(Dis)honesty in individual and collaborative settings

A behavioral ethics approach

Leib, M.

Publication date
2021

Document Version
Other version

License
Other

Chapter 5

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

This chapter is based on Leib, M., Köbis, N.C., Soraperra, I., Weisel, O., & Shalvi, S. (Dis)honesty in collaborative settings: A meta-study. Working paper.

In many situations, people may be tempted to forgo their honesty and violate ethical rules for profit. At times, such ethical rule violations may have severe financial and societal consequences. For instance, when Volkswagen's engineers illegally manipulated emissions test results to meet US standards, they allowed for the release of highly polluting vehicles to roads around the world. In the US alone, the damage from those vehicles was estimated at 59 early deaths and $450 million (Barrett et al., 2015). Due to the potential severe consequences ethical rule violations may entail, studying the extent to which ethical rule violations occur, and the process underlying them, is important.

Indeed, in the last decade, researchers across many disciplines have been studying what shapes ethical rule violation, and in particular, dishonesty. Recently, this work has been summarized in several literature reviews (Gino & Ariely, 2016; Jacobsen, Fosgaard, & Pascual-Ezama, 2018; Shalvi, Gino, Barkan, & Ayal, 2015), and meta-analyses (Gerlach, Teodorescu, & Hertwig, 2019; Abeler, Nosenzo, & Raymond, 2019; Köbis, Verschuere, Bereby-Meyer, Rand, & Shalvi, 2019). Most prior work assessed (dis)honesty in *individual* settings, where people affect their own profits and engage in experimental tasks alone. However, in life, people frequently *interact* and *collaborate* with others. For instance, engineers collaborate with colleagues to find solutions for complex problems, and academics work on research papers together. Often, individuals can collaborate and be honest at the same time, yet sometimes the two moral obligations—be honest and collaborative—are in conflict. In such situations, individuals have to give up one moral obligation to pursue the other. Consider the Volkswagen case we started with. To collaborate with their colleagues and obtain team bonuses, engineers had to dishonestly help their colleagues manipulate the emissions test results (Goodman, 2015).

Here, we add to prior work that aggregates data on (dis)honesty in *individual* settings, and conduct the first meta-study on (dis)honesty in an *interactive*, *collaborative* setting where honesty and collaboration clash. Specifically, we gather and analyze 51,640 decisions, made by 3,264 participants across 43 conditions in the dyadic die-rolling task (Weisel & Shalvi, 2015). In this sequential task, over multiple rounds, participants can increase their financial profits by dishonestly coordinating with one another. Focusing on the dyadic die-rolling task, we tackle new theoretical questions as well as set up a research agenda for future work based on our findings.

(Dis)honesty in collaborative, interactive settings

Two recent meta-analyses statistically summarized the large and ever-growing literature on (dis)honesty (Gerlach et al., 2019; Abeler et al., 2019). In line with the idea that people want to think of themselves as honest and good (Mazar, Amir, & Ariely, 2008), and thus lie only to the extent to which they can justify their lies (Shalvi et al., 2015), these meta-analyses revealed that people lie, but just a little. Indeed, Abeler and colleagues (2019) found that participants lied to increase their profits only by 23.4% of what they could. Gerlach and colleagues (2019) further revealed that various situational factors (e.g., lab vs. field, the existence of normative cues, experimental deception) as well as personal factors (e.g., age, gender) affect lies. These meta-analyses focused on individual settings, where participants engage in various experimental tasks alone, and for the most part, affect only their own financial outcomes.

Whereas focusing on individual settings is valuable, assessing (dis)honesty in collaborative settings, where people interact with others, is important as well. Collaboration helps individuals complete tasks and achieve complex goals they cannot achieve alone, and many decisions, also in the ethical domain, are made in collaboration with others. To shed light on this rather neglected setting, recent work has started studying (dis)honesty in collaborative settings (Cohen, Gunia, Kim-Jun, & Murnighan, 2009; Conrads, Irlenbusch, Rilke, & Walkowitz, 2013; Conrads et al., 2017; Gross, Leib, Offerman, & Shalvi, 2018; Muehlheusser, Roider, & Wallmeier, 2015; Kocher, Schudy, & Spantig, 2017; Pulfrey, Durussel, & Butera, 2018; Sutter, 2009; Weisel & Shalvi, 2015).

One task that is commonly used to assess (dis)honesty in collaborative settings is the dyadic die-rolling task, which is the focus of our analysis. First introduced by Weisel and Shalvi (2015), in the task, participants are paired in dyads. In each dyad, one participant is randomly assigned to the role of the first mover, and the other is assigned to the role of the second mover. Across multiple rounds, the first mover rolls a die in private and reports the die-roll outcome. The second mover learns about the first mover's report, and then rolls a die in private and reports the outcome as well. At the end of each round, both dyad members learn about each other's reports and payoff and move to the next round of the task. The payoff scheme is such that if both dyad members report the same outcome (a double), they get paid according to the double's worth. If they report different outcomes (a non-double), they do not get paid. For instance, if the first and the second mover both report 1, each earns €1. If the first and second mover both report 6, each earns €6. If the first mover reports 1 and the second

mover reports 6, they both earn €0. Because the die-roll outcomes are private, participants can either be honest and report the die-roll outcomes they actually observed or lie to secure higher profit. Given that both movers' payoff is determined by both reports, in the dyadic die-rolling task, participants can secure high mutual gains by engaging in collaborative dishonesty.

Implementing the task, Weisel and Shalvi (2015) found that dyads lie quite a lot. Specifically, dyads reported doubles in 81.5% of the cases, much higher than the 16.66% anticipated if participants were honest. Further, the proportion of doubles in the dyadic task was much higher than in an individual version of the task, in which the same participant took the role of both the first and second mover and engaged in the task alone (54.8% doubles).

Whereas several other tasks assess (dis)honesty in collaborative settings, analyzing participants' behavior *only* in the dyadic die-rolling task has several advantages. First, the dyadic die-rolling task is commonly employed, providing a high number of observations on collaborative dishonesty. Since first used, follow-up work has explored how social norms (Wouda, Bijlstra, Frankenhuis, & Wigboldus, 2017; Soraperra et al., 2017), punishment and framing (Della Vella, Gino, & Piovesan, working paper), the ability to choose one's partner (Gross et al., 2018), and partners' characteristics (Ścigała et al., 2019; Gross & De Dreu, working paper; Verwijmeren, van Lent, & Bijlstra, working paper; Burghoorn, De Irueta Florentina, Bijlstra, & Verwijmeren, working paper) affect behavior in the dyadic die-rolling task.

Second, the dyadic die-rolling task is sequential and repeated. As such, it advances prior valuable work that captured collaborative dishonesty employing one-shot and/or simultaneous decisions (Barr & Michailidou, 2017; Cohen et al., 2009; Conrads et al., 2013; 2017; Kocher et al., 2017; Muehlheusser et al., 2015; Pulfrey et al., 2018; Sutter, 2009). The sequential and repeated nature of the dyadic die-rolling task allows us to tap into the interaction between the two partners and capture the dynamics of behavior over time. Thus, analyzing behavior in the dyadic die-rolling task captures not only the prevalence of (dis)honesty in a collaborative, interactive setting, but also the dynamics that develop between the first and second mover over time.

Third, focusing on a single task provides the advantage of exploring one dependent variable, instead of several different dependent variables. Focusing on one variable allows clearer conclusions to be drawn about (dis)honesty in the collaborative setting (see a similar approach in meta-analyses focusing on one experimental task; Capraro, 2018; Engel, 2011; Oosterbeek, Sloof, & Van De Kuilen, 2004; Johnson & Mislin, 2011). Doing so allows us to

avoid collapsing different dependent variables across several experimental tasks, which at times can lead to less accurate conclusions regarding the behavior of interest.

Here, we gather and analyze all decisions made by participants who engaged in the dyadic die-rolling task. We do so to pursue two main goals. The first goal is to address new and interesting theoretical questions regarding collaborative dishonesty. The second goal is to identify new theoretical questions and set up an agenda for future research to better understand (dis)honesty in collaborative settings.

