in red).

Results revealed participants used irrelevant yet tempting information when doing so was justifiable. Specifically, participants did not report the most profitable outcome on the screen. Instead, they either reported the target die or made a “mistake” and reported the die next to the target. A closer look at the pattern of such “mistakes” revealed they were rather self-serving. Participants were more likely to report the die next to the target when it was higher than the target (and thus tempting and more profitable) than when it was lower than the target (and thus non-tempting and less profitable). Further, the higher the ambiguity, the more self-serving “mistakes” participants made. In a control condition, where participants were paid only if they correctly identified the target die, fewer mistakes were made overall. Further, mistakes were rather equally distributed between trials in which the die next to the target was higher and lower than the target.

Whereas people clearly use tempting yet irrelevant information to justify their lies, the extent to which they do so under different motivations remains an open question. Does the motivation to maximize gains versus minimize losses affect self-serving mistakes? Do these motivations affect the attentional process underlying self-serving mistakes? Chapter 2 addresses these questions.

Socially motivated justifications: Lying for the greater good

Socially motivated justifications apply to situations where one's lies affect others' outcomes, mostly in positive ways. The prosocial consequences of one's dishonesty allow people to do wrong but feel moral. After all, when lies serve a greater good, they are easily justifiable. Indeed, people feel less guilty and evaluate themselves as better people when they lie to benefit themselves and others than when they lie to benefit only themselves (Gino, Ayal, & Ariely, 2013). Surprisingly, third parties even evaluate people who dishonestly benefit others as more moral, benevolent, and trustworthy than those who are honest but selfish (Levine & Schweitzer, 2014; 2015).

In line with such positive evaluation, past work found people forgo their honesty for a variety of prosocial reasons. People lie to benefit the less wealthy (Gino & Pierce, 2010b; Okeke & Godlonton, 2014), to restore equality (Atanasov & Dana, 2011; Gino & Pierce, 2009, 2010a), and to secure high pay for charity (Ayal, 2015; Lewis et al., 2012). Further, people lie more when doing so benefits both themselves and others, as compared to only themselves (Gino, Ayal, & Ariely, 2013; Wiltermuth, 2011).

Whereas people lie to benefits themselves and others, will they lie to *only* benefit others? Will they lie to only harm others? And does being treated fairly or unfairly affect subsequent other-helping and harming lies? Chapter 3 provides answers to these questions.

(Dis)honesty in collaborative settings

Until recently, the vast majority of the literature studying (dis)honesty focused on individual decision-making, where people engage in a task alone. A growing body of work assesses (dis)honesty in collaborative settings where people can interact and/or coordinate with others on joint dishonest acts to boost mutual pay (Cohen, Gunia, Kim-Jun, & Murnighan, 2009; Conrads, Irlenbusch, Rilke, & Walkowitz, 2013; Conrads et al., 2017; Kocher, Schudy, & Spantig, 2017; Sutter, 2009; Weisel & Shalvi, 2015). Moving toward studying (dis)honesty in collaborative settings is not only theoretically interesting but also practically important. People work in groups and collaborate with one another on a regular basis. Further, working together allows people to complete tasks and achieve complex goals they could not if they worked alone. As such, many, if not most, important decisions are made by groups rather than by individuals. Therefore, understanding how collaborative settings shape honesty and ethical decision-making is especially crucial.

Collaborative settings might pose a unique tension between two moral obligations—to be honest and collaborative. Although in many situations, people can be both honest and collaborative, in some situations, the two are in conflict. Consider the Volkswagen case that I began the introduction with. To collaborate with their colleagues and increase the profits for their team and entire company, engineers had to work together to dishonestly manipulate the emissions test results (Goodman, 2015). Situations in which such tension exists are the main focus of research on collaborative dishonesty and of chapters 4 and 5 in this dissertation.

Assessing how people trade off honesty and collaboration, Weisel and Shalvi (2015) developed the dyadic die-rolling task. In the task, participants are paired in dyads and assigned the role of a first mover or a second mover. Across multiple rounds, the first mover rolls a die in private and reports the outcome. Then, the second mover learns about the first mover's report, rolls an independent private die, and reports the outcome as well. Then, the first and second mover learn about each other's reports and payoff, and move to the next round. If the first and second mover report the same outcome (a double), they both get paid according to the double's worth, with higher doubles corresponding to higher pay. If the first and second mover report different outcomes (a non-double), they do not get paid; see Figure 1.2. For instance, if the first and second mover report a 1, they earn €1. If the first and second mover report a 6, they earn €6. If the first mover reports a 1 and the second mover reports a 6, they earn €0.

