HAL
open science

A Reference Point Bias in Judging Cheaters

Sophie Clot, Gilles Grolleau, Lisette Ibanez

To cite this version:

Sophie Clot, Gilles Grolleau, Lisette Ibanez. A Reference Point Bias in Judging Cheaters. Journal of Economic Psychology, 2022, 89, pp.102485. 10.1016/j.joep.2022.102485 . hal-03528307

HAL Id: hal-03528307

https://hal.inrae.fr/hal-03528307v1

Submitted on 17 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Reference Point Bias in Judging Cheaters

By SOPHIE CLOT*, GILLES GROLLEAU AND LISETTE IBANEZ

Sophie Clot (Corresponding author)
Department of Economics, University of Reading, United Kingdom
E-mail : s.clot@reading.ac.uk

Gilles Grolleau (Co-author)
ESSCA School of Management
CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France
E-mail : grolda@gmail.com

Lisette Ibanez (Co-author)
CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France
E-mail : ibanez@supagro.inrae.fr

Do observers judge differently a wrongdoer when s/he does not exploit a situation to its maximum extent? Using a social intuitionist perspective and taking into account the reference point bias, we hypothesize that people will judge a wrongdoing less severely when the situation is not exploited to its fullest extent. We run two experimental surveys in France and examine whether various wrongdoings performed in the business realm (overcharging travel expenses, overstating work hours, pollution) are judged less severely when different reference points are suggested: (i) no explicit reference point is mentioned, (ii) a reference point is mentioned and the maximum extent is reached, (iii) a reference point is mentioned but the maximum extent is not reached and (iv) the participant is invited to elicit a reference point corresponding to what s/he considers as the maximum extent. Our findings support that participants judge a wrongdoer less severely, when a reference point mentioning that s/he has not exploited the situation to its fullest extent is indicated or elicited. Our findings suggest that partial cheaters could emphasize their self-restraint to mitigate judgment and punishment if they get caught. We draw some managerial and policy implications.

Key-words *Ethics, experimental survey, moral judgment, reference points.*

JEL numbers: C91; K42.

The authors declare that they have no conflicts of interest.

Data was collected in accordance with the ethical standards. Informed consent was obtained from all individual participants involved in the study.

A Reference Point Bias in Judging Cheaters

In 2018, an accountant in Belgium was found guilty of embezzling public funds to an amount of €500,000 and argued for his defense that "*[he] could have embezzled much more*" (La, 2018). Interestingly, he benefited from judges' leniency. In the same vein, a former Egyptian Culture Minister, Farouk Hosni who has been acquitted of embezzlement charges argued that "*he could have embezzled much more during his tenure as minister*" (Al-Masry Al-Youm, 2013).

In these two anecdotes, the transgressors do not deny the wrongdoing *per se* but emphasize that they could have embezzled "much more". Interestingly, they seem to have benefited from an excessively lenient judgment, as reported by media. Does the statement that someone has not behaved badly to a maximum extent likely to soften the moral judgment of his wrongdoing? Imagine an employee that has decided to overcharge his travel and accommodation expense reports for an amount of €200, without any detection and punishment risk. Now, as an observer, you are asked to judge his unethical behavior under four different conditions (i) you have no extra information (ii) you know that he cannot overcharge more than €200 without taking the risk of being detected and sanctioned (iii) you know that he could overcharge up to €350 without any risk, but solicited only €200 (iv) you know that he overcharged €200 and you are asked to indicate what you consider the maximum extent the employee could have solicited, without taking the risk of being detected and sanctioned. Noteworthy, the ethical rules that are not respected and the victim (e.g., the employer) are the same, and the lie size or amount overcharged is identical in the four situations. Nevertheless, the reference points provided differ across situations, namely, (i) no reference point, (ii) a reference point or maximum extent is mentioned and reached, (iii) an identical reference point is mentioned but the maximum extent is not reached, and (iv) a reference point is elicited by participants corresponding to what they consider the maximum extent. These reference points correspond to realistic circumstances, where onlookers may be tempted to judge a wrongdoing, not in an absolute fashion, but from a reference point as emphasized by the seminal contribution of Kahneman and Tversky (Tversky and Kahneman, 1981). As an intuitive benchmark, we have a condition without any explicit indication about a reference

point. Two conditions allow to capture whether it is just the explicit presence of the maximum limit that generates a possible attenuating effect or the fact that this limit has not been fully exploited by the wrongdoer. A fourth condition elicits participants reference points by asking them to what could constitute the maximum limit in the considered situation. This additional treatment is very interesting to understand whether observers anticipate that people are partial or full cheaters (Fischbacher and Föllmi-Heusi, 2013) and whether this anticipation influences subsequent judgements. Our main research question can be formulated as follows: *Do these reference points, and especially the situation where the individual is presented as a partial cheater, lead onlookers to express a more lenient judgment?*

This issue is crucial given the considerable amounts at stake. If we consider only expense reimbursement frauds, they account for 21 percent of fraud in small businesses, and 11 percent in large businesses, according to the Association of Certified Fraud Examiners 2018 Report to the Nations (ACFE, 2018). Even if the amounts seem trivial, these moderate losses can add up quickly and reach considerable levels. For instance, business travel expense fraud is estimated to cost U.S. organizations $1.9 billion per year (Cohn, 2018). Understanding the factors that influence the onlookers' judgment on these offenses can help to prevent them. It is well-admitted that individuals attach importance to other people's opinion and judgment upon themselves and their actions. These judgments and opinions and resulting behaviors (e.g., ostracism) can even constitute an important component of the deterrent power of some punishment schemes. Interestingly, if the judgment of others such as colleagues or fellow citizens can be to some extent manipulated by perpetrators to artificially reduce the perceived seriousness of their wrongdoings and subsequent punishment in a predictable fashion, it is crucial to unveil this psychological mechanism. The proposed mechanism is almost free and based on the introduction of an arbitrary reference point where the perpetrator both admits his wrongdoing but emphasizes that s/he has not exploited the situation to its full extent. Given the likelihood of its application to a very broad range of situations with consequential behavioral outcomes, a better understanding of how this *a priori* irrelevant contextual factor shapes the judgment is required.

In most justice systems, fairness and equality are considered as two fundamental guidelines. A natural implication from a purely rational approach is that offenders who commit a similar wrongdoing should receive similar judgments and sentencing outcomes, regardless of irrelevant details. Understanding whether the moral judgments vary according to a suggested reference point is very important. Indeed, manipulating arbitrarily reference points used by onlookers is relatively easy (e.g., Ariely et al., 2003; see also Pittarello et al., 2013). This manipulation can potentially lead to substantial differences in moral judgment and sanctions of objectively equivalent offences. It can ease moral transgressions and lead to subtle influence of observers. Also, this apparent inconsistency in moral judgments can lead people to question the fairness and legitimacy of systems when similar transgressions yield different reactions from observers and different punishments from authorities.