The unique nature of the dyadic die-rolling task, together with our access to the raw data of all studies, allowed us to tackle the following questions: (i) Is (dis)honesty stable (as suggested by, e.g., Kartik, Tercieux, & Holden, 2014; Matsushima, 2008; Ashton et al., 2004), or alternatively affected by the (dis)honesty of one's interaction partner? (ii) How does participants' (dis)honesty develops over time? (iii) How do financial incentives affect (dis)honesty in the collaborative settings? Further, we make the first steps into comparing individual and collaborative settings meta-analytically by retrieving the prevalence of (dis)honesty in individual settings as captured by Abeler and colleagues' (2019) meta-analysis and assessing (iv) whether the prevalence of (dis)honesty differs between individual and collaborative settings. Lastly, we test (v) how situational factors (lab vs. field study, fixed roles in the task, fixed partners throughout the task), as well as personal factors (gender and age) affect dishonesty in the collaborative setting.

The current meta-analysis

Methods

Literature search. In September 2018, we conducted an online database search to identify eligible studies. We searched in Web of Science, PsycINFO, and Econlit for articles that include the following combination of search terms and Boolean operators in the text of the article: cooperat* OR coollaborat* AND dishonesty OR cheating OR lying OR deception OR unethical behavior OR corruption OR honesty OR ethical behavior. We further searched for articles that contain the terms "corrupt collaboration," "joint unethical acts," "joint dishonest acts," "dyadic die-rolling task," and "dyadic die-rolling game." Because the task was introduced in 2015, we restricted our search to articles that were published in 2015 onwards. Additionally, we screened all the articles that cite Weisel and Shalvi (2015).

Simultaneously, we sent calls inviting colleagues to send us published or unpublished work that employed the dyadic die-rolling task. Namely, we posted calls on the mailing lists

of the *Economic Science Association, Society of Judgement and Decision-Making, European Association for Decision Making, Academy of Management* (Organizational Behavior), *Society for Personality and Social Psychology, European Association for Social Psychology*, and the *International Association for Research in Economic Psychology*, as well as personally emailed colleagues we knew may have used the task.

In November 2018, we conducted an updated online search, in which we added the word "coordination" as a synonym for "cooperation" to the list of search terms. We further searched the outlined keyword combinations on Google Scholar. The search covered journal articles, book chapters, PhD dissertations, master's theses, and working papers. Our main pre-registered[13] inclusion criteria were that (i) participants' behavior was financially incentivized, (ii) each participant was part of a group of at least two individuals, (iii) the task used a random device to affect participants' actual outcome, (iv) more than one group member had the opportunity to misreport the actual outcome in the task, (v) at least one member of the group received a higher payoff if two or more members of the group coordinate on a specific rule, and (vi) in case participants' actual die-roll outcome was known to the experimenter, no sanctioning mechanism for being suspected or detected as a liar was present.

Following the inclusion criteria, we excluded from the analyses one article in which participants interacted with a simulated, pre-programmed "partner" instead of another actual participant (Ścigała et al., 2019). Further, while conducting the literature search, we found that all articles, except for one, implemented the dyadic die-rolling task in groups of two individuals. One article (Rilke, Danilov, Irlenbusch, Weisel, & Shalvi, working paper) employed a version of the task in which three participants interacted with one another and were paid when they reported triples instead of doubles. Changing the group size from two to three may have affected participants' efficiency concern and level of prosocial motivation. Thus, because only one article employed the task with a group size of three, and to meaningfully interpret our results, we included only the dyadic version of the task, as originally introduced by Weisel and Shalvi (2015). Figure 5.1 provides a PRISMA chart (Moher, Liberati, Tetzlaff, & Altman, 2009) that gives an overview of the identification, screening, and selection process.

After identifying the relevant articles, we contacted the authors and received the raw data from all of them. Overall, our analyses draw on the raw data of 43 conditions across nine

[13] https://osf.io/89dte/?view_only=dddc7e01b48b431cb3a5aaa1b8ca3437

articles (four published) employing the dyadic die-rolling task. Conditions vary on various factors such as the incentive structure, the composition of the dyad, the dyad members' ability (vs. inability) to select their interaction partners and more. Overall, our data set included 51,640 reports, made by 3,264 individuals (1,632 first movers and 1,632 second movers).

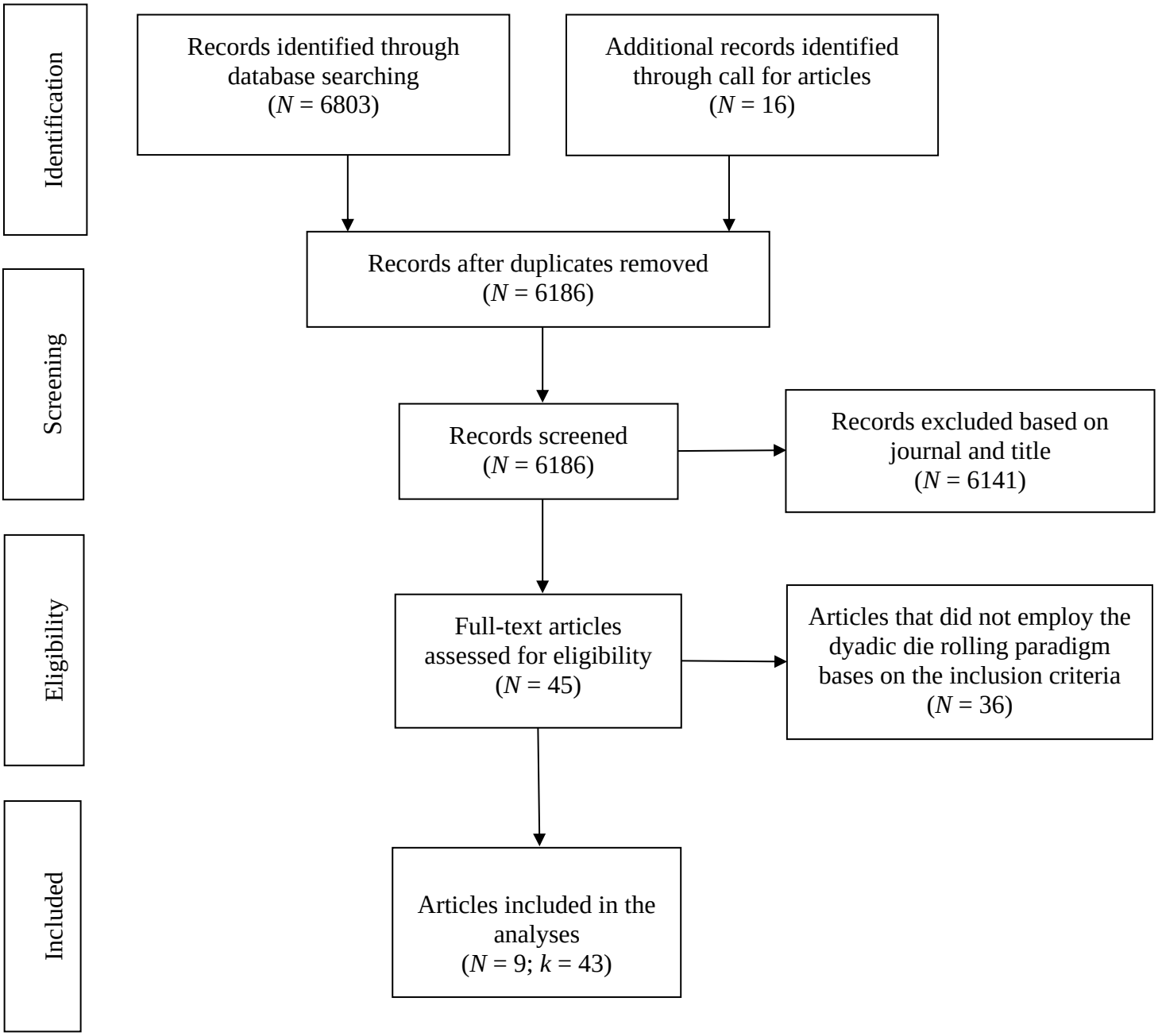

Figure 5.1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) chart describing the literature searching and screening process. *N* indicates number of papers; *k* indicates number of conditions.

