Economia Politica
https://doi.org/10.1007/s40888-025-00361-6

ORIGINAL PAPER

How does dishonesty diffuse? An experiment comparing spread and concentrated incentives

Beatrice Braut[1] · **Marco Piovesan**[2]

Received: 13 February 2024 / Accepted: 1 February 2025

Abstract
We experimentally test if the propensity to cheat changes when subjects face different distributions of the same monetary incentive. Such a manipulation represents a costless intervention in the many situations where multiple reports are required simultaneously, like timesheets or expense reimbursements. In our experiment, subjects have to roll a die in private five times and report the outcomes. In the LAST treatment, only the fifth reported roll determines the payoff, whereas in the SUM treatment, all five reports contribute proportionally to the payoff. The possible reward from reporting remains the same, but the two treatments differ in the effort required to obtain it: inflating one report or many. We find that subjects report significantly higher numbers in LAST compared to the average report in SUM. Looking at the total of reports, it does not differ by treatment since subjects do not inflate the reported rolls when there is no incentive. Dishonesty differs as there are more opportunities, having the positive effect of reducing the total profit obtained from cheating, but the negative effect of spreading the lies across all the opportunities. The total amount of cheating made remains stable across conditions, but it costs less when the incentive is spread on multiple reports.

Keywords Dishonesty · Lying · Self-reporting · Incentive

JEL classification C90 · D03 · D82

✉ Beatrice Braut
beatrice.braut@unige.it

Marco Piovesan
marco.piovesan@univr.it

[1] Department of Economics, University of Genoa, Via Vivaldi, 5, 16126 Genova, Italy

[2] Department of Economics, University of Verona, Polo Santa Marta, 37129 Verona, Italy

Published online: 18 February 2025

1 Introduction

Dishonest behavior has severe consequences for societies, with costs up to several billion Euro every year (e.g., European Commission, 2014). Small acts of dishonesty - like overreporting the number of hours worked or stealing from the office supply cabinets - are extremely costly for society and can undermine people's trust, leading to further erosion of honesty (Bucciol and Montinari, 2019). Unfortunately, dishonesty is a complex phenomenon to fight. According to Becker (1968), we have two main tools for this goal: reducing the benefit of dishonest action and increasing the probability (or magnitude) of punishment. More recently, behavioral economics has shown that there is a larger set of individual and contextual factors that also affect the likelihood of dishonesty (Abeler et al., 2019; Khalmetski & Sliwka, 2019; Heck et al., 2018; Gneezy et al., 2018; Pierce & Balasubramanian, 2015). In this context, benefits interact with many other aspects and their amount is not the most important factor: lying does not increase with the size of the incentives (Kajackaite & Gneezy, 2017).

In this work, we test if giving the same amount of money in different combinations can alter truth-telling. The experiment focuses on self-reporting of outcomes on which the payments directly depend. Self-reporting is a common tool in organizational monitoring and it represents the setting where cheating is studied the most, as demonstrated by the widespread use of the die-under-the-cup experimental paradigm. (Fischbacher & Föllmi-Heusi, 2013; Abeler et al., 2019). When an organization or government has to observe the outcome of many subjects it is common practice to ask employees and citizens to self-report, given the amount of monitoring costs on such large scales. Understanding how to reduce the amount stolen when unmonitored, and how to improve self-reporting accuracy, is crucial for the efficiency of these systems.

In a situation where subjects are asked to report multiple times, we test if the behavior when only one reporting opportunity is incentivized differs from when the same incentive is spread across all the opportunities. In any setting where self-reporting is required multiple times, it is always possible to change the payoff relevance of the specific reports. If the same monetary incentive has different effects depending on how it is distributed, without altering the number of reports to be done, this can offer a costless tool to influence (dis)honest behavior in contexts like the compilation of timesheets or the declaration of expenses.

We run a framed field experiment with 217 university students who play a modified version of the die roll paradigm (Fischbacher & Föllmi-Heusi, 2013). Our two main treatments differ in the number of reports that are payoff-relevant: payoffs depend only on one die roll (the last of five) or on five die rolls (their sum). In both treatments, subjects are asked to report the outcome of a roll five times. In this way, we simulate a situation where the objective is to know the (true) value of each report. Moreover, we keep constant the variability of the outcomes observed (asking subjects to roll the die before starting the task) and therefore the possible justifications for being dishonest (Shalvi et al., 2011). In different sub-conditions, we also manipulate the framing of the task (gain vs. loss framing) and the size of the incentives (small, medium, or large).