A reference point can indicate a gain or loss framing. Grolleau et al. (2016) found people cheated more when decisions were framed in terms of losses rather than gains[1], but they did not investigate whether the frame also affects the moral judgments of onlookers. As far as we know, the literature is silent on how people judge a similar wrongdoing when they are informed that the transgressor exploits or not the situation to its fullest extent. At the same time, understanding whether onlookers' judgment can be influenced so easily is crucial. To fill this gap, we test whether onlookers judge similarly or differently a similar offense when various reference points are stated. In other words, may people judge more favorably a transgression that violates the same ethical principles and creates the same amounts of harm for the same victims if they have been previously informed that the transgressor has not fully exploited the situation?

The remainder of the paper is organized as follows. In the next section, we develop our conceptual framework by crossing the literature on moral judgment and reference points and draw the

[1] This result has been supported by additional works (e.g., Schindler & Pfattheicher, 2017; Huynh, 2020) but also questioned by other studies (e.g., Heinicke et al., 2019; Soraperra et al., 2019). Heinicke et al. (2019) indicated that a possible explanation between their results and those from previous studies could be stake sizes, that were considerably higher in Grolleau et al. (2016) and Schindler and Pfattheicher (2017). In short, more studies are needed to corroborate these results.

main behavioral hypotheses. The following sections are devoted to two experimental surveys, i.e. i) a preliminary study on a limited sample (N=341) and ii) a main study on a larger online sample (N=1517). For each experimental survey, we follow the same organization; participants and methods, design, results and implications. We also develop the research limitations and suggest avenues for further research. The last section concludes.

Moral judgement and Reference points

From a rationalist perspective, moral judgments are the results of conscious moral reasoning, a process that involves careful, rational thinking and the consistent application of general moral rules or principles. Such reason-based models imply the recognition of moral issues, and then application of moral reasons, rules, or frameworks to reach a moral judgment (Rest, 1986; see also Jones, 1991). Among usual rationalist models, consequentialism focuses on the outcomes by seeking the maximization of well-being whereas formalism focuses on the duty to follow rules and principles. This perspective led scholars such as Jones (1991) to suggest a linear correlation between objective harm and severity of ethical judgment. But these models have difficulty in explaining moral judgments that do not correspond to good moral reasoning as in the case of moral dumbfounding[2].

This rationalist dominance has been challenged by the social intuitionist perspective (Haidt, 2001, 2007) which posits that most moral judgments do not result from a conscious reasoning but rather involves quick, automatic, intuitive, and affective processes. The social intuitionist model emphasizes that most moral judgments result from intuitions that are strongly influenced by contextual factors. Rather than eliminating moral reasoning, the social intuitionist model argues that it occurs *ex post*, that is, after a judgment has been reached and serves to rationalize the *ex ante* intuition (Haidt, 2001, 2007). Haidt argues clearly that "moral reasoning does not cause moral judgment;

[2] Interestingly, Stanley et al. (2019) argue that moral dumbfounding does not provide the requisite evidence for concluding that reasons and reasoning play only a minor role in bringing about moral judgments and decisions. Appealing to the risk of causing harm that did not actually occur offers a reason-based justification of moral wrongness judgments.

rather, moral reasoning is usually a post hoc construction, generated after a judgment has been reached" (p. 814)[3].

These contextual factors can be incidental or irrelevant to the situation being judged, such as the cleanliness of the environments in which individuals form their judgments (Schnall et al., 2008), the time of the day at which moral judgments are made (Gunia et al., 2008), or even the creativity with which the transgression has been performed (Wiltermuth et al., 2017). We consider another kind of contextual factors that can strongly affect the intuitive judgment, precisely, reference points or anchors (Tversky and Kahneman, 1981; Ariely et al., 2003). Reference points, or the reference dependence concept, illustrate how apparently irrelevant changes in the presentation of a decision problem can influence decision makers (Tversky and Kahneman, 1981). For example, Kahneman and Tversky (1981) observed individuals' preferences reversal for a policy program to control a disease just by presenting a scenario that would prevent 200 of 600 deaths as a scenario with 400 lives being lost versus a scenario with 200 lives being saved. While the figures of the problem are exactly the same in both cases, the framing of the problem significantly influenced individuals' preference towards one program. Introducing reference points thus enables to change the frame (loss versus gain) in which individuals make decision with a potential impact on the subsequent decision. Imagine an employee who is expecting a €3,000 annual pay rise, but instead receives €1,000. This is very likely to be experienced as a disappointing news by the employee who will perceive it as a €2,000 loss. However, it would have been a €1,000 gain for an individual who was not expecting a pay rise, which means that his reference point was equal to zero. Moreover, in a gain-loss utility framework, people give more weight to the most common reference points, i.e., the status quo and extreme values, and less weight to information in the middle (Baillon et al., 2020).

Reference dependence was first raised by the literature on decision making under uncertainty, and as such, most of the research is applied to risk preferences (Sprenger and O'Donoghue, 2018). To a lesser but still significant extent, reference-dependent preferences have also been applied to

[3] Moreover, rather than just opposing these two perspectives, recent contributions in business ethics (e.g., Provis, 2017) adopt a more conciliatory view and consider them as complements.

bargaining situations (Kristensen and Gärling, 1997). In the latter, an initial offer will act as a reference point and will tie up future negotiations. Overall, the impact of reference dependence on moral judgment is missing from the literature. The closest research outcomes we could refer to relates to individual unethical behaviors. This literature shows that the maximum amount available that can be stolen acts as a reference point, with people maintaining a positive self-image by cheating less than this artificial reference point. The framing in this case consists in transforming stolen money into saved money (which is the difference between the maximum amount available and the amount that has been taken away). Several scholars argue that some people commit unethical actions and keep a positive self-image, by reassuring themselves that they could have done worse (Mazar et al., 2008; Fischbacher and Föllmi-Heusi, 2013). Many people cheat, but do not cheat to the maximum extent (Fischbacher and Föllmi-Heusi, 2013; Garrett et al., 2016; Abeler et al., 2019). For instance, using an ingenuous cheating dice game, Fischbacher and Föllmi-Heusi (2013) found a high share of cheating subjects to be partial cheaters. Concretely, they did not report the payoff-maximizing draw, but preferred a lower draw, *i.e.*, one that allows them to maintain a positive self-image (Mazar et al., 2008). Garrett et al. (2016) also found that participants could have cheated much more than they actually did. In sum, there is substantive anecdotal and experimental evidence that people cheat, but do not do it to the fullest extent possible (Abeler et al., 2019)[4].