Analytical approach. In all 43 conditions, unique dyads engaged in the task for multiple rounds (between 3 and 30 rounds, $M = 18.74$, $SD = 6.59$). In the majority of the conditions ($k = 29$), dyads and roles were fixed. Each participant interacted with only one partner and in one role, first or second mover, throughout the entire task. In these conditions, each dyad composed of a unique first and second mover was given a unique dyad number. The dyad number was the independent unit of observation ($n = 1048$ dyads and units of observations).

In four conditions, partners were fixed but roles were not. That is, each participant interacted with the same partner throughout the task, in the role of the first mover in some rounds and in the role of the second mover in other rounds. In those conditions, the independent unit of observation was all the rounds in which the same two participants interacted, regardless of their roles ($n = 198$ dyads, 99 units of observations).

In six conditions, roles were fixed but partners were not. That is, each participant was either the first or the second mover but interacted with different partners throughout the task. In four out of the six conditions, participants were assigned to groups of four: two first movers and two second movers. Throughout the task, each first mover interacted with both second movers, and each second mover interacted with both first movers. In those conditions, the independent unit of observation was all the rounds in which all four participants interacted with one another. In the remaining two conditions (in Gross et al., 2018), participants either interacted with or could potentially interact with all the other participants in the opposite role in a session of 20 participants. Following Gross and colleagues' (2018) procedure, in these conditions, the independent unit of observation was the session. Overall, across the six conditions, we had 1,207 dyads and 45 units of observations.

Lastly, in the four remaining conditions, neither roles nor partners were fixed. Thus, each participant interacted with several partners—in some rounds in the role of the first mover, and in other rounds in the role of the second mover. In those conditions, the independent unit of observation was the session containing all the participants who interacted with one another, regardless of their role ($n = 504$ dyads, 42 units of observations). For brevity, we refer to the independent unit of observation as "group" throughout the results section.

To analyze the data, we fitted mixed-effects linear regression models (when the dependent variable was continuous) and mixed-effects logistic regression models (when the dependent variable was binary). To account for the nested structure of the data, we specified

random intercepts for each group (independent unit of observation) and condition, when we assessed behavior across multiple rounds of the dyadic task. When we assessed behavior in one round only (e.g., first round of the task), we specified random intercepts for each condition.

Dependent variables. In the dyadic die-rolling task, first movers roll a die and report the outcome. Then, second movers roll a die and report the outcome as well. The majority of the conditions in our data set (k = 30) employed an "aligned outcomes" incentive structure. That is, if the dyad reported a double, both the first and second mover got paid. If the dyad did not report a double, both the first and second mover did not get paid. Further, the higher the double's worth, the higher the amount the dyad members earned. As such, first and second movers have different roles in the task. The first movers "set the stage" by determining the *potential* amount the dyad can earn. The second movers "get the job done" (or not) by determining whether the dyad *will* earn that amount (if they match the first mover's reports) or not (if they do not match the first mover's report). Given the different roles that first and second movers have in the task, we focus on different dependent variables for each role.

First movers. Assessing the behavior of the first movers, we focus on the die-roll outcome they report, between 1 and 6. A condition in which the mean die-roll outcome is 6 is a condition in which all first movers lied to the full extent in all rounds to maximize their (potential) pay. A condition in which the mean die-roll outcome is 3.5 is a condition in which, on average, first movers reported honestly across all rounds. A condition in which the mean die-roll outcome is 1 is a condition in which all first movers lied to the full extent to minimize their (potential) pay. Whereas assessing whether a first mover lied or reported honestly in a given round is not possible, on average, the die-roll outcome of honest first movers should be 3.5. We thus compare the die-roll outcomes to 3.5 as a proxy for (dis)honesty. We further interpret higher die-roll outcomes as indicating higher levels of dishonesty.

Second movers. Assessing the behavior of the second movers, we focus on the binary outcome of whether second movers reported a double. Although assessing whether a second mover lied or reported honestly in a given round is not possible, honest second movers should match first movers' reports in 16.66% of the cases. As a proxy for (dis)honesty, we thus compared the proportion of doubles that second movers reported to the expected proportion if second movers were honest, 16.66%. Further, we interpret a higher chance of reporting doubles as a higher likelihood of dishonesty.

Beyond the "aligned outcomes" incentive structure, in the remaining conditions (k = 13) other incentive structures were implemented. For instance, in some conditions, first movers were paid if the dyad reported a double (with higher doubles = higher pay), but second movers received a fixed payment regardless of what the dyad reported. In other conditions, first movers received a fixed payment, and second movers were paid if the dyad reported a double (with higher doubles = higher pay). Thus, in the remaining conditions, the dyad members' incentives were misaligned. Whereas the incentive structures in those conditions vary, in all conditions, at least one dyad member received higher pay when the dyad reported a double versus a non-double. We thus included all conditions in our analyses and specified random intercepts for each condition to control for the different incentive schemes.

Results

Do participants lie in the dyadic die-rolling task?

First movers. We calculated an average die-roll outcome for each first mover. The average die-roll outcome varied between 1.05 and 6, with a mean of 4.20 (SD = 0.48; Mdn = 4.18) across all conditions. A one-sample Wilcoxon signed-rank test revealed the average die-roll outcome was larger than 3.5—the expected average die-roll outcome if first movers are, on average, honest, $z = 5.50$, $p < .001$.

Second movers. We calculated the proportion of doubles for each second mover. The proportion of doubles varied between 0% and 100%, with an average of 52.03% (SD = 16.85%; Mdn = 49.75%) across all conditions. A one-sample Wilcoxon signed-rank test revealed the proportion of doubles was larger than 16.66%—the expected proportion of doubles if second movers are, on average, honest, z =5.71, $p < .001$.

Heterogeneity. Similar to meta-analysis on (dis)honesty in individual settings (Gerlach et al., 2019), the heterogeneity of first and second movers' reports was high (I^2, the percentage of effect-size variance that is not caused by sampling error: for first movers, 93.9%, and for second movers, 93.5%; τ^2, the variance between conditions: for first movers, .197, and for second movers, .028). One reason for the high heterogeneity is that the different conditions (i.e., manipulations) across studies strongly affected participants' reports. For instance, in one study (Gross & De Dreu, working paper), participants were pre-classified as rule followers or violators and then assigned into dyads where two rule followers (violators) interacted with one another. In a condition in which two rule followers interacted with each

other, reports were closer to honesty (first movers' average die-roll reports: 3.86; second movers' proportion of doubles: 36.51%), whereas in a condition in which two rule violators interacted with each other, reports were further away from honesty (first movers' average die-roll reports: 4.28; second movers' proportion of doubles: 58.33%).

Outliers analysis did not change the I^2 much (first movers: eight conditions removed, I^2 changed from 93.9% to 89.1%; second movers: nine conditions removed, I^2 changed from 93.5% to 83.2%). Leave-one-out analysis in which we removed each condition and re-calculated the heterogeneity and effect size revealed that no one particular condition affected the effect size (first movers: I^2 ranged between 92.03% and 93.48%; effect size ranged between 4.17 and 4.21; second movers: I^2 ranged between 92.45% and 93.65%; effect size ranged between 50.95% and 52.38%). Thus, analyses suggest the high heterogeneity is driven by the overall variance in manipulations across conditions. Our analyses throughout the paper are conducted on the raw data and control for such heterogeneity by specifying random intercepts for conditions and groups. Nevertheless, similar to previous work that statistically summarizes previous findings (Gerlach et al., 2019), results obtained here should be interpreted with care.

Does partners' (dis)honesty correlate?