Weisel and Shalvi, (2015) found that participants reported doubles in 81.5% of the cases. This proportion was much higher than the expected proportion of doubles if participants were honest (16.66%). This proportion was also higher than the proportion of doubles in an individual condition, in which the same participant took the role of both the first and second mover (54.8%). Clearly, when collaboration is pitted against honesty, people choose to collaborate over being honest. But how do people select their partners in such settings? Will dishonest people prefer to interact with dishonest partners? With whom do honest people prefer to interact with—honest or dishonest partners? And how does the ability (vs. inability) to choose a partner affect the spread and efficiency of dishonesty? Chapter 4 addresses these questions. Lastly, chapter 5 employs meta-analytical tools and explores dishonesty in collaborative settings further. By analyzing 51,640 decisions, made by 3,264 participants across 43 conditions employing the dyadic die-rolling task, the chapter tests the following questions: Is partners' behavior correlated? How does (dis)honesty in collaborative settings develop over time? And is it affected by financial incentives? Further, chapter 5

makes initial steps to meta-analytically compare the prevalence of (dis)honesty in individual and collaborative settings.

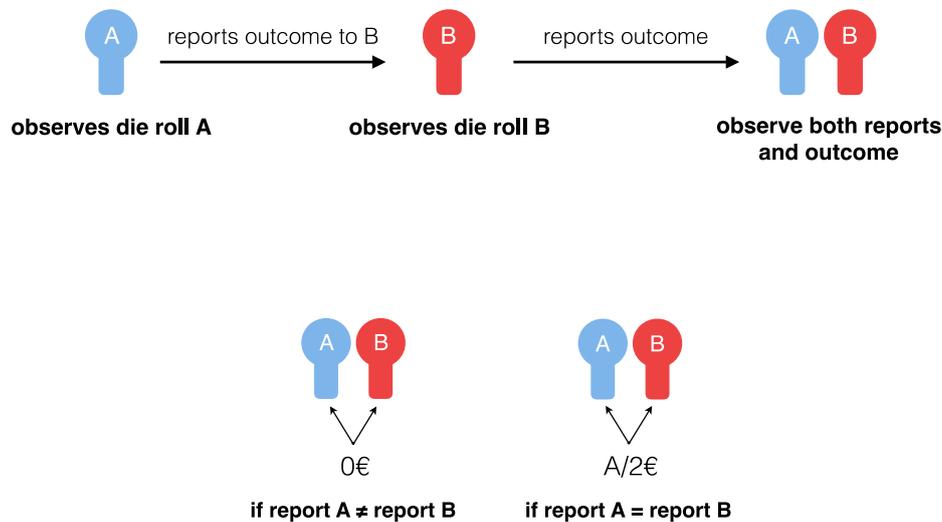


Figure 1.2. The procedure of the dyadic die-rolling task. The first mover (in Blue, A) observes a die-roll outcome and reports an outcome to the second mover (in Red, B). The second mover (B) then observes an independent die-roll outcome and reports an outcome. Then, both the first and second movers learn about both reports and the corresponding payoff. If the number the second mover (B) reports matches the number the first mover (A) reports (= a double), both participants earn money (higher doubles = higher pay). Otherwise, both participants earn no extra pay.

Overview of empirical chapters

The four empirical chapters in this dissertation test how gain versus loss framing (chapter 2) and prior (un)fair experience (chapter 3) affect (dis)honesty in individual settings, where people act alone. Then, the dissertation moves toward collaborative settings and assesses behavior in situations where collaboration and honesty clash. Chapter 4 examines how people choose their partners, and the effect of the ability to choose a partner on collaborative dishonesty. Chapter 5 further explores the dynamics emerging in collaborative dishonesty employing meta-analytical tools.

Results reveal honesty is rather malleable. The motivation to minimize losses, versus maximize gains, drives people to make more self-serving mistakes (chapter 2). Further, after experiencing fair, unfair, and no prior treatment, people are likely to lie to help others. Only after experiencing an unfair treatment, a rather small fraction of people lie to harm others (chapter 3). Testing how people select their partners in collaborative settings reveals dishonest partners are quite popular, because both dishonest and honest people prefer to interact with

dishonest partners. Dishonest people seek and find a “partner in crime.” Honest people engage in “ethical free riding”—they intentionally benefit from their partner’s rule violation without violating the rules themselves (chapter 4). Lastly, conducting the first meta-study on collaborative dishonesty reveals partners’ behavior is correlated—if one person lies, the partner is more likely to lie as well. Further, lies increase as the task progresses and as the financial incentives to lie get bigger. Initial comparisons between individual and collaborative settings further reveal lies are more prevalent in collaborative settings (chapter 5).

Chapter 2: Loss framing increases self-serving mistakes (but does not alter attention)

Chapter 2 employs the ambiguous-die paradigm and focuses on a subtle type of dishonesty, namely, self-serving mistakes (see Figure 1.1). The chapter consist of two financially incentivized experiments in which participant can make self-serving mistakes to either increase their gains or decrease their losses. Further, participants’ eye movements are tracked to explore the attentional mechanism underlying self-serving mistakes. Results reveal participants are twice as likely to make self-serving mistakes to decrease their losses than to increase their gains. Eye-tracking data show that tempting, self-serving information does not attract more attention when facing a loss, compared to an equal-sized gain. Rather, in loss framing, people are more likely to use and report the tempting, self-serving information they attend to.