Interestingly, people are usually unaware of the possibility that their views, especially in the valuation domain, might change because of various reference points (Ariely et al., 2003). Ariely et al. (2003) found experimentally that initial valuations of familiar products and simple hedonic experiences are strongly influenced by arbitrary and irrelevant “anchors” such as a person's social security number. Even if the previous authors have not considered ethical judgment, it seems intuitively convincing that this vulnerability to automatic influences could apply to other domains, such as moral judgment. For instance, the same unethical act can be perceived differently if one

[4] This insight echoes research where participants found partial confessions attractive, but failed to anticipate the emotional costs associated with partially confessing, making individuals practicing partial confessions people feeling worse than not confessing or fully confessing (Peer et al., 2014).

compares it with the honesty reference (no wrongdoing at all) or with the maximum extent to which the wrongdoing could have been performed. If the suggested reference point is the lack of any wrongdoing, committing a wrongdoing is likely to be perceived as a loss compared to this ideal situation (see Kern and Chugh, 2009). Kern and Chugh (2009) found that individuals engaged in more unethical behavior when a decision was presented in a loss frame than when the same decision was presented in a gain frame. Conversely, if the suggested reference point is the maximum extent to which the wrongdoing could have been practiced, the same wrongdoing can appear as a gain. Moreover, social judgments critically depend on which comparison standards are made accessible in the judgmental situation and the way reference points are violated (Ockenfels et al., 2015). Interestingly, Shalvi et al. (2011) examined how observing desired counterfactuals impacted unethical decisions. A counterfactual can be perceived as a reference point. Shalvi et al. (2011) suggested that the degree of lying depends on the extent to which self-justifications are available. These previous contributions emphasize that reference points play a pivotal role in how individuals will evaluate, judge and behave in domains related to their own judgments and decisions. Our contribution differs mainly because reference points are used to manipulate the moral judgment of onlookers on the wrongdoer's actions. Moral judgments do not occur in a vacuum and judges can be overly influenced by reference points that are provided by the involved individuals in a self-serving way. More precisely, the same embezzlement can be judged more severely if the action is compared to the no-embezzlement scenario, rather than to the full or even partial embezzlement scenario. In contrast, we argue that it will be judged less severely if the judges or onlookers are influenced by an anchor about the maximum possible embezzlement or about other embezzlers who embezzled much more than the considered embezzler. Simply said, we hypothesize that anchors about the maximum extent (or to a higher extent) to which the wrongdoing could have occurred make onlookers judging the perpetrator less severely, if the indicated threshold has not been reached.

Given that dishonesty is morally wrong and considered as unethical in most societies, if not all, we predict that individuals will make a negative judgment on a dishonest act. Nevertheless, their moral judgments are likely to be reinforced or tempered by a number of contextual factors, such as the

suggested reference points. Based on the preceding discussion, we formulate and test the following hypotheses:

- ✓ *H1: Individuals will judge the same dishonest act as more unethical when no reference point is mentioned compared to a situation where a reference point is mentioned and the maximum extent is not reached.*
- ✓ *H2: Individuals will judge the same dishonest act at the same unethical level when no reference point is mentioned compared to when a reference point is mentioned and the maximum extent is reached, and compared to when the reference point is not mentioned but elicited by the participant himself.*
- ✓ *H3: Individuals will judge the same dishonest act as more unethical when a reference point is mentioned and the maximum extent is reached compared to a situation where the reference point is mentioned but where the maximum extent is not reached.*

Indeed, we contend that not stipulating a reference point can simply suggest individuals to use the honesty reference (*i.e.*, no cheating at all) as a natural and intuitive comparison basis, especially when one considers that most individuals are intrinsically or morally motivated to not engage in dishonest actions (e.g., Gibson et al., 2012; Rosenbaum et al., 2014).

Preliminary study

Participants & methods

As a first investigation, we performed an experimental survey with a simple between-subjects design (Weber, 1992; Croson et al., 2007) involving three identical unethical scenarios with the same end-result. All scenarios were about dishonest actions in the business realm.

Bystanders in the metropolitan area of Montpellier, a medium-sized city in the South of France (more than 470, 000 inhabitants in the metropolitan area in 2017), were randomly solicited to fill a pen-and-paper questionnaire (see Supplementary online materials, SOM1, for the complete experimental design). Respondents participated on a voluntary basis and did not receive a monetary compensation. Each subject participated in only one treatment and was exposed to the three scenarios from this treatment in a fixed order. Participants to the survey were only asked to indicate the ethicality of each wrongdoing on a 7-point Likert scale from 1 (very ethical) to 7 (very unethical). The questionnaires were administered in February 2019 and March 2020, before the pandemics related restrictions. Besides the severity judgment on the three scenarios, we also collected some individual socioeconomic data, namely, gender, age, education level, and financial situation. We gathered a sample of 341 complete observations of which 45% were women. The average age of participants equals 26, 45% of respondents feel financially comfortable and 36% possess a master degree or more.

Design

We used a between design with three scenarios. In order to ensure a high level of realism (Weber, 1992), we selected three plausible scenarios that are all about dishonest actions in the business realm. The scenarios were tested beforehand on twelve individuals (not included in the final sample). By debriefing these individuals, we get qualitative feedback on whether the scenarios were realistic and easy to understand. As a result, some sentences were better phrased (see Supplementary online materials, SOM1, for the complete experimental design). Concretely, we consider (1) an employee who found a way to overcharge his travel expenses to get higher reimbursements from his company (2) an employee who found a way to report more working hours than what s/he actually did and therefore earns a higher salary (3) and a company manager who decided to unlawfully dump 100 liters of a polluting substance into a river near his production plant, in order to make a financial profit. These three scenarios constitute a preliminary robustness check across various domains. Scenarios 1 and 2 have been selected as the dishonest actions are related to personal monetary gains without visible social costs (financial loss for the firm) whereas in scenario 3, the wrongdoing has a high social component (environmental loss). Moreover, in the third scenario, the victim is a natural entity

(e.g. plant, animal, or ecosystem) that is not able to defend itself contrary to a human victim (i.e., higher hierarchical level of management in the two first scenarios) (Martin et al., 2016). Consequently, we may expect that moral judgments will be more severe when the victim has a helpless status (see Dijker, 2010; Schein and Gray, 2018).

More precisely, each unethical act violates the same ethical principles and creates the same amounts of harm for the same victims. In this preliminary study we considered three treatments[5]: without information about the fullest extent possible (T0), with information about the fullest extent possible that is not reached (T1), and with information about the fullest extent possible that is reached (T2).

We tested the three following hypotheses:

- H1: T0 > T1 (Cheating in T0 is judged as more unethical than cheating in T1)
- H2: T0 ≃ T2 (Cheating in T0 is judged as unethical as cheating in T2)
- H3: T2 > T1 (Cheating in T2 is judged as more unethical than cheating in T1)

Results and Discussion

We report average judgments of the three scenarios, as well as the number of observations per treatment in Table 1. Onlookers judge the illegal dumping of a polluting substance into a river to be more unethical compared to over-reporting working hours or over-charging travel expenses in the workplace (Hotelling's T-squared test of equality of means is rejected at more than 99%). This comparison suggests that people are more sensitive to dimensions that affect the public good and 'silent' victims that cannot defend themselves compared to private bads that harm private entities and individuals who can defend themselves.

Comparing the judgment ratings with and without indication regarding the maximum limit (see Table 1) shows that participants do not judge similarly an identical behavior when various reference points

[5] For an illustration of the way treatments were presented to subjects see Box 1, which describes the different treatments for the scenario "Overcharging" of the study 2.

are provided (Kruskal-Wallis test for equality of populations is rejected for all scenarios). For the three scenarios, people are most lenient when a limit is indicated but not reached (T1). The most severe judgement is when no information is given on the maximum limit (T0) for scenarios 1 and 2 only.