To test whether participants' (dis)honesty is correlated with their partners', we conducted two types of analyses. First, as in Weisel and Shalvi (2015), we analyzed the data on the dyadic level by classifying participants as brazen liars or not, and assessing whether an association exists between the brazenness of the first and second movers. Second, for a more fine-tuned analysis, we tested the association between the partners' behavior on a round-by-round basis. Specifically, we tested whether (i) higher die-roll outcomes reported by first movers are correlated with a higher likelihood of second movers to report a double, and whether (ii) second movers reporting of a double (versus not) is correlated with higher die-roll outcomes reported by first movers in the next round.

Dyad-level analysis. Analyzing the entire sample (43 conditions, with 3,264 participants, and 51,640 observations), we classified participants as brazen liars or not. As in Weisel and Shalvi (2015), we classified each first mover as a brazen liar if this participant reported the highest die-roll outcome, 6 in all rounds, and as non-brazen otherwise. Further, we classified each second mover as a brazen liar if this participant reported a double in all rounds, and as non-brazen otherwise. Overall, 17.07% of first movers and 30.03% of second movers were brazen liars.

If first and second movers do not affect each other, the likelihood of one person being a brazen liar should not depend on whether one's partner was a brazen liar or not. If, on the other hand, partners do affect each other, participants should be more likely to be brazen liars when their partners are brazen liars than when they are not. Indeed, results revealed that when first movers were brazen liars, second movers were more likely to be brazen liars as well. When the first mover was brazen, 69.90% of second movers were brazen as well (353 of 505 cases). When the first mover was not brazen, only 21.81% of second movers were brazen (535 of 2452 cases), $\chi^2(1) = 458.45$, $p < .001$. Similarly, when second movers were brazen liars, first movers were more likely to be brazen liars as well. When the second mover was brazen, 39.75% of first movers were brazen as well (353 of 888 cases). When the second mover was not brazen, only 7.34% of first movers were brazen (152 of 2069 cases), $\chi^2(1) = 458.45$, $p < .001$. Thus, analyzing the data on a dyadic level suggests that when one participant maximally lied to obtain the highest pay, her partner was more likely to do so as well. Results remain the same when employing an alternative classification procedure for brazenness, specifically, when first and second movers were classified as brazen liars if they reported 6s (for first movers) or doubles (for second movers) in at least 90%, 80%, and 70% of rounds.

Lastly, we defined four dyad types: (i) Brazen first mover, brazen second mover; (ii) brazen first mover, non-brazen second mover; (iii) non-brazen first mover, brazen second mover; and (iv) non-brazen first mover, non-brazen second mover; see Figure 5.2. Out of all dyads (n = 2,957), 11.94% were brazen first-mover, brazen second-mover dyads, and 64.83% were non- brazen first-mover, non-brazen second-mover dyads. Moreover, 18.09% were non-brazen first-mover, brazen second-mover dyads, and only 5.14% were brazen first-mover, non-brazen second-mover dyads.

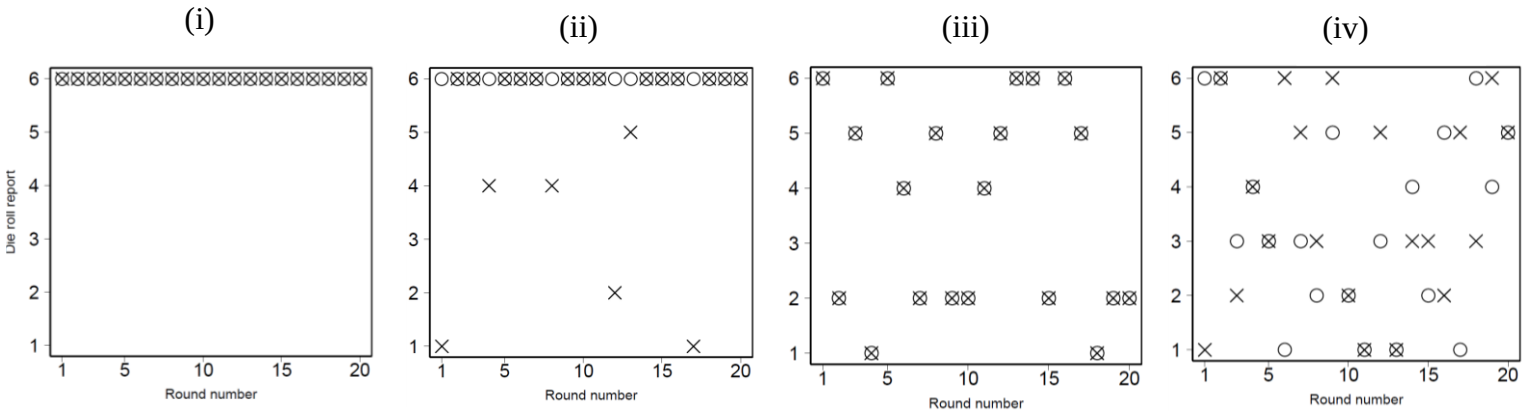

*Figure 5.2***.** Dyad types. The Y-axis represents the die-roll outcome participants reported (1-6), and the X-axis represents the round number. An "O" represents a first mover's reports; an "X" represents a second mover's reports on the same round. (i) Brazen first mover, brazen second mover; (ii) brazen first mover, non-brazen second mover; (iii) non-brazen first mover, brazen second mover; and (iv) non-brazen first mover, non-brazen second mover. The dyad types presented here are taken from the data gathered in this paper.

Round-by-round analysis. In the round-by-round analysis, we focused on conditions in which the roles of first and second movers were fixed. Doing so allowed us to isolate the association between partners' behaviors, without considering the experience each participant had engaging in the task in their partner's role. Further, because we were interested in the association between the dyad members' behavior over multiple rounds, we restricted our analysis to conditions in which the experimenters recorded the round number. Overall, we conducted the analysis on 31 conditions, with 2,408 participants, and 41,240 observations, split evenly between first and second movers.

For every unique dyad, we recoded the round numbers such that all rounds in which unique dyads interacted with each other were coded from round number = 1 to the last round of that dyad's interaction. For instance, if first mover A1 and second mover B1 interacted with one another for six rounds, their round numbers were recoded as one through six. If, then, first mover A1 interacted with second mover B2 for nine rounds, the new dyad's round numbers were recoded as one through nine. Overall, the analysis drew on 2,137 unique dyads who interacted with each other for 8.32 rounds, on average (*SD* = 6.32).

Analysis revealed partners' behavior is correlated. We conducted a linear regression predicting first movers' die-roll reports from whether second movers reported a double (vs. not) in the previous round, specifying a random intercept for each dyad, group, and condition. Results revealed first movers' die-roll reports were higher when second movers reported a

double in the previous round (M = 4.59, SD = 1.69) than when they did not (M = 3.94, SD = 1.72), b = .107, t = 3.86, p < .001. This effect remained significant when controlling for first movers' own report in the previous round, and the round number; see Table 5.1.

We further conducted a logistic regression predicting whether second movers reported a double from the first movers' report, specifying a random intercept for each dyad, group, and condition. Results revealed that the higher the first movers' die roll reports were, the more likely second movers were to report a double, b = .241, z = 19.18, p < .001. This effect remained significant when controlling for second movers' own report in the previous round and the round number; see Table 5.1. Estimating the average marginal effect of first movers' reports on second movers' likelihood of reporting a double, controlling for second movers' own behavior and the round number, revealed that an increase in first movers' die-roll outcome by 1 (e.g., first movers reporting 5 instead of 4) is associated with an increased likelihood of second movers reporting a double by 3.38%.

Table 5.1. *Association between first and second movers' behavior*

	Model 1: First movers' die roll report (1-6) at T:	Model 2: Second movers' report (0 = non-double,1 = double) at T:
Round number	.004*	.007^{+}
	(.002)	(.004)
First movers' die roll report (T-1)	.024***	
	(.007)	
Second movers' report (T-1) 0 = non-double, 1 = double	.098***	-.290***
	(.027)	(.053)
First movers' die roll report (T)		.242***
		(.013)
Random intercept for	Dyads, Groups Conditions	Dyads, Groups Conditions
# Observations	18483	18483
# Unique dyads	2137	2137
# Groups	1060	1060
# Conditions	31	31

Note. A linear regression analysis predicting first movers' die-roll report (model 1) and a logistic regression analysis predicting second movers' reporting of a double (model 2) from the round number, the counterpart's past report, and the participant's own past report. $^{***}p$ < .001, $^{**}p$ < .05, $^{*}p$ <.05, ^{+}p < .10.