We find that subjects claim higher payments when the incentive is only attached to one report. In LAST the mean of the single incentivized report is higher than the mean of the five incentivized reports in SUM. On average, subjects' gains from reporting are lower when all reports count for payment. When we instead compare the total of all reports made, we find that it does not differ by treatment. Subjects do not inflate the reports for which they are not paid in LAST. The amount of cheating is overall the same but differently distributed. Following the monetary incentives, subjects cheat a little in each opportunity or a lot in the case of only one incentive. The distribution of incentives can change the distribution of dishonesty, but it does not alter its existence. In LAST, subjects on average steal more money with one report but they tell the truth in the others. Regression analysis confirms that the frame of the task, the size of the incentives, and personal characteristics do not influence the result.

Our work contributes to the debate on how dishonest spread relates to incentives. It is related to the lab experiment of Casal and Filippin (2023) where the authors compare one-shot and repeated reporting tasks designed to measure dishonesty. Contrary to our result, Casal and Filippin (2023) find only small difference in reporting between a treatment where subjects have to report a single number and one in which they report the sum of numbers. The authors suggest that this result is the balancing of the effect of observing many outcomes (positive) with the effect of reputational concerns (negative). However, in our experimental design, subjects report every single number, while in their design, subjects reported only once the sum of all the numbers. This peculiarity can make a significant difference since it changes the number of cheating acts to be done, affecting reputational concerns. Also, in our design, subjects see five outcomes in both treatments whereas, in Casal and Filippin (2023), the number of outcomes is not held constant. Consequently, in our study, the two treatments do not differ in the number of outcomes observed in theory. The fact that in our work subjects can observe more than one report and that not all the reports are relevant for payment can be another relevant difference between the two designs. In our experiment, subjects could feel licensed to be dishonest after seeing an initial convenient report that does not count. This would be in line with evidence observing higher cheating when providing justifications for dishonest acts comes easy (Shalvi et al., 2011, 2012).

Rilke et al. (2016) first analyzed a similar problem in the context of monitoring subjects' performance, comparing two policies: one with a series of reports and one with an overall report. They find that a one-by-one monitoring policy resulted in more cheating than an all-at-once policy, in line with their predictions but contrary to people's intuition (elicited in a separate sample in an incentivized manner). Our work differs from their paper since, in our treatments, we change the way payments are made and not the way reporting is done. Also, we use a die-rolling game instead of a trivia game. Subjects cheat differently when it is about effort or luck (Kajackaite, 2018). Moreover, in our case, cheating cannot be detected at the individual level but an a priori distribution of outcomes exists (and is the same for each subject).

Our experiment compares two situations where the reporting opportunities and the size of the incentives are identical, but the relationship between reports and incentives changes. We focus on a situation where the interest is to know the true value of multiple reports. Our simple and conservative intervention has multiple consequences: from one side it alters the size and the number of lies to be made, and there is evidence that each lie implies a cost and smaller lies are easier to tell (Shalvi et al., 2015); from the other side, it alters the salience of reputational concerns (Dufwenberg & Dufwenberg, 2018) and there is evidence that subjects care about being seen as honest. Ultimately, the two situations differ in the fact that there are or are not reports that do not count for the payment. By testing if and how cheating differs in these two situations, we can draw insights on how to design incentive schemes in tempting situations of multiple reporting like timesheets, deductions, and expense reimbursement reporting in general. Given that subjects do not inflate reports that are not payoff-relevant (Charness et al., 2019), knowing if the amount cheated in the single report is comparable to the cheating spread across the total of reports is key to evaluating the cost of the incentive schemes.

2 Experiment

2.1 Task

Subjects played a modified version of the die-roll paradigm originally introduced by Fischbacher and Föllmi-Heusi (2013), similar modifications have been implemented in other studies on dishonesty, such as Shalvi et al. (2012). In this task, subjects are asked to roll a die in private and are paid proportionally to what they report: a higher number gives the right to a higher payoff (with 0 as the lowest payoff). Cheating is not only possible but also undetectable at the individual level. In our experiment, subjects are asked to perform this task five times and how many of these are payoff-relevant depends on the treatment as explained in the next paragraph.