Table 1. Average judgement in the 3 treatments (S.D. in brackets)

Treatments		Average judgment rating (S.D)		
	Nb of Obs.	Scenario 1 (Overcharging)	Scenario 2 (Over-reporting)	Scenario 3 (Pollution)
Average judgement for all scenarios	341	4.85 (1.75)	4.95 (1.7)	6.28 (1.4)
T0: No reference point at all	101	5.13 (1.75)	5.35 (1.66)	6.42 (1.31)
T1: Cheated amount < Maximum	144	4.49 (1.75)	4.64 (1.65)	6.04 (1.56)
T2: Cheated amount = Maximum	96	5.09 (1.65)	5 (1.74)	6.5 (1.19)
p-value (Kruskal-Wallis test)		**0.0031****	**0.002****	**0.011***

*p < .05, **p < .01 and ***p < .001

Table 2 reports pairwise comparisons of judgements according to treatments. To account for the multiplicity of null hypotheses tested, we apply the procedure of List et al. (2019) (Theorem 3.1). First, looking at how individuals judge dishonesty when no explicit reference point is given (T0), we find that the average judgment is significantly more severe compared to when the maximum limit is indicated but not reached (T1). This result is robust across the 3 scenarios and provides preliminary empirical support for the hypothesis H1. Second, the moral judgments are quite similar in T0 (no threshold indicated) and T2 (threshold indicated and exhausted) in the three scenarios. In other words, the average judgement with no limit indicated (T0) is not more severe compared to the judgment when the maximum limit is reached (T2). This result provides preliminary support for the hypothesis H2. Third, our results show that in the three scenarios, a similar dishonest act is judged as less unethical when a reference point is given but the maximum extent is not reached (T1) compared to a situation where the same maximum extent is indicated and reached (T2), but this difference is significant in only two of them (S1, S3). This result provides partial empirical support for the hypothesis H3.

Table 2. Multiple hypothesis testing (List et al., 2019) for equality of judgement

Scenario 1 (Overcharging)						
		p-values				
Compared treatments	Difference in means	Unadjusted	Adjusted			
		Remark 3.1	Thm. 3.1	Remark 3.7	Bonferroni	Holm
T0 vs T1	**0.6357**	**0.004****	**0.01*****	**0.01*****	**0.012***	**0.012***
T0 vs T2	0.035	0.875	0.875	0.875	1	0.875
T2 vs T1	**0.6001**	**0.0063****	**0.011***	**0.0063****	**0.019***	**0.013***

Scenario 2 (Over-reporting)						
		p-values				
Compared treatment	Difference in means	Unadjusted	Adjusted			
		Remark 3.1	Thm. 3.1	Remark 3.7	Bonferroni	Holm
T0 vs T1	**0.7076**	**0.0003*****	**0.0003*****	**0.0003*****	**0.001*****	**0.001*****
T0 vs T2	0.035	0.152	0.152	0.152	0.455	0.152
T2 vs T1	0.3611	0.113	0.19	0.113	0.338	0.225

Scenario 3 (Pollution)						
		p-values				
Compared treatment	Difference in means	Unadjusted	Adjusted			
		Remark 3.1	Thm. 3.1	Remark 3.7	Bonferroni	Holm
T0 vs T1	**0.374**	**0.048***	0.087	**0.048***	0.145	0.09
T0 vs T2	0.084	0.639	0.639	0.639	1	0.639
T2 vs T1	**0.458**	**0.014***	**0.038***	**0.038***	**0.042***	**0.042***

*p < .05, **p < .01 and ***p < .001

When the limit is stipulated but not reached, the moral judgment of a similar wrongdoing is less severe compared to a situation where no limit is indicated. In addition, we find partial support that the moral judgement of a similar wrongdoing is less severe when a limit is stipulated but not reached compared to when the limit is indicated and reached. We also found partial empirical support that the moral judgments are quite similar when the limit is indicated and exhausted compared to a situation where no limit is mentioned. These preliminary results suggest that wrongdoers could use self-serving reference points to affect third party evaluators and obtain less severe judgements. Nevertheless, our preliminary experimental survey has some limitations that were addressed in a second study. First, our sample size was somewhat limited and getting results on a larger sample was crucial to reinforce their

reliability. Second, we did not investigate how participants react if they were invited to elicit a reference point corresponding to the maximum extent. This extension can allow a better understanding of how various reference points affect the moral judgment of wrongdoers. Third, the preliminary study lacked some control variables that could be relevant such as the participant's ethical position (Forsyth, 1980) or her/his professional situation.

Main Study

Participants & Methods

To overcome some limitations of our preliminary study, we added a supplementary treatment (rf. T3 in Box 1) and we conducted the survey experiment among a random sample of the French population via the French platform Foule Factory[6]. Alike the U.S. Amazon Mechanical Turk, this platform gives the opportunity for an integrated participant compensation system which complies with the French regulations on minimal wage, a large participant pool and a streamlined process of study design, participant recruitment, and data collection (Buhrmester et al., 2011).

Participants received a fixed payment of 1€ for filling in the questionnaire that lasted on average 5 minutes (corresponding to an average salary of 12€/hour). Each subject participated in only one treatment and was exposed to the three scenarios from this treatment in a random order. Participants to the survey were asked to indicate the ethicality of each wrongdoing on a 7-point Likert scale from 1 (very ethical) to 7 (very unethical). The survey experiment was run in July 2021.

We run a power analysis to determine the minimum sample size required based on data collected during the preliminary study. The analysis revealed that we needed a sample size ranging from 234 to 660 observations according to the heterogeneous effect sizes between our different treatments and among the three scenarios (alpha=0.05; beta=0.80). We collected a sample of 1517 observations. Our sample is composed of 46% of men, 20% of respondents are older than 50 years,

[6] Data available at https://osf.io/xhnwa/

and 32.5% feel financially comfortable. 30% have a master degree or more, and 70% have a professional activity. We constructed dummy variables for all our control variables. We also collected information on each participant's ethical position. We used a shortened version of the Ethical Positional Questionnaire of Forsyth (1980), where participants were asked to indicate on a 9-point Likert scale their level of (dis)agreement with several statements related to either idealism or relativism. We computed two scores for each of these categories of ethical orientation and created a dummy variable to distinguish subjects with a high level of ethical idealism and those with a high level of ethical relativism, using the classical median split. (see Supplementary online materials, SOM3, for more detailed information on the sample; Tables A and B).

Design

As in the preliminary study, we designed an experimental survey with a simple between-subjects design (Weber, 1992; Croson et al., 2007) involving the same three identical unethical scenarios with the same end-result, i.e. (1) an employee who found a way to overcharge his travel expenses to get higher reimbursements from his company (2) an employee who found a way to report more working hours than what s/he actually did and therefore earns a higher salary (3) and a company manager who decided to unlawfully dump 100 liters of a polluting substance into a river near his production plant, in order to make a financial profit (see Supplementary online materials, SOM2, for the complete experimental design). We designed the three scenarios in a way that facilitates comparison, by equalizing the damage of the wrongdoing to 200€/week for all. More precisely, for each kind of unethical act, the same ethical principles are transgressed, and an identical amount of harm is created for the same victims. We consider four treatments (see Box 1): no reference point (T0), a reference point is mentioned but the maximum extent is not reached (T1), with information about the fullest extent possible that is reached (T2), and finally without any information about the limit but each participant is invited to indicate the maximum extent to which one could cheat without attracting attention (T3).