Because in the dyadic die-rolling task participants engage in the task over multiple rounds, endogeneity issues allow us to assess how partners' behavior is correlated. We are not able, however, to assess the influence one participant has on her partner (Manski, 1993). Restricting our analyses to the first and second rounds of the interaction allows us to assess how participants affect each other, because we can control for all previous reports, resolving the endogeneity issues. When restricting the analyses to the first and second rounds of the task result revealed that first movers' reports affect second movers' likelihood of reporting a double (b = .190, p < .001). However, the effect of second movers' reports (a double vs. not) on first movers' subsequent die-roll outcome, controlling for first movers' own report in the first round, was not significant (p = .963).

How does behavior develop over time?

A linear regression specifying a random intercept for each group and condition revealed that the higher the round number, the higher first movers' die-roll report, b = .008, t = 5.26, p < .001. Similarly, a logistic regression predicting whether second movers reported a double from the round number, specifying a random intercept for each condition and group, revealed that the higher the round number, the higher the likelihood of second movers reporting a double, b = .012, z = 4.79, p < .001. Thus, throughout the duration of the task, both first and second movers lied more.

How do financial incentives affect behavior in collaborative settings?

In individual settings, where participants engage in a task alone, financial incentives do not seem to affect participants' dishonesty. Abeler and colleagues (2019) assessed whether participants' reports were affected by the maximum amount (in 2015 USD) they could earn from lying. In the experiments summarized by Abeler and colleagues, the maximum payoffs from misreporting participants could earn, per round, was between $.041 and $50.81 across the various conditions. Results revealed participants' reports did not vary as a function of the maximum amount they could earn from misreporting. In another meta-analysis on dishonesty in individual settings, focusing on other tasks, Gerlach and colleagues (2019) found similar results. We thus adopted the same approach as these two meta analyses and tested whether the reports of first and second movers in the dyadic die-rolling task varied as a function of the maximum amount they could earn from misreporting (in 2015 USD).

First movers. Figure 5.3 presents the average die-roll report in each condition, as a function of the maximum amount first movers could earn (assuming second movers matched

first movers' reports, in 2015 USD). A linear regression predicting the die-roll report from the maximum amount, specifying a random intercept for each group and condition, revealed that the more first movers could earn from lying, the higher their reports were, $b = 0.83$, $t = 3.21$, $p = .002$.

Whereas the dyadic die-rolling task is a repeated task with multiple rounds, the vast majority of tasks that assess dishonesty in individual settings are one-shot. Thus, we further restricted our analysis to the first round of the dyadic die-rolling task. The same analysis, focusing only on the first round of the dyadic die-rolling task, revealed a trend showing that the more first movers could earn from lying, the higher their reports were, $b = 0.51$, $t = 2.03$, $p = .053$.

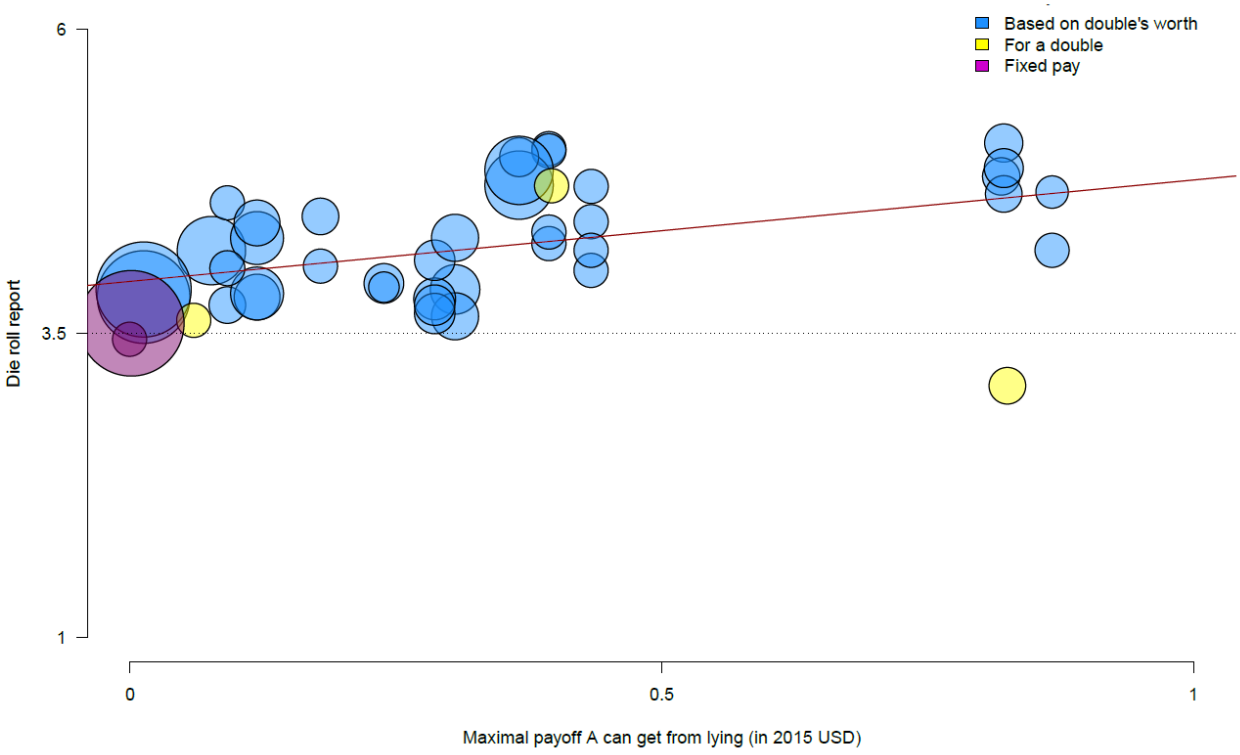

Figure 5.3. Average die-roll reports of first movers as a function of the maximum amount first movers could earn from lying (assuming second movers matched their reports; maximum minus minimum pay, converted to 2015 USD). The dashed line represents the average die-roll report expected if first movers were honest (3.5). Every bubble represents a condition, and the size of the bubble corresponds to the number of first movers in that condition. Blue bubbles represent conditions in which first movers were paid based on the dyad reporting a double and the double's worth. Yellow bubbles represent conditions in which first movers were paid only based on whether the dyad reported a double, regardless of the double's worth. Purple bubbles represent conditions in which first movers received a fixed payment regardless of their reports.

Second movers. Figure 5.4 presents the average proportion of doubles second movers reported in each condition, as a function of the maximum amount second movers could earn by matching first movers' reports. A logistic regression predicting whether second movers reported a double (vs. not) from the maximum amount they could earn, specifying a random intercept for each group and condition, revealed the more second movers could earn, the more likely they were to report a double, $b = 4.09$, $z = 20.51$, $p < .001$. The same analysis focusing on the first round of the dyadic die-rolling task revealed the same results: the more second movers could earn from reporting a double, the more likely they were to report it, $b = 2.02$, $z = 4.75$, $p < .001$. Thus, in the dyadic setting, people seem to be sensitive to the financial incentive to lie, whereas in the individual setting, they are not.