2.2 Treatments

Our treatment manipulation concerns the payment rule: in Treatment **LAST**, we paid participants according to only the last outcome reported (the fifth); in Treatment **SUM**, we paid participants according to the sum of the five outcomes separately reported. Figure 1 illustrates our treatment manipulation: the report sheet for LAST is shown on the left, while the SUM sheet is on the right. Table 1 illustrates the relationship between payoffs and the possible reports of LAST and SUM. Since SUM

LAST

Die-Roll	Outcome	
1st roll	___	TRIAL
2nd roll	___	TRIAL
3rd roll	___	TRIAL
4th roll	___	TRIAL
5th roll	___	Determining payoff

SUM

Die-Roll	Outcome	
1st roll	___	
2nd roll	___	
3rd roll	___	Sum = ___
4th roll	___	
5th roll	___	

Fig. 1 Report sheets in the two treatments. *Note* Report sheets in treatments LAST (on the left) and SUM (on the right)

Table 1 Report-payment scheme for the two treatments

LAST Report	6	-	5	-	4	-	3	-	2	-	1	-
Payoff	**π**	**9/10 π**	**8/10 π**	**7/10 π**	**6/10 π**	**5/10 π**	**4/10 π**	**3/10 π**	**2/10 π**	**1/10 π**	**0**	**0**
SUM Report	28,29, 30	26, 27	24, 25	22, 23	20, 21	18, 19	16, 17	14, 15	12, 13	10, 11	8, 9	5, 6, 7

Note Reports and corresponding payoffs in LAST and SUM treatment. **π** indicates the maximum payment possible, i.e. 20, 50, or 100 DKK

Table 2 Distribution of the sample across experimental variations

	20 DKK	50 DKK	100 DKK	Total
LAST	Gain 18; Loss 19	Gain 19; Loss 19	Gain 18; Loss 20	113
SUM	Gain 16; Loss 19	Gain 14; Loss 19	Gain 17; Loss 19	104

Note Subjects distribution across experimental variations in LAST and SUM

involves a wider range of reportable numbers, we adjusted the payoffs to equalize the expected values between treatments (with a maximum discrepancy of 5 DKK), even though this increased variation in SUM's payoffs.

In each treatment, we also vary the size of payoffs and framing. The maximum payoff subjects can win is 20, 50, and 100 DKK, and the instructions' frame can be either that or they will gain the amount of money, or they will avoid losing it. While these additional variations help us assess the robustness of the results, they are not the primary focus of this study. Table 2 provides the distribution of subjects in the two treatments across the different variations.

2.3 Procedure

The experiment took place as a framed field experiment in the University of Copenhagen campus (see pictures in Appendix B). Our sample includes 217 subjects frequenting the University (39% females, mean age 23, 113 subjects in Treatment LAST). Subjects received written instructions in English (Appendix C), the die, and the paper sheets where to submit the report. When the loss framing was implemented, participants received the maximum amount of money in an envelope, from which they would take the amount they earned, leaving the remainder behind. The roll, the report, and the payment took place in private.

2.4 Hypothesis and analysis plan

By employing our simple treatment manipulation, we introduce some differences between the two treatments that can influence treatment behavior. To gain the same amount obtained by over-reporting by 1 point in LAST, in SUM it is necessary to over-report by more than 1 point (at least 3). In LAST the expected distribution of (honest) reports is uniform with a mean equal to 3.5. In SUM instead, the expected distribution is normal with a mean equal to 17.5. While in LAST there is only one way to cheat, in SUM subjects can get the same outcome using different cheating strategies (e.g. subjects can get a 25 by reporting 5, 5, 4, 6, 5 or 3, 4, 6, 6, 6). There are 6 possible outcomes (and 6 payoffs) in LAST while in SUM there are 26 outcomes and the corresponding payoffs are also more numerous.

Combining the differences, SUM subjects face a more complex situation where there are multiple and more nuanced possibilities for cheating. In SUM, on one hand, there is the possibility of engaging in smaller lies, easier to justify (Ayal et al., 2015; Shalvi et al., 2015), but, on the other hand, the underlying theoretical distribution of outcomes can exacerbate the aversion to being perceived as a cheater (Dufwenberg & Dufwenberg, 2018). The present work aims to test if one of the two payment schemes is more valid to promote honesty in situations where multiple reports are required. The focus of the study is on multiple simultaneous reports where dishonesty is an issue. While the literature on dishonesty is concentrated mainly on singular reports, here, we provide insights into how the shape of incentives can alter dishonesty when facing multiple reports.