Box 1. Description of the overcharging scenario across treatments

- **T0 (No maximum limit indicated)**: An employee found a way to overcharge his company travel expenses in order to obtain higher reimbursements. He overcharges his travel expenses by € 200 / week without attracting attention.
- **T1 (A maximum limit is indicated but the cheated amount is far below)**: An employee found a way to overcharge his company travel expenses in order to obtain higher reimbursements. He overcharges his travel expenses by € 200 / week, knowing that he could have overcharged up to € 350 without attracting attention.
- **T2 (A maximum limit is mentioned and the cheated amount is equal to it)**: An employee found a way to overcharge his company travel expenses in order to obtain higher reimbursements. He overcharges his travel costs by € 200 / week, knowing that he could not have overcharged more than € 200 without attracting attention.
- **T3 (Participants' maximum limit is elicited**): An employee found a way to overcharge his company travel expenses in order to obtain higher reimbursements. He overcharges his travel expenses by 200 € / week. What do you think is the maximum amount he could overcharge for travel expenses in a week without attracting attention? _____ €

For sake of clarity, we described in Box 1 the variations in the overcharging scenario across the four treatments. Our three hypotheses can be operationalized as follows:

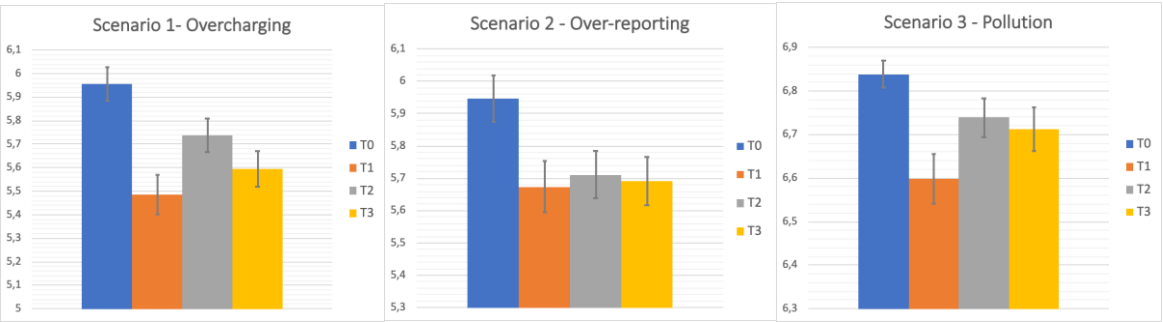
- H1: T0 > T1 (Cheating in T0 is judged as more unethical than cheating in T1)
- H2: T0 ≃ T2≃T3 (Cheating in T0 is judged as unethical as cheating in T2 and in T3)
- H3: T2 > T1 (Cheating in T2 is judged as more unethical than cheating in T1)

Results and Discussion

The 1517 participants were randomly assigned to one of the treatments (T0, T1, T2, or T3). We gathered 360 observations for treatment T0, 371 for treatment T1, and respectively 391 and 395 for

the two control treatments T2 and T3. Average judgments for the three scenarios are reported in Table 3. We observe that, on average, onlookers judge the illegal dumping of a polluting substance into a river to be more unethical compared to over-reporting working hours or over-charging travel expenses in the workplace (Hotelling's T-squared test of equality of means is rejected at more than 99%, t=26.5 and t=27.7 respectively), whereas over-reporting working hours and over-charging travel expenses are judged in a similar way (t=0.98). Moreover, 87.2% of respondents recommend the highest unethical grade, i.e. the 7 level, for illegal dumping of polluting substances, whereas only 42.1% (respectively 43.8) of respondents recommend the highest unethical grade for over-charging travel expenses (respectively over-reporting working hours). These results confirm findings of the preliminary study. People are more sensitive to dimensions that affect the public good and 'silent' victims that cannot defend themselves compared to private bads that harm private entities and individuals who can defend themselves. As stated by Martin et al. (2016), "nature and natural entities cannot speak for themselves" which might explain why people might feel more responsible for defending them and recommend more severe judgments.

Figure 1. ***Average judgment for the 3 scenarios***



Comparing the judgment ratings with and without any indication regarding the maximum limit (see Table 3) shows that participants do not judge similarly an identical behavior when various reference points are provided (Kruskal-Wallis test for equality of populations is rejected for all scenarios). For the 3 scenarios, people are most lenient when a limit is indicated but not reached (T1). The most severe judgment is when no information is given on the maximum limit (T0) for all

scenarios. Interestingly, we observe that judgments are less severe when the maximum limit is elicited (T3) than when no limit is mentioned (T0). However, results show that there is no correlation between the indicated limit and the judgment for all scenarios (ρ =0.046 for scenario 1, ρ =0.031 for scenario 2 and ρ =0.042 for scenario 3).

Table 3. ***Average judgment in the 4 treatments (S.D in brackets)***

Treatments	***Average judgment rating (S.D)***			
	Nb of Obs.	Scenario 1 (Overcharging)	Scenario 2 (Over-reporting)	Scenario 3 (Pollution)
Average judgement for all scenarios	1517	5.69 (1.49)	5.75 (1.47)	6.72 (0.92)
T0: No reference point at all	360	5.96 (1.34)	5.95 (1.36)	6.84 (0.60)
T1: Cheated amount < Maximum	371	5.49 (1.61)	5.68 (1.52)	6.6 (1.11)
T2: Cheated amount = Maximum	391	5.74 (1.42)	5.71 (1.46)	6.74 (0.89)
T3: Cheated amount = Elicited	395	5.59 (1.53)	5.69 (1.50)	6.71 (1.00)
p-value (Kruskal-Wallis test)		**0.0003*** **	**0.03***	**0.003** **

*p < .05, **p < .01 and ***p < .001

Table 4 reports pairwise comparisons of judgments according to treatments. To account for the multiplicity of null hypotheses tested, we apply the procedure of List et al. (2019) which has much greater ability to detect truly false null hypotheses (i.e. reduce type-I errors). We can make two important observations.

First, looking at how individuals judge dishonesty when no explicit reference point is given (T0), we find that the average judgment is significantly more severe compared to when the maximum limit is indicated but not reached (T1) for the three scenarios. This result provides convincing empirical support for hypothesis H1. Second, we did not find consistent support for hypothesis H2. Indeed, except for the scenario 3 (pollution) the moral judgments are lower when a maximum limit is indicated and exhausted (T2) or elicited (T3) than when no maximum level is mentioned (T0). Third, although the numbers go in the predicted directions, we find limited empirical support for hypothesis H3, only for scenario 1 (overcharging), where the equality of judgments is rejected at the 5% level for the un-adjusted p-values. Simply said, cheating when a maximum limit is mentioned and exhausted is not systematically judged as more unethical than cheating when a maximum limit is indicated but the

cheated amount is far below. It suggests that the key lever in attenuating the moral judgment is the mention of a maximum limit, rather than the fact that it has been exhausted.