Testing how the motivation to avoid losses versus secure gains shapes self-serving mistakes is important on several accounts and can be especially useful for organizations. First, self-serving mistakes are a subtle type of dishonesty and might be more common than overt, blunt lies (Gino & Bazerman, 2009). Thus, this understudied type of dishonesty deserves more attention in the literature as a whole (for an exception, see Exley & Kessler, 2018). Second, organizations vary in the way they set and communicate goals. Some organizations might highlight the potential bonus employees can earn by reaching a certain goal, whereas other organizations might highlight the possibility of losing a potential bonus to a competitor if a goal is not reached. Findings obtained in chapter 2 may assist organizations to set goals and incentive schemes in ways that will curb self-serving “mistakes” and foster honesty and accuracy.

Chapter 3: Dishonest helping and harming after (un)fair treatment

Chapter 3 assesses the extent to which individuals engage in dishonest behavior aimed at helping or harming others after they experience (un)fair treatment. Across three financially

incentivized experiments, recipients in a dictator game received a fair or unfair amount and then could dishonestly inflate or deflate their counterparts' pay. Results show dishonest helping is a common and robust behavior. Individuals lie to help others after fair, unfair, and no prior treatment. Dishonest harming, however, is less prevalent. Only after unfair treatment, some, but not all, individuals engage in dishonest harming. Assessing self-reported emotions reveals that dishonest harming is associated with high levels of anger and disappointment, and low levels of gratitude. Interestingly, the source of (un)fairness, whether it is intentional or not, does not attenuate behavior, suggesting dishonest reactions to (un)fairness are driven by the mere (un)fair treatment, and not by a motivation to reciprocate an (un)fair counterpart. As people regularly interact with one another and treat each other in various (also fair and unfair) ways, understanding how such social interactions shape our (dis)honesty and other-oriented behavior is important.

Chapter 4: Ethical free-riding: When honest people find dishonest partners

Whereas chapters 2 and 3 focus on individual settings, where people act alone and affect their own (in chapter 2) or others' (chapter 3) pay, chapters 4 and 5 focus on collaborative settings, where people can lie together to secure mutual pay. Chapter 4 tests how people choose their partners in such collaborative settings. Participants engaged in the dyadic die-rolling task, in which they were assigned a partner and could increase their mutual profits by coordinated lying. After several interactions, participants were either (i) free to choose whether to stay with or switch their partner, or (ii) forced to stay with or (iii) forced to switch their partner. Results reveal both dishonest and honest people exploit the freedom to choose a partner. Dishonest people seek a dishonest partner—a “partner in crime.” Honest people, by contrast, engage in “ethical free riding”: they refrain from lying but also from leaving dishonest partners, taking advantage of their partner's lies. Further results reveal that allowing individuals to select their partners (vs. forcing them to switch them) increases the efficiency and contagiousness of lies. Our findings thus suggest that to curb collaborative dishonesty, relying on people's honesty is insufficient. Encouraging honest individuals not to engage in ethical free riding is essential.

The findings obtained in chapter 4 may have implications for organizations. In many organizations, employees collaborate with one another. Some organizations allow employees to choose their collaborators, for instance, by implementing voluntary job-rotation policies (implemented in, e.g., the European Commission; Fontaine & Tang, 2006). Other organizations force employees to change collaborators once in a while, for instance, by

implementing mandatory job-rotation policies (implemented in, e.g., the United Nations; Fontaine & Tang, 2006). Because different organizations implement different policies, understanding how people choose their partners when they can, and how the ability to choose partners affects (dis)honesty in collaborative settings, is important. Such understanding is especially important for organizations with a high risk of collusion and corruption.

Chapter 5: (Dis)honesty in collaborative settings: A meta-study

Lastly, chapter 5 reports the first meta-study on (dis)honesty in a collaborative setting. With recent meta-analyses' summarized findings on honesty in individual settings (Gerlach et al., 2019; Abeler et al., 2019; Köbis, Verschuere, Bereby-Meyer, Rand & Shalvi, 2019), and with the literature focusing more and more on (dis)honesty in collaborative and group settings, summarizing the knowledge and data collected so far on collaborative dishonesty is therefore useful. Analyzing 51,640 decisions, made by 3,264 participants, across 43 conditions in the dyadic die-rolling task reveals that in collaborative settings, individuals' behavior is affected by their partners: if one person lies, the partner is more likely to lie as well. Further, participants lie more as the task progresses and as the financial incentive to lie increases. Initial comparisons between individual and collaborative settings further reveal dishonesty is more prevalent in collaborative settings. Beyond providing aggregated results on collaborative dishonesty, chapter 5 further identifies new theoretical questions and sets up an agenda for future research to better understand (dis)honesty in collaborative settings.

Note that all chapters are based on separate scientific papers that have been published or are in preparation for submission. Thus, some theoretical overlap exists between the chapters, and each chapter may be read independently.