In the following, we will compare the observed reporting to the theoretical underlining distribution, therefore testing if cheating is present, and compare the overreporting in the payoff-relevant situations between treatments. This allows us to test if our simple manipulation of the payment scheme can be effective in changing cheating behavior and reducing the amount dishonestly gained. Focusing uniquely on our main treatment manipulation allows us to have a sample of around 110 independent observations per arm. Power calculations indicate that our pooled sample is able to detect a difference of 0.43 at the conventional values of 0.8 power and alpha 0.05. We will also run a regression to ensure that the result remains robust when considering the different levels of payment, the frame, and the subject's characteristics like field of study, gender, and age.

3 Results

In LAST, subjects report significantly higher with respect to the theoretical uniform distribution (average value 4.12, p-value < 0.001 for Kolmogorov-Smirnov test and Pearson chi-squared test). None of the first four reports, which are not incentivized, significantly differs from the expected uniform distribution (3.68, p-value 0.39; 3.49 p-value 0.41; 3.51 p-value 0.92; 3.54 p-value 0.99, Pearson chi-squared tests), in particular, none are lower. On average, subjects did not engage in underreporting to compensate for the higher values of the incentivized report and they did not over-report without a monetary incentive. In SUM, subjects report a sum of numbers significantly higher than expected (18.85 corresponding to an average of 3.77; p-value < 0.001, Wilcoxon signed-rank test). Moreover, the sum they report corresponds perfectly to the sum of the numbers reported. Subjects followed the rules and did not exploit alternative ways to cheat, such as inflating the sum instead of the reported numbers. Subjects do engage in overreporting when the report is payoff-relevant. This confirms the finding that subjects do not cheat where there is no reason to do so (Charness et al., 2019) and that they do not balance their over-reporting by consistently underreporting on other occasions.

Regarding incentivized reports, subjects cheat significantly more when only one report is incentivized compared to when all reports are (fifth report of 4.12 vs. mean of the five reports of 3.77, p-value = 0.005, Wilcoxon rank-sum test). This does not necessarily mean that in SUM they are cheating less often, but that they are stealing less in total. SUM treatment is, therefore, more efficient in minimizing the amount of extra profit "gained". Figure 2 shows the distribution of the values of the fifth outcome reported by the subjects in LAST (left panel) and the distribution of the values of the sum reported by subjects in SUM (right panel).

Comparing the sum of the five numbers reported in LAST and in SUM, we find that they do not differ from each other statistically (18.34 vs. 18.85, p-value = 0.33, Wilcoxon rank-sum test), while when we compare the fifth values we find them slightly higher in LAST (4.12 vs. 3.7, p-value = 0.08, Wilcoxon rank-sum test). This suggests that regardless of the incentive structure, the amount of cheating on the total is similar as if the dishonesty in the fifth opportunity in LAST was cut in pieces and redistributed in all the five opportunities in SUM. Incentivizing and monitoring each report instead of only a defined one just splits dishonesty into parts. Figure 3 shows the distribution of the values of the total sum of reports in the two treatments. The two distributions do not differ by treatment (two-sample Kolmogorov-Smirnov test of the