Table 4. ***Multiple hypothesis testing (List et al., 2019) for equality of judgment ratings for the three scenarios***

Scenario 1 (Overcharging)						
		p-values				
Compared treatments	Difference in means	Unadjusted	Adjusted			
		Remark 3.1	Thm. 3.1	Remark 3.7	Bonferroni	Holm
T0 vs T1	**0.47**	**0.0003*** **	**0.0003*** **	**0.0003*** **	**0.002** **	**0.002** **
T0 vs T2	0.219	**0.027* **	0.087	0.068	0.162	0.108
T0 vs T3	**0.361**	**0.0003*** **	**0.0003*** **	**0.0003*** **	**0.002** **	**0.0017** **
T1 vs T2	0.251	**0.028* **	0.071	0.071	0.168	0.084
T1 vs T3	0.11	0.11	0.350	0.350	1	0.350
T2 vs T3	0.142	0.181	0.307	0.181	1	0.361

*p < .05, **p < .01 and ***p < .001

Scenario 2 (Over-reporting)						
		p-values				
Compared treatments	Difference in means	Unadjusted	Adjusted			
		Remark 3.1	Thm. 3.1	Remark 3.7	Bonferroni	Holm
T0 vs T1	0.273	**0.012* **	**0.05* **	**0.05* **	0.07	0.07
T0 vs T2	0.236	**0.023* **	0.076	0.06	0.136	0.09
T0 vs T3	0.256	**0.013* **	0.055	**0.035* **	0.078	0.065
T1 vs T2	0.037	0.723	0.93	0.93	1	1
T1 vs T3	0.017	0.874	0.874	0.874	1	0.874
T2 vs T3	0.02	0.856	0.975	0.856	1	1

*p < .05, **p < .01 and ***p < .001

Scenario 3 (Pollution)						
		p-values				
Compared treatments	Difference in means	Unadjusted	Adjusted			
		Remark 3.1	Thm. 3.1	Remark 3.7	Bonferroni	Holm
T0 vs T1	**0.241**	**0.0007*** **	**0.002** **	**0.002** **	**0.004** **	**0.004** **
T0 vs T2	0.1	0.627	0.154	0.063	0.376	0.188
T0 vs T3	0.127	0.041	0.148	0.1	0.246	0.205
T1 vs T2	0.141	0.053	0.16	0.127	0.316	0.211
T1 vs T3	0.113	0.148	0.249	0.148	0.888	0.296
T2 vs T3	0.028	0.695	0.695	0.695	1	0.695

*p < .05, **p < .01 and ***p < .001

As a robustness check, we run a multivariate regression model considering simultaneously the three dependent variables, and including respondents ethical position and socio-economic variables as control variables (see Table 5).

Table 5. ***MANOVA multiple-variate regression for judgment ratings (S.D. in brackets)***

Variable	Scenario 1 (Overcharging)	Scenario 2 (Over-reporting)	Scenario 3 (Pollution)
Treatment			
T1 (limit not reached)	**-0.402*** ** (0.104)	**-0.218*** (0.104)	**-0.224*** ** (0.068)
T2 (limit reached)	-0.153 (0.103)	-0.188 (0.103)	-0.094 (0.067)
T3 (limit elicited)	**-0.32** ** (0.102)	**-0.224*** (0.103)	-0.117 (0.067)
Control variables			
Gender	**-0.283*** ** (0.074)	**-0.251*** ** (0.074)	-0.061 (0.048)
Older than 50	**0.481*** ** (0.093)	**0.396*** ** (0.094)	**0.167** ** (0.061)
Financially comfortable	-0.138 (0.081)	-0.07 (0.081)	-0.042 (0.053)
High education level	0.052 (0.081)	0.151 (0.082)	0.042 (0.053)
Professional activity	**0.437*** ** (0.081)	**0.302*** ** (0.081)	**0.284*** ** (0.053)
Ethical Orientation (=1 if higher than the median)			
Idealist	**0.341*** (0.141)	0.126 (0.141)	-0.103 (0.092)
Relativist	**0.383** ** (0.142)	**0.531*** ** (0.142)	**0.179*** (0.092)
Constant	**5.363*** ** (0.124)	**5.345*** ** (0.125)	**6.8*** ** (0.081)
Control for order effect	Yes	Yes	Yes
F	17.31***	12.03***	4.64***
R2	0.12	0.088	0.036
Number of observations	1517	1517	1517
		Multivariate Test, F	
Wilks' lambda		7.12***	
Pillai's trace		6.92***	

*p < .05, **p < .01 and ***p < .001

The model, i.e. the three linear regressions taken together, controls for order effects and is statistically significant for both the Wilks' lambda and Pillai's trace criteria (both p-values are < 0.0001). The regression results are consistent with our previous results. More precisely and in line with hypothesis H1, the moral judgment is more lenient when individuals are told that the wrongdoing level is far below the maximum allowed for all three scenarios. Interestingly, we observe less severe judgments if the maximum limit is elicited for scenarios 1 and 2. Only a few socio-economic variables

influence moral judgments, and these effects are consistent across scenarios. For all scenarios, ethical judgments are increased for older people, for people who have a professional activity, and for people with a high level of ethical relativism. In scenarios 1 and 2, being a male is associated with more lenient judgments.

We also analyzed the determinants of attributing the highest possible unethical grade, that can reflect the fact that some respondents might have felt restrained by the proposed scale. Accordingly, we constructed a binary variable (1 if the ethical judgment equals 7, 0 otherwise) and run probit regressions to investigate what factors contribute to choose this highest grade of unethical judgment and calculate marginal effects of significant variables. The results, not reported here but available in the SOM3, are in line with the previous results.

To summarize, H1 is supported while H2 and H3 are not consistently supported. An integrative explanation for the entire pattern of results is that mentioning reference points related to the possible maximum level to which a wrongdoing can be performed has an attenuating effect on moral judgements. This attenuating effect is the strongest when a maximum limit is indicated and not exhausted. The presence of a reference point seems to prevail over its origin (given or elicited) or level (i.e., whether the maximum extent has been reached or not).

General Discussion

When the limit is stipulated but not reached, the moral judgment of a similar wrongdoing is less severe compared to a situation where no limit is indicated. Regarding hypothesis 2 that was unsupported, a probable explanation of our result (i.e., the moral judgments are lower when a maximum limit is indicated or elicited (T1, T2 and T3) than when no maximum level is mentioned (T0)) might be the dilution effect or the scope bias. Indeed, using more than one source of evidence might result in biased regressive judgments (Smith et al., 1998; Sanborn et al., 2020). In sum, more information is not necessarily better, as an overload of information increases the cognitive burden of searching the best possible outcome and therefore may lead to second-best decisions (Kahneman,

2003). For instance, giving customers more choices reduces their likelihood to make a purchase (Iyengar and Lepper, 2000), or adding green information can harm the perceived quality of products (Grolleau et al., 2019). Similarly, individual judgments could be affected by irrelevant information or evidence and result in a dilution effect, leading individuals to make more regressive decisions or leaner judgments, e.g. the auditors' fraud risk assessments become less extreme if irrelevant evidence is added to the case information (Hackenbrack, 1992; Eilifsen et al., 2019).