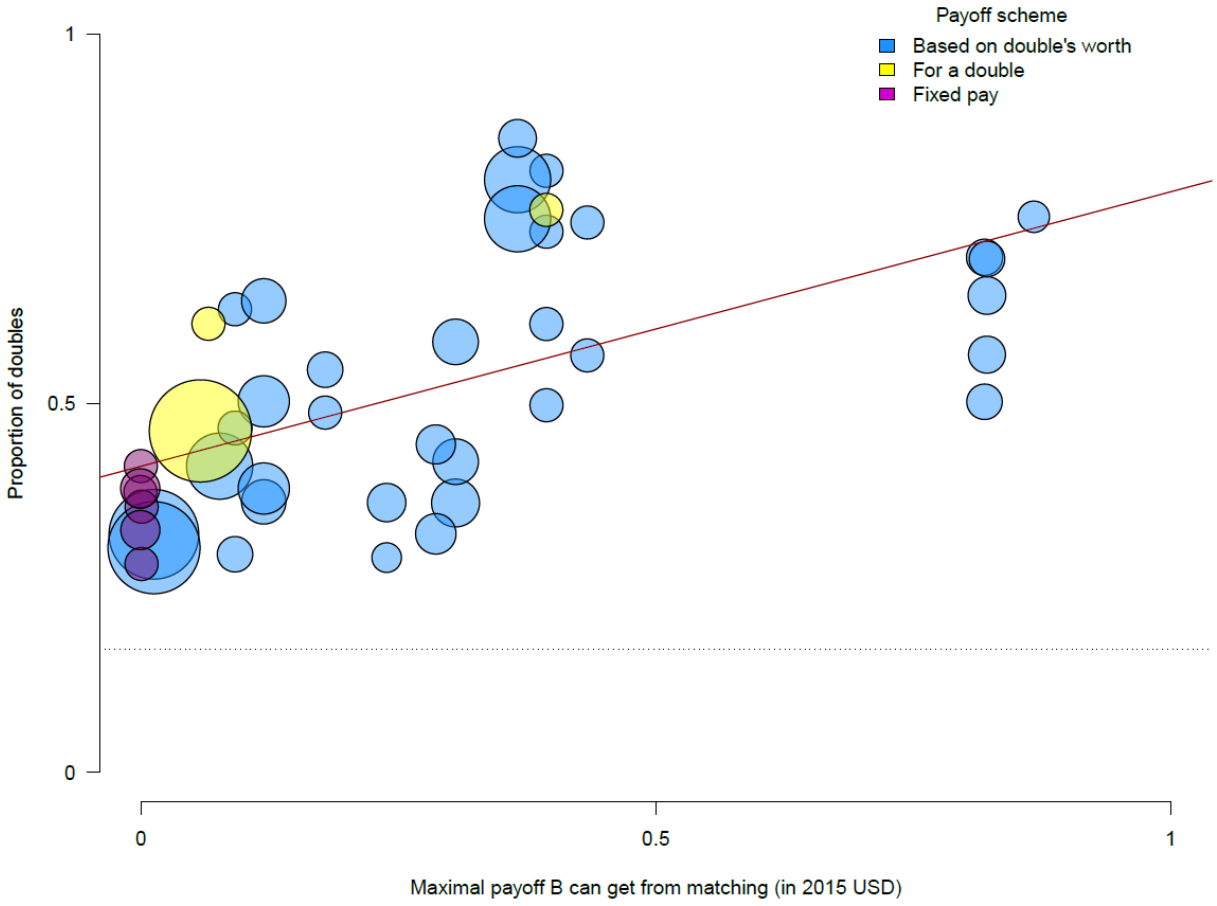

Figure 5.4. Proportion of doubles reported by second movers, as a function of the maximum amount they could earn from matching the first movers' reports in each condition (assuming first movers reported the highest die roll outcome, 6; maximum minus minimum pay, converted to 2015 USD). The dashed line represents the proportion of doubles expected if second movers reported honestly (16.66%). Every bubble represents a condition, and the size of the bubble corresponds to the number of second movers in that condition. Blue bubbles represent conditions in which second movers were paid based on the dyad reporting a double and the double's worth. Yellow bubbles represent conditions in which second movers were

paid only based on whether the dyad reported a double, regardless of the double's worth. Purple bubbles represent conditions in which second movers received a fixed payment regardless of their reports.

Does the prevalence of (dis)honesty differ between collaborative and individual settings?

We compare participants' reports in the dyadic die-rolling task to equivalent individual die-rolling tasks. Because the first and the second mover have different roles and thus different dependent variables, we compare the reports of each mover with a different, equivalent individual task. Further, because the majority of the individual die-rolling tasks are one-shot, whereas the dyadic die-roll task is a repeated task, we compare the reports in the first round of the dyadic die rolling task with the reports in the individual tasks. We compared the data we gathered on the dyadic die-rolling task with the data published by Abeler and colleagues (2019) who conducted a meta-study on (dis)honesty in individual settings.

First movers. In the dyadic die-rolling task, first movers roll a die and report the outcome. The equivalent task in an individual setting is the commonly used die-rolling task (e.g., Fischbacher & Föllmi-Heusi, 2013; Shalvi, Dana, Handgraaf, & De Dreu, 2011) in which participants roll a die privately, report the outcome, and receive pay based on the outcome they report. We thus compared the reports of first movers in the dyadic die-rolling task with the reports of participants in this individual task. Specifically, we focused on the version of the individual task in which the payoff scheme was the same, that is, where reporting higher die-roll outcomes corresponded to higher pay (k = 51). Similarly, in the dyadic die-rolling task, we focused only on conditions in which first movers were financially incentivized to report high die-roll outcomes, that is, conditions in which first movers were paid only if the dyad reported a double, and where the higher the double's worth, the higher the payment first movers received (k = 34).

A linear regression predicting the standardized score, specifying a random intercept for each condition, revealed no difference between the die-roll reports in the individual task (M = 4.14, SD = 1.64) and the first round of the dyadic task (M = 4.11, SD = 1.74), b = .041, t = .404, p = .687; see Figure 5.5A. Controlling for the maximum amount (in 2015 USD) participants could earn from misreporting did not change the results, p =.418.

The comparison between the individual task and the reports of first movers in the dyadic task provides initial evidence for how participants behave in the different settings. However, such comparison has limitations. In the individual task, participants report the die

roll outcome and are certain that they will earn the payoff that corresponds to their reports. In the dyadic task, however, first movers report an outcome but depend on the second movers to report the same outcome in order to get paid. That is, in the dyadic task, first movers are uncertain their reports will be realized. Thus, the two settings vary in participants' certainty about their dishonesty becoming profitable. The difference in participants' certainty does not exist when considering the behavior of second movers. Because second movers determine whether the dyad will get paid (by matching first movers' report), second movers are certain their reports will affect their pay, just as participants are in the individual setting.

Second movers. In the dyadic die-rolling task, second movers learn about first movers' reports, and then roll a die and report the outcome. If second movers report a double, the dyad gets paid. As such, the most comparable task in an individual setting is a task in which participants are instructed to report one outcome out of six possible outcomes, and get paid only if they report a certain outcome, for example, an individual task where participants are given a target number (e.g., 5) and are then asked to roll a die in private. They then report whether the outcome of the die roll was the target number. If they report "yes," they get paid; if they report "no," they do not get paid (Kajackaite & Gneezy, 2017; Hilbig & Hessler, 2013; Hilbig & Zettler, 2015). Other examples are tasks in which participants are instructed to think of a number between 1 and 6, then roll a die in private/observe a die-roll outcome on a computer screen, and report whether the outcome they observed is the one they thought of. If yes, they get paid; if not, they do not get paid (Kajackaite & Gneezy, 2017; Barfort, Harmon, Hjorth, & Olsen, 2015).

Both for participants in these individual tasks and for second movers in the dyadic task, the chance of honestly receiving pay was16.66%. However, in both cases, participants could lie to secure their pay. Comparing reports of participants in these individual tasks with the reports of second movers, we focused on conditions in the dyadic die-rolling task in which the second movers had a financial incentive to report a double (k = 35), that is, conditions in which second movers received higher pay for reporting a double than for reporting a non-double.

A logistic regression predicting whether participants reported the relevant-for-pay outcome, specifying a random intercept for each condition revealed second movers were more likely to report the relevant-for-pay outcome in the first round of the dyadic task (M = 46.41%, SD = 49.88%) than participants in equivalent individual tasks (M = 40.61%, SD = 49.12%), b = .629, z = 3.73, p < .001; see Figure 5.5B. Specifically, second movers reported

the relevant-for-pay outcome 14% more [46.41/40.61 = 1.14] than participants in the individual tasks. This difference remained significant when controlling for the maximum amount (in 2015 USD) participants could earn from reporting the relevant-for-pay outcome, b = .674, z = 3.42, p < .001.