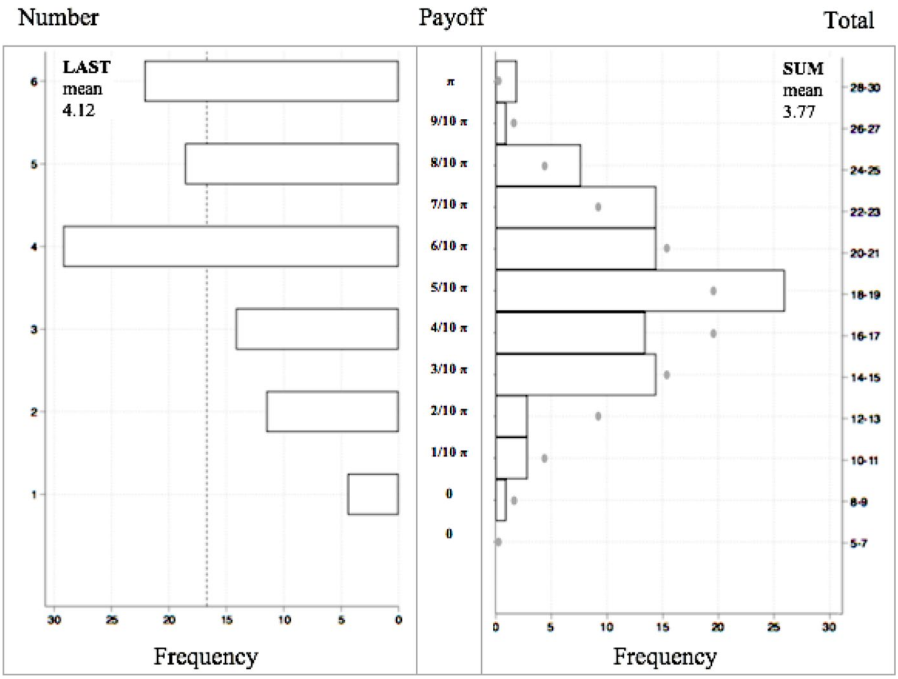

Fig. 2 Distribution of claims in LAST (on the left) and in SUM (on the right). *Note* On the left panel the distribution of the reports in LAST, on the right panel the distribution of the total of reports in SUM. The external axes indicate the numbers reported (from the single roll or the total of five rolls). The central panel indicates the payment categories. π indicates the maximum payment achievable, a gain of 20, 50 100 DKK, or an equivalent not-loss. Note that in SUM there are more feasible payment categories. The dotted line indicates the expected frequency in LAST and the dots represent the expected frequencies in SUM

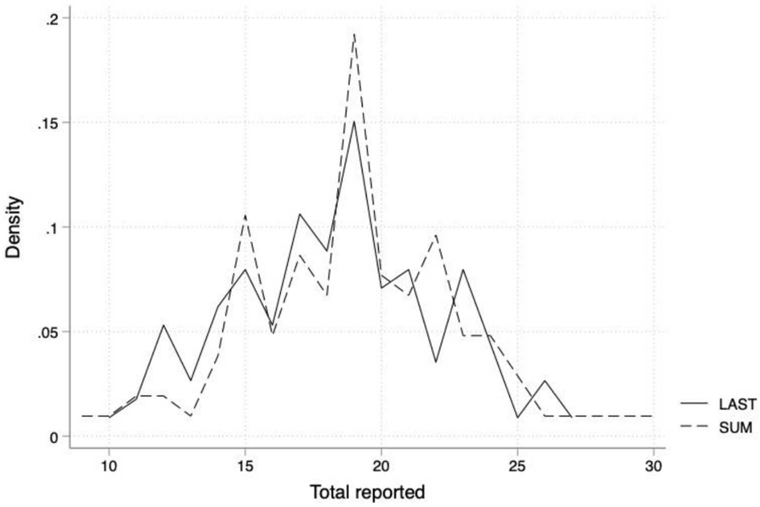

Fig. 3 Distribution of the total in LAST (solid) and in SUM (dashed). *Note* The figure shows the distribution of the total sum of reported claims for the two treatments. The solid line represents the LAST treatment, while the dashed line represents the SUM treatment

equality of distributions, p-value = 0.818). This evidence suggests that, overall, the total reports were altered in both cases, so subjects did not refrain from cheating even in SUM where it required more effort. While cheating happens in both treatments, thanks to the payoff distribution, overreporting is more evident and concentrated in LAST.

We further check that the background variables - the gain or loss frame, the size of payoffs, and the individual characteristics– gender, age, and field of study- have similar effects in both treatments. Using regression analysis– an ordered probit for LAST, and an OLS for SUM– we find that none of these variables significantly affects reporting in any treatment (Appendix A). We can therefore affirm that our treatment effect is not driven by a mediated effect of the background variables or of the individual characteristics.

4 Conclusion

In this work, we experimentally test if dishonesty differs in a situation where there is only one opportunity to profitably cheat (LAST) or where there are more (SUM), but the total amount to earn the required number of reports remain the same. We find that dishonesty is not cut down when there are fewer opportunities: subjects report, on average, a number higher than expected when they have the incentive to cheat once (treatment LAST) compared to when they have incentives to do it multiple times (treatment SUM). Moreover, we find evidence that dishonesty splits: while the mean of the incentivized reports is significantly different, the sum of all the reports made does not differ between the two treatments. Subjects engage in smaller lies and spread them in all the opportunities if they have to misreport more times for the same benefit, such that there is no difference looking at the total reported.