Regarding H3, even if results do not reach statistical significance for all scenarios, we found a clear tendency that cheating in T2 (a maximum limit is mentioned and the cheated amount is equal to it) is judged more unethical than cheating in T1 (a maximum limit is indicated but the cheated amount is far below). This partial support could indicate the effect of a limit is likely to be stronger when it is not exhausted, which corresponds well to the anecdotal examples reported in the introduction.

In other words, even if the mention of a reference point or maximum limit (regardless of its origin) has an attenuating effect on moral judgment, this effect is strongest when the maximum limit has not been exhausted. Interestingly, this attenuating mechanism is likely to apply to a wide variety of unethical acts in the business realm. Our findings extend the theory of "partial cheaters" by suggesting that wrongdoers could also decide to cheat a little as a strategy to mitigate punishment if they get caught.

The main implication of our results is that perpetrators of wrongdoings can be tempted to use self-serving reference points to affect third party evaluators and obtain less severe judgments. If offenders and their defenders anticipate the effect of reference points, they can attempt to manipulate them, even arbitrarily, to diminish the moral severity of judgments of outside observers upon their acts. Interestingly, this tactic can be used in a wide variety of contexts such as stealing, lying, cheating, embezzling or "cooking the books".

In addition to this external motivation and even if our protocol design does not allow us to test it, we would like to mention an additional and internal motivation that can push individuals to use self-serving reference points. Indeed, while our findings show that many people cheat "a little" or less

than the maximum amount to mitigate punishment if they get caught, we speculate that they can also choose this strategy as a way to maintain a positive self-image. For instance, Hilbig and Hessler (2013; see also Shalvi et al., 2011) showed experimentally that dishonesty threatens one's self view and as a consequence, people avoid major lies. In their meta-analysis on dishonest behavior, Gerlach et al. (2019, p. 20) proposes to devote more attention to the desire to appear honest by distinguishing whether participants want to appear honest to themselves (producing guilt when not following some internalized norms) or to others (producing shame when not following some social norms). Of course, we do not claim that influencing third party evaluations solicit the same psychological mechanisms and reasons than those underpinning the willingness to preserve a positive self-image while lying. A promising extension of our contribution will be to test experimentally whether wrongdoers consciously emphasize to themselves that they did not exploit the situation to the maximum extent to maintain their positive self-image and maybe to examine the respective contribution of each mechanism.

Given the power of this tactic and its detrimental and contagious consequences, it seems important to suggest some ways to de-bias individuals who are amenable to this bias (Lilienfeld et al., 2009). A natural de-biasing candidate is to inform outside observers or judges at the adequate timing on the reference point bias. An extension could be to examine whether people are less subject to this reference bias if they are informed and cautioned just before. Another strategy can be to use simultaneously counter reference points such as the honesty norm, by creating a kind of rivalry among reference points.

A collateral implication suggested by our findings is that some interested individuals may attempt to manipulate the domain to which the wrongdoing belongs (e.g., public versus private domains, helpless versus powerful victim) in order to influence the moral judgment of onlookers. For instance, stating a company theft can be requalified as a theft affecting an employee fund or even as a theft that will prevent environmental investments can be intentional in order to affect the judgment of observers. In short, the way a similar wrongdoing is qualified may constitute a powerful influence

weapon, but further research is required to elaborate and test whether and how the domain qualification affects the moral judgment of observers.

Limitations and future research

Our two survey experiments have several limitations. We do not measure a moral judgment in a real-world situation. An appropriate lab or field experiment where judges are incentivized to punish wrongdoers can significantly enrich the analysis. Rather than providing definitive replies, we consider our contribution as a stepping stone that paves the way to further research on this fascinating topic.

The way the reference point is created, framed and how far any departure from this reference point impacts judgment constitute insightful extensions. For instance, even if we found supporting evidence that third party evaluators attenuate their moral judgment when they are reminded that the wrongdoer has not fully exploited the situation, we did not examine whether and how the discrepancy size between the achieved extent and the maximum extent matters. It remains unclear whether a small difference is enough to activate the above-mentioned effect or whether a minimum size is necessary. Computing this "elasticity" could usefully inform managers and policymakers on where to direct their attention and avoid a one-size-fits all approach. Moreover, does this tactic remain effective if used repetitively, e.g., in case of multiple wrongdoings, and for the same individual?

Last but not least, we observed a different moral judgment level when the wrongdoing was related to human versus natural entities. Further investigations are needed to better understand the rationale behind this judgment discrepancy. Are judgments on unethicality subject to the nature of the victim (e.g., public versus private, human versus non-human entity)? In the case of a public good, people may perceive themselves as inside observers (potentially bearing the impact of the wrongdoing) while in a private good scenario, people may perceive themselves as outside observers (no direct relationship between the wrongdoing and themselves).

Conclusion

Thanks to two survey experiments, we investigated the effect of mentioning a maximum limit on moral judgments of wrongdoings. The main finding of our study is that reference points related to the maximum level to which a wrongdoing could be performed matter in moral judgement. The moral judgments of wrongdoings are lower when a maximum limit is indicated or elicited compared to situations where no maximum level is mentioned. Participants judge a wrongdoer less severely, when it is indicated that s/he has not exploited the situation to its fullest extent. Our findings suggest that partial cheaters can emphasize their self-restraint to mitigate judgment and punishment if they get caught.

References

Abeler, J., Nosenzo, D., and Raymond, C. (2019). Preferences for truth-telling. *Econometrica* 87(4), *1115-1153*.

Al-Masry Al-Youm (2013). Former culture minister acquitted of embezzlement charges, *Egypt Independent*, January, 5, https://www.egyptindependent.com/former-culture-minister-acquitted-embezzlement-charges/

Ariely, D., Loewenstein, G., Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *Quarterly Journal of Economics*, 118(1), *73-105*.

Association of Certified Fraud Examiners (ACFE). (2018). *Report to the Nations*, https://www.acfe.com

Baillon, A., Bleichrodt, H., Spinu, V. (2020). Searching for the Reference Point. *Management Science*, 66(1), *93-112*.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3–5.

Cohn, M. (2018). Expense fraud estimated to cost $1.9B annually, *Accounting Today*, May, 31, https://www.accountingtoday.com/news/business-travel-expense-fraud-estimated-to-cost-19b-annually

Cook, N.J., Grillos, T., Andersson, K.P. (2019). Gender quotas increase the equality and effectiveness of climate policy interventions. *Nature Climate Change*, 9, *330-334*.

Croson, R., Anand, J., Agarwal, R. (2007). Using experiments in corporate strategy research. *European Management Review*, 4, *173–181*.

Dijker, A.J.M. (2010), Perceived vulnerability as a common basis of moral emotions. *British Journal of Social Psychology*, 49, 415-423.

Eilifsen, A., Kochetova, N., Messier, W. (2019). Mitigating the Dilution Effect in Auditors' Judgments Using a Frequency Response Mode. *Behavioral Research in Accounting*, 31 (2), 51–71.

