University of Reading

The Reference Point Bias in Judging Cheaters

by Sophie Clot, Gilles Grolleau and Lisette Ibanez

The 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)

CEREN EA 7477, Burgundy School of Business, Université Bourgogne Franche-Comté, Dijon, France

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 the situation to its maximum extent? Using a social intuitionist perspective and taking into account the reference point bias, we hypothesize that people will judge less severely a moral wrongdoing when the situation is not exploited to its fullest extent. Thanks to an experimental survey in France, we examine whether various wrongdoings performed in the business realm (overcharging travel expenses, overstating work hours, pollution) are judged less severely when differing reference points are suggested: (i) no explicit reference point is mentioned, (ii) the maximum extent is reached, (iii) the maximum extent is not reached. Our findings support that participants judge less severely a wrongdoer, when it is indicated that s/he has not exploited the situation to its fullest extent. In addition of maintaining their self-concepts, our findings suggest that partial cheaters can also emphasize their self-restraint to mitigate judgement 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 Reference Point Bias in Judging Cheaters

1. Introduction

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. Does the statement that someone has not behaved badly to the 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, you have to judge his unethical behavior under three 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. 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 three situations. Nevertheless, the reference points provided differ across situations, namely, (i) no reference point, (ii) a maximum extent that has been exhausted, and (iii) a maximum extent that has not been reached. *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' judgement on these offenses can help to prevent them, especially if these factors are predictable and can be manipulated in a self-serving way.

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). 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 manipulation of observers. Also, this apparent inconsistency in moral judgements 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. Inducing a gain frame rather than a loss one regarding a wrongdoing can lead observers to express more lenient moral judgement (Grolleau et al., 2016). 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' judgement 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 integrating the literature on moral judgement and reference points and draw the main behavioral hypotheses. Section 3 is devoted to the empirical strategy. Section 4 conveys the main results

and discusses some managerial and policy implications. Section 5 concludes and suggests avenues for further research.

2. Conceptual framework and main hypotheses

From a rationalist perspective, moral judgements 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 judgement (Rest, 1986). 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. But these models have trouble to explain moral judgements that do not correspond to good moral reasoning as in the case of moral dumbfounding[1].

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; rather, moral reasoning is usually a post hoc construction, generated after a judgment has been reached" (p. 814)[2].

---

[1] 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.

[2] 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.

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 presentation or framing of the problem significantly influenced individuals' preference towards one program. Introducing reference point 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, much 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 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 judgement 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-concept 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 argued that some people commit unethical actions and keep a positive self-concept, 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 fullest extent possible (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-concept (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; La, 2018).

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). For instance, the same unethical act can be perceived differently if one 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). 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). 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 nudged with 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 a similar dishonest act as more unethical when no explicit reference point is given compared to a situation where the maximum extent is indicated but not reached.*

✓ *H2: Individuals will judge a similar dishonest act at the same unethical level when no limit is indicated and when the maximum extent is indicated and reached.*

✓ *H3: Individuals will judge a similar dishonest act as less unethical when a reference point is given but the maximum extent is not reached compared to a situation where the maximum extent is indicated and 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 not to engage in dishonest actions (e.g., Gibson et al., 2012; Rosenbaum et al., 2014).

3. Empirical strategy

In order to test our three hypotheses, we designed an experimental survey with a simple between-subjects design (Weber, 1992; Croson et al., 2007) involving several identical dishonesty or unethical scenarios or vignettes with the same end-result. More precisely, each unethical act violates the same

ethical principles and creates the same amounts of harm for the same victims. There are three treatments (Table 1): 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).

Table 1. Experimental design

T0	T1	T2
No maximum limit indicated	A maximum limit is indicated but the cheated amount is far below	A maximum limit is indicated and the cheated amount is equal to it

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 (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)

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 scenario was pretested on a small convenience sample (N=10) not included in the final sample to improve its understanding and readability. 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 hours of work than 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 (higher hierarchical level of management in the two first scenarios) (Martin et al., 2016). Consequently, we may expect that moral judgements will be more severe when the victim has a helpless status. The scenarios were tested beforehand on twelve individuals (data not included in the final sample), notably to improve realism, understanding and reader friendliness (see Appendix 1). Bystanders in the metropolitan area of Montpellier, a medium-sized city in the South of France, were randomly solicited to fill a pen-and-paper questionnaire. Respondents participated on a voluntary basis and did not receive a monetary compensation. Each subject participated in only one treatment. The questionnaires were administered in February 2019 and March 2020. Besides the severity judgment on the three scenarios, we also collected some individual socioeconomic data, namely, gender, age, education level, and financial situation.