In our experiment we find that in LAST over-reporting is huge but limited to the only incentivized report; in SUM over-reporting is smaller but it is diffused across reports. This reveals that in choosing the scheme there is an important trade-off between how many lies and their size, and that a stable limit to dishonesty seems to exist: no matter the distribution of the incentives the total deviation from the true sum is the same. Further research is needed to determine whether such a limit is present in other experimental conditions and to test small variations of the design looking for the perfect recipe to promote honesty in multiple reporting settings. In particular, our results underline the importance of finding a modification able to balance the amount stolen and the honest reports. Such studies would benefit from collaboration with organizations seeking to reduce overreporting by clients or employees.

In conclusion, a simple change in how the same incentive is distributed across the same reporting opportunities has non-negligible effects and these differ by looking at the level of overreporting (and therefore over earnings) or their distribution. These results open the way to finding an optimal distribution of incentives to minimize dishonesty and its costs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40888-025-00361-6.

Funding Open access funding provided by Università degli Studi di Genova within the CRUI-CARE Agreement.

Data availability The data are available from the corresponding author upon request.

Declarations

Ethical approval This work was supported and approved by the Department of Economics of the University of Copenhagen and the Laboratory for Experimental Economics.

Conflict of interest The authors declare no conflict of interest.

References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *87*, 1115–1153.

Ayal, S., Gino, F., Barkan, R., & Ariely, D. (2015). Three principles to REVISE people's unethical behavior. *Perspectives on Psychological Science*, *10*(6), 738–741.
Becker, G. S. (1968). Crime and punishment: An economic approach. *The economic dimensions of crime* (pp. 13–68). Palgrave Macmillan.
Bucciol, A., & Montinari, N. (Eds.). (2019). Dishonesty in behavioral economics *Perspectives in behavioral economics and the economics of behavior*, Academic Press.
Casal, S., & Filippin, A. (2023). The effect of observing multiple private information outcomes on the inclination to cheat. *Economic Inquiry*.
Charness, G., Blanco-Jimenez, C., Ezquerra, L., & Rodriguez-Lara, I. (2019). Cheating, incentives, and money manipulation. *Experimental Economics*, *22*, 155–177.
Dufwenberg, M., & Dufwenberg, M. A. (2018). Lies in disguise–A theoretical analysis of cheating. *Journal of Economic Theory*, *175*, 248–264.
European Commission (2014). EU Anti-Corruption Report. Available at https://www.tagesschau.de/wirtschaft/eukorruptionsbericht100.pdf
Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *The Economic Journal*, *114*(497), 641–659.
Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise: An experimental study on cheating. *Journal of European Economic Association*, *11*, 525–547.
Gneezy, U., Kajackaite, A., & Sobel, J. (2018). Lying aversion and the size of the lie. *American Economic Review*, *108*(2), 419–453.
Heck, D. W., Thielmann, I., Moshagen, M., & Hilbig, B. E. (2018). Who lies? A large-scale reanalysis linking basic personality traits to unethical decision making. *Judgment and Decision Making*, *13*(4), 356.
Kajackaite, A. (2018). Lying about luck versus lying about performance. *Journal of Economic Behavior & Organization*, *153*, 194–199.
Kajackaite, A., & Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior*, *102*, 433–444.
Khalmetski, K., & Sliwka, D. (2019). Disguising lies—image concerns and partial lying in cheating games. *American Economic Journal: Microeconomics*, *11*(4), 79–110.
Pierce, L., & Balasubramanian, P. (2015). Behavioral field evidence on psychological and social factors in dishonesty and misconduct. *Current Opinion in Psychology*, *6*, 70–76.
Rilke, R. M., Schurr, A., Barkan, R., & Shalvi, S. (2016). One-by-one or all-at-once? Self-reporting policies and dishonesty. *Frontiers in Psychology*, *7*, 113.
Shalvi, S., Dana, J., Handgraaf, M. J., & De Dreu, C. K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, *115*(2), 181–190.
Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science*, *23*(10), 1264–1270.
Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science*, *24*(2), 125–130.