Fischbacher, U., Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11, *525-547*.

Forsyth, D. R. (1980). A taxonomy of ethical ideologies. *Journal of Personality and Social Psychology*, 39(1), *175–184*.

Garrett, N., Lazzaro, S.C., Ariely, D., Sharot, T. (2016). The brain adapts to dishonesty. *Nature Neuroscience*, 19, *1–6*.

Gerlach, P., Teodorescu, K., Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, 145(1), *1-44.*

Gibson, R, Tanner, C., Wagner, A. (2012). Preferences for truthfulness: heterogeneity among and within individuals. *American Economic Review*, 103(1), *532–548*.

Grolleau, G., Kocher, M. G., Sutan, A. (2016). Cheating and loss aversion: Do people cheat more to avoid a loss? *Management Science*, 62(12), *3428-3438.*

Grolleau, G, Mzoughi, N, Sutan, A. (2019). Does advertising the green benefits of products contribute to sustainable development goals? A quasi-experimental test of the dilution effect. *Business Strategy and the Environment*, 28, *786– 793*.

Gunia, B.C., Barnes, C.M., Sah, S. (2014). The morality of larks and owls unethical behavior depends on chronotype as well as time of day, *Psychological Science*, 25 (12), *2272-2274*.

Hackenbrack, K. (1992). Implications of seemingly irrelevant audit evidence in judgment. *Journal of Accounting Research*, 30(1), 126-36.

Haidt, J. (2001). The emotional dog and its rational tail. *Psychological Review*, 108, *814-834*.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), *998–1002*.

Heinicke, F., Rosenkranz, S. Weitzel, U. (2019). The Effect of Pledges on the Distribution of Lying Behavior: an Online Experiment. *Journal of Economic Psychology*, 73, *136-151*.

Hilbig, B. E., Hessler, C. M. (2013). What lies beneath: How the distance between truth and lie drives dishonesty. *Journal of Experimental Social Psychology*, 49(2), 263-266.

Huynh, T.L.D. (2020). Replication: Cheating, loss aversion, and moral attitudes in Vietnam. *Journal of Economic Psychology*, 78, 102277.

Iyengar, S., Lepper, M. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), 995–1006.

Jones, T. (1991). Ethical Decision Making by Individuals in Organizations: An Issue-Contingent Model. *Academy of Management Review*, 16(2), 231–248.

Kahneman, D. A. (2003). perspective on judgment and choice: mapping bounded rationality. *American Psychology*, 58, 697–720

Kern, M.C., Chugh, D. (2009). Bounded ethicality: The perils of loss framing. *Psychological Science*, 20(3), *378-384*.

Kristensen, H., Gärling, T. (1997). The effects of anchor points and reference points on negotiation process and outcome. *Organizational Behavior and Human Decision Processes*, 71(1), *85–94*.

La., J. (2018). Le comptable a volé 500 000 euros à la Communauté française : "J'aurais pu détourner beaucoup plus", *La Libre Belgique*, Avril, 17, https://www.lalibre.be/belgique/le-comptable-a-vole-500-000-euros-a-la-communaute-francaise-j-aurais-pu-detourner-beaucoup-plus-5ad62ea3cd702f0c1aef0251.

Lilienfeld, S.O., Ammirati, R., Landfield, K. (2009). Giving Debiasing Away Can Psychological Research on Correcting Cognitive Errors Promote Human Welfare? *Perspectives on Psychological Science*, 4(4), *390-398*.

List, J.A., Shaikh, A.M., Xu, Y. (2019). Multiple hypotheses testing in experimental economics. *Experimental Economics*, 22, *773–793*.

Martin, J-L, Maris, V., Simberloff, D. (2016) The need to respect nature and its limits challenges society and conservation science. *Proceedings of the National Academy of Sciences*, 113(22), *6105-6112*.

Mazar, N., Amir, O., Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45, *633-644*.

Ockenfels, A., Sliwka, D., Werner, P. (2015). Bonus Payments and Reference Point Violations, *Management Science*, 61(7), *1496-1513*.

O'Donoghue, T., Sprenger, C. (2018). Chapter 1 - Reference-Dependent Preferences, in *Handbook of Behavioral Economics: Applications and Foundations* 1, North-Holland, Vol. 1, *1-77*.

Peer, E., Acquisti, A., & Shalvi, S. (2014). "I cheated, but only a little": Partial confessions to unethical behavior. *Journal of Personality and Social Psychology*, 106(2), *202-217*.

Pittarello, A., Rubaltelli, E. & Rumiati, R. (2013). You can't be better than me: The role of the reference point in modulating people's pursuit of wealth, Journal of Economic Psychology, 37, *65-76*.

Provis, C. (2017). Intuition, analysis and reflection in business ethics. *Journal of Business Ethics*, 140(1), *5-15*.

Rest, J. (1986). *Moral development: Advances in research and theory*. New York: Praeger.

Rosenbaum, S.M, Billinger, S., Stieglitz, N. (2014). Let's be honest: a review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, 45, *181–196*.

Sanborn, A., Noguchi, T., Tripp, J., Stewart, N. (2020). A dilution effect without dilution: When missing evidence, not non-diagnostic evidence, is judged inaccurately. *Cognitio*n, 196, https://doi.org/10.1016/j.cognition.2019.104110.

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, 22(1), *32–70.*

Schindler, S., & Pfattheicher, S. (2017). The frame of the game: Loss-framing increases dishonest behavior. *Journal of Experimental Social Psychology*, 69, *172-177*.

Schnall, S., Haidt, J., Clore, G., Jordan A. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34 (8), *1096-1109*.

Shalvi, S., Dana, J., Handgraaf, M. J. J., & De Dreu, C. K. W. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, 115(2), *181-190*.

Smith, H.D., Stasson, M.F. & Hawkes, W.G. (1998). Dilution in legal decision making: Effect of non-diagnostic information in relation to amount of diagnostic evidence. *Current Psychology*, 17, *333–345* https://doi.org/10.1007/s12144-998-1015-6

Soraperra, I., Weisel, O., & Ploner, M. (2019). Is the victim max (Planck) or Moritz? How victim type and social value orientation affect dishonest behavior. *Journal of Behavioral Decision Making*, 32(2), *168–178*.

Sprenger, C. & O'Donoghue, T. (2018). Reference-Dependent Preferences. In Handbook of Behavioral Economics: Applications and Foundations. Volume 1, 1-77.

Stanley, ML., Yin, S. Sinnott-Armstrong, W. (2019). A reason - based explanation for moral dumbfounding. *Judgment and Decision Making*, 14 (2), 120–129.

Tversky, A., Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, *453–458*.

Weber, J. (1992). Scenarios in Business Ethics Research: Review, Critical Assessment, and Recommendations, *Business Ethics Quarterly*, 2(2), *137–159*.

Wiltermuth, S.S., Vincent, L.C., Gino, F. (2017). Creativity in unethical behavior attenuates condemnation and breeds social contagion when transgressions seem to create little harm. *Organizational Behavior and Human Decision Processes*, 139, *106-126*.