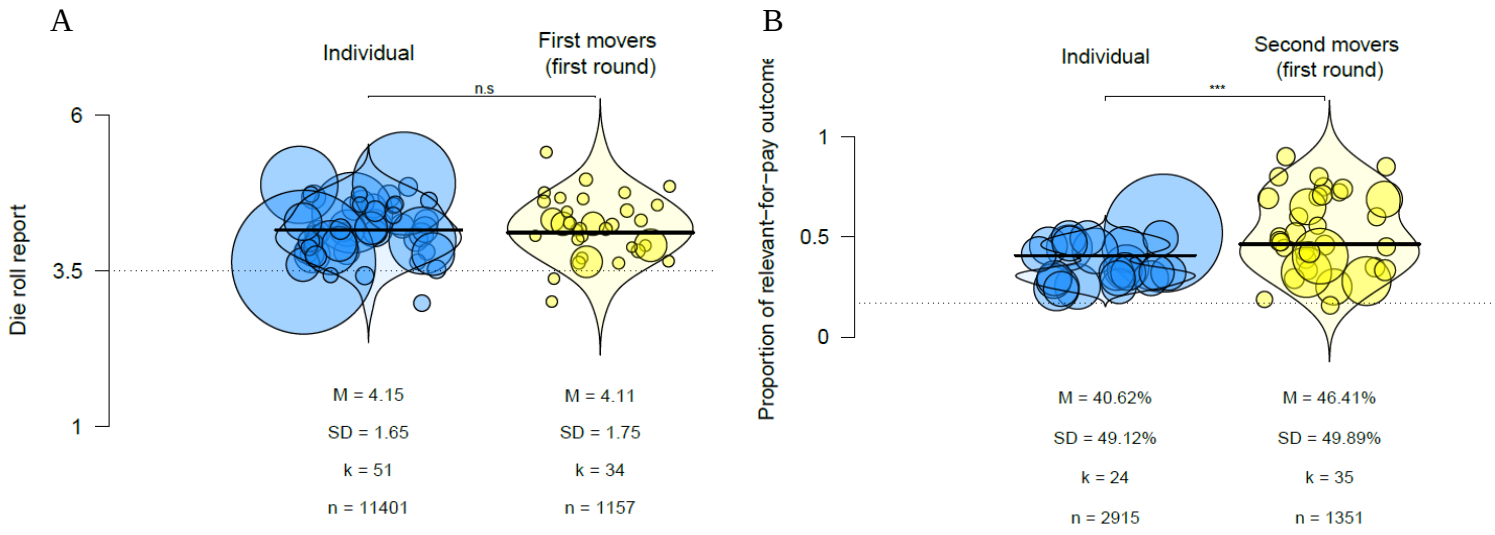

Figure 5.5. The (A) mean die-roll report and (B) proportion of reporting the relevant-for-pay outcome in individual tasks (data retrieved from Abeler et al., 2019), and the first round of the dyadic die-rolling task. Dashed lines represent honesty (a mean of 3.5 for first movers and the equivalent individual task, and 16.66% of reporting the relevant-for-pay outcomes for second movers and the equivalent tasks). Each bubble represents a condition, and the size of the bubble is proportional to the number of participants in that condition. In the dyadic die-rolling task, we excluded four conditions in which the round number was not recorded. $^{***}p < .001$.

Situational and personal factors. Lastly, we tested whether situational and personal factors affect behavior in the dyadic die-rolling task. Specifically, we tested whether (i) a lab versus field study, (ii) having a fixed role (as a first or second mover) throughout the task, and (iii) having a fixed partner throughout the task affect first and second movers' reports. Further, we tested whether participants' gender and age are associated with their reports.

Situational factors. We ran a linear regression predicting first movers' die-roll reports from whether the study was run in the lab, and whether participants had a fixed role and a fixed partner, specifying a random intercept for groups and conditions. Results revealed a significant effect for lab studies, b = .57, t = 2.85, p = .006. First movers reported higher die-roll outcomes when the study was run in the lab (M = 4.34, SD = 1.73, k = 35) than when it was run in the field (M = 3.83, SD = 1.73, k = 8). Further, we found a trend for a fixed role, b

= .38, t = 1.89, p = .065. First movers' die-roll outcomes were somewhat higher when they had a fixed role throughout the task (M = 4.28, SD = 1.74, k = 35), compared to not (M = 4.07, SD = 1.73, k = 8). We found no effect for a fixed partner, p = .454.

Similarly, for second movers, we ran a logistic regression predicting second movers' likelihood of reporting a double from whether the study was run in the lab, and whether participants had a fixed role and fixed partner, specifying a random intercept for groups and conditions. Results revealed a significant effect for lab, b = 1.61, z = 4.42, p < .001. Second movers were more likely to report a double when the study was run in the lab (M = 59.46%, SD = 49.09, k = 35) than when it was run in the field (M = 32.95%, SD = 47.01, k = 8). Further, we found a significant effect for a fixed role, b = .97, z = 2.65, p = .007. Second movers were more likely to report a double when they had a fixed role throughout the task (M = 56.46%, SD = 49.80, k = 35) than when they did not (M = 45.55%, SD = 49.58, k = 8). We found no effect for a fixed partner, p = .297.

Personal factors. A linear regression predicting first movers' die-roll reports from participants' gender and age, specifying a random intercept for each participant, group, and condition, revealed that, on average, males reported higher die-roll outcomes (M = 4.53, SD = 1.72) than females (M = 4.21, SD = 1.75), b = .25, t = 4.27, p < .001. We found an age trend, suggesting that the older first movers were, the somewhat lower their die roll reports were, b = -.008, t = -1.74, p = .081.

Similarly, for second movers, we ran a logistic regression predicting whether second movers reported a double from participants' gender and age, specifying a random intercept for each participant, group, and condition. Results revealed that, on average, males were more likely to report a double (M = 65.48%, SD = 47.54) than females (M = 57.99%, SD = 49.35), b = .53, z = 3.88, p < .001. We found an effect for age, showing that the older second movers were, the less likely they were to report a double, b = -.02, z = -2.15, p = .031.

Discussion

In line with the recently growing interest in group ethical decision-making (Cohen et al., 2009; Conrads et al., 2013; 2017; Muehlheusser et al., 2015; Kocher et al., 2017; Pulfrey et al., 2018; Sutter, 2009), here we conduct the first meta-study of (dis)honesty in a collaborative, interactive setting. Focusing on the dyadic die-rolling task (Weisel & Shalvi, 2015), we meta-analyze the dynamics that develop between the dyad members throughout the task. Consistent with prior work (Gross et al., 2018; Weisel & Shalvi, 2015), we find partners'

behavior is correlated. Specifically, first movers are more likely to be brazen liars when second movers are brazen liars than when they are not. Similarly, second movers are more likely to be brazen liars when first movers are brazen liars than when they are not. Looking at how dyads' relationship develops over time, we further find that when first movers report high die-roll outcomes, second movers are more likely to report doubles, and that when second movers report doubles (vs. not), first movers increase their reports in the next round.

The reasons for the association between the behavior of first and second movers remain an open question. Two main mechanisms might drive this effect. The first is participants' response to financial incentives. When a first mover reports a rather high die-roll outcome, the second mover's temptation to report a double increases because matching a high outcome is more profitable than matching an alternative, lower outcome. Further, when a second mover matches the first mover's report, the first mover might expect the second mover to match the next report as well. The expectation of the next report to be matched may motivate the first mover to report higher die-roll outcomes in the next round. Thus, the mere financial profits participants can gain from lying might drive, at least to some extent, the association between the first and second movers' reports. Our findings that participants in the dyadic die-rolling task are sensitive to the financial incentives to lie support this mechanism.

The second mechanism that might account for the association between partners' behavior is the exposure to social norms. People are more likely to violate the rules and lie when they see others engage in such behavior (e.g., Köbis et al., 2015; 2019; Gino, Ayal, & Ariely, 2009; Leib & Schweitzer, 2020). Learning that one's partner lied suggests such behavior is common and perhaps even justifiable (Gino, Ayal, & Ariely, 2009; Leib & Shalvi, 2020). As such, the mere exposure to one's partner's lies allows people to lie without spoiling their moral self-image (Mazar et al., 2008). One interesting avenue for future work is to explore the extent to which each of the mechanisms—the financial incentives to lie and the exposure to one's partner's behavior—drives the association between the partners' behaviors. An experimental setting in which the financial incentives remain fixed but the exposure to one's partner's behavior varies seems like a promising first step toward addressing this question.

Testing the development of behavior over time, we find that as the task progresses, both first and second movers lie more. Based on prior work, this finding does not seem to be the case in individual settings. Meta-analytical findings by Abeler and colleagues (2019) show that overall, repeated individual tasks lead to lower reports than one-shot tasks. Further,

additional work measuring tax evasion (Kogler, Mittone, & Kirchler, 2015) and bribery (Köbis, van Prooijen, Righetti, & Van Lange, 2017) find no support for a gradual increase in ethical rule violations over time. However, because some prior work did find ethical rule violation increases over time (Fischbacher & Föllmi-Heusi, 2013; Welsh, Ordóñez, Snyder, & Christian, 2015), future work should directly compare the effect of time on (dis)honesty in collaborative and individual set-ups.

As we mentioned above, our analysis further revealed that in the dyadic die-rolling task, participants are sensitive to the financial incentive to lie. The more money they can earn from doing so, the higher the die-roll outcomes first movers report and the more doubles the second movers report. This finding is in contrast to the results obtained from previous meta-analyses on (dis)honesty in individual settings showing that in individual die-rolling tasks, the maximum amount participants can earn from misreporting does not affect their reports (Abeler et al., 2019; Gerlach et al., 2019). Additionally, conducting initial comparisons between the prevalence of (dis)honesty in individual and collaborative settings, we found that compared with participants in equivalent individual tasks, second movers in the first round of the dyadic die-rolling task reported the relevant-for-pay outcome 14%% more. This finding is in line with prior work suggesting people lie more in groups than alone (Cohen et al., 2009; Conrads et al., 2013; Gross et al., 2018; Kocher et al., 2017; Weisel & Shalvi, 2015)

Taking into account these findings, another question that remains open is why people (i) lie more and are (ii) more sensitive to financial incentives in the dyadic die-rolling task than in equivalent individual tasks. When people collaborate and interact with each other, different motivations and psychological mechanisms come into play than when they act alone. These mechanisms might explain the differences we found between the two settings. First, as we mentioned above, in a collaborative setting, people are exposed to their partner's lies, which might free them to lie as well, in turn increasing the prevalence of dishonesty.

Second, in the collaborative setting, one's lies lead to a prosocial consequence, which is absent from the individual setting. Lying in the dyadic die-rolling task benefits not only the person who lies but also one's partner. Prior work found that people lie more when lies benefit others as well as themselves than when lies benefit only themselves (Gino, Ayal, & Ariely, 2013; Wiltermuth, 2011). Further, people lie to benefit others even when they cannot benefit themselves in the process (Leib, Moran, & Shalvi, 2019). Thus, the prosocial consequence of dishonesty in the collaborative setting seems to be responsible, at least to an

extent, for the higher levels of dishonesty and sensitivity to incentives in the collaborative setting.

A third mechanism that might come into play in collaborative but not individual settings is diffusion of responsibility (Darley & Latané, 1968). When engaging in a task with others, the dyad members share the responsibility for (dishonest) outcomes they generate. When engaging in a task alone, however, they do not share the responsibility. Therefore, sharing responsibility for joint dishonest outcomes might serve as a justification to lie (Shalvi et al., 2015) and reduce the psychological cost of lying (Lundquist, Ellingsen, Gribbe, & Johannesson, 2009). In turn, the lower psychological cost of lying might push people to lie more in the dyadic than in the individual task, especially when financial incentives are high. An important next step for future research is to isolate the mechanisms that shape behavior in individual and collaborative settings and test the extent to which exposure to social norms, prosocial consequences, and diffusion of responsibility contribute to the higher (i) level of dishonesty and (ii) sensitivity to incentives in the collaborative setting.

We further find that several situational and personal factors drive behavior in the dyadic die-rolling task. Our results are consistent with meta-analysis in individual settings (Gerlach et al., 2019), revealing males lie more than females, young participants lie more than old ones, and overall dishonesty is higher in the lab than in the field. Unique to the dyadic die-rolling task, we further assessed how having a fixed role (vs. not) and a fixed partner (vs. not) affect participants' behavior. Aggregated results reveal fixed roles lead to a higher level of dishonesty, but fixed partners do not. Future work could systematically vary whether roles and partners are fixed in one experimental setup (Abbink, 2004; Gross et al., 2018) to corroborate the findings obtained here.

Another question to explore is whether, in situations where people can profit from dishonesty, they prefer to act alone or in collaboration with others. People might anticipate that collaborative settings will tempt them to lie more, and in order to not succumb to such temptation, they prefer to self-select into individual settings (Fishbach & Woolley, 2015; Shalvi, Handgraaf, & De Dreu, 2011a; 2011b). Further, people might be especially concerned about their reputation as honest individuals and thus prefer to engage in individual tasks, where others do not observe their actions. Thus, individuals who wish to avoid tempting situations, as well as those who want to lie, might shy away from collaborative settings. Alternatively, collaborative settings might be more attractive to individuals because they enable individuals to lie more, and might even allow some individuals to profit from others'

rule violations without violating the rules themselves (ethical free riding; Gross et al., 2018). Testing how attractive individual and collaborative settings are, and the extent to which the ability to choose between the two situations shapes people's (dis)honesty, presents an intriguing question to explore.

In the current analyses, we test how people behave in situations in which honesty and collaboration are pitted against each other. Another open question is how people with different personality traits behave in such situations. Could otherwise desirable traits, such as agreeableness, lead to more dishonesty in collaborative settings? Cooperating and meeting other's expectations is especially important for agreeable people (Ashton et al., 2004). As such, could agreeable people be more inclined to lie when working with a partner than when working alone? Further, could agreeable individuals lie even more when they have a higher impact on their partner's outcome, for instance, as a second mover (compared to a first mover) in the dyadic die-rolling task? Understanding how personality traits drive behavior in situations where collaboration and honesty are in conflict is both theoretically interesting and practically relevant for hiring and placement policies.

Lastly, another promising direction for future research is to discover the ways in which both collaboration and honesty can be maintained. Interacting and collaborating with others provides many advantages. In collaborations, people can exchange private information, bring together their unique skills, and achieve high-quality decisions and innovation (Fraidin, 2004; Laughlin, Hatch, Silver, & Boh, 2006; West & Anderson, 1996). As such, collaboration should be maintained and encouraged. At the same time, honesty should be sustained as well. Identifying the mechanisms that contribute to the higher prevalence of dishonesty in collaborative than in individual settings is thus a first step toward finding interventions that can help maintain high levels of both collaboration and honesty. For instance, if the diffusion of responsibility indeed drives the higher level of dishonesty in collaborative settings, a promising next step could be to highlight personal responsibility in collaborative settings, and experimentally test whether such intervention helps mitigate collaborative dishonesty.

Conclusions

Many, if not most, important decisions occur in collaboration with others. Thus, understanding how collaborative settings shape the overall prevalence and the dynamics of ethical decision-making is important. We focus on a collaborative setting, as captured by the dyadic die-rolling task, in which honesty and collaboration are pitted against each other. Meta-analyzing behavior in the task, we find that in collaborative settings, partners' behavior

is correlated. If one person lies, the partner is more likely to lie as well. Further, we find participants lie more as the task progresses and as incentives to lie increase. Our findings further reveal people lie more in the collaborative task than in equivalent individual tasks. As the first meta-study on (dis)honesty in collaborative, interactive settings, our findings further open up new intriguing avenues for future research aiming to better understand collaborative dishonesty.