4. Results and discussion

We gathered a sample of 341 useable observations of which 45% were women. Some descriptive statistics regarding the sample are provided in Table 2. We also report average judgments of the three scenarios. Participants to the survey were asked to indicate the ethicality of the wrongdoing on a 7-point Likert scale from 1 (very ethical) to 7 (very unethical). 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%). 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. Control variables are all dummy variables.

Table 2. Descriptive statistics regarding the sample

Variables	N	Mean	S.D.	Min	Max
Control variables					
Gender (men=1)	341	0.55	0.5	0	1
Older than 25	341	0.35	0.48	0	1
Financially comfortable	341	0.44	0.5	0	1
High education level	341	0.36	0.36	0	1
Year (=2020)	341	0.43	0.5	0	1
Dependent variables					
Judgement ratings					
- Scenario 1	341	4.85	1.75	1	7
- Scenario 2	341	4.95	1.7	1	7
- Scenario 3	341	6.28	1.4	1	7

Comparing the judgment ratings with and without 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 three scenarios, people are most lenient when a limit is indicated but not reached (T1). Conversely, the most severe judgement is when no information is given on the maximum limit (T0), except for scenario 3.

Table 3. Average judgement in the 3 treatments

Treatments	Average judgment rating (S.D)			
	Nb of Obs.	Scenario 1	Scenario 2	Scenario 3
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**

(*), (**), (***) refer to the significance of the Kruskal-Wallis test at the 10%, 5% and 1% levels respectively.

Table 4 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) in order to control for not rejecting true null hypotheses. 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 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 convincing empirical support for the hypothesis H2. Third, our results show that in two scenarios out of three (S1, S3) 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). This result provides partial empirical support for the hypothesis H3.

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

		Scenario 1				
		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				
		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				
		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****

(*), (**) and (***) stand for significance at the 10, 5 and 1% level, respectively.

As a robustness check, we run an ordered probit model, including socio-economic variables as control variables (see Table 5). The regression results are consistent with our previous results. More precisely, the moral judgment is significantly more lenient when individuals are told that the wrongdoing level is far below the maximum allowed. Only a few socio-economic variables influence moral judgments, but these effects are inconsistent across scenarios. In scenario 2, being older than 25 years is associated with more severe judgments. In scenario 3, being a male is associated with more lenient judgements. This latter effect is consistent with the well-known higher environmental awareness of women (Cook et al., 2019).

Table 5. Ordered probit regression for judgement ratings

Variable	Scenario 1	Scenario 2	Scenario 3
Treatment			
T1 (limit not reached)	**-0.402***	**-0.6***	**-0.486***
T2 (limit reached)	-0.016	-0.232	0.091
Control variables			
Gender	0.1	0.007	**-0.265***
Older than 25	0.355	**0.264****	0.242
Financially comfortable	0.065	-0.119	-0.163
High education level	-0.112	-0.131	-0.236
Year (= 2020)	-0.12	**-0.285****	0.024
Log Likelihood	-619.003	-601.93	-351.48
LR Chi2	15.67***	28.12***	17.96***
Pseudo R2	0.0125	0.0228	0.0249
Number of observations	341	341	341

(*), (**) and (***) stand for significance at the 10, 5 and 1% level, respectively.

To go further in the analysis, we also analyzed the determinants of attributing the highest possible unethical grade. Indeed, recommending the highest unethical grade can reflect the fact that some respondents might have felt restrained by the proposed scale, by being unable to select an even stronger unethicality level. Accordingly, we use a binary variable (1 if the ethical judgment equals 7, 0 otherwise)

and run a probit regression to investigate what factors contribute to choose this highest grade of unethical judgment and calculate marginal effects (dy/dx) of significant variables (Table 6).

Table 6. Probit regression for the highest unethical grade

Variables	Scenario 1		Scenario 2		Scenario 3	
Treatment	Coef.	dy/dx	Coef.	dy/dx	Coef.	dy/dx
T1 (limit not reached)	**-0.45****	**-0.122****	**-0.842*****	**-0.225*****	**-0.526*****	**-0.18*****
T2 (limit reached)	-0.1		-0.252		0.076	
Control variables						
Gender	0.121		-0.115		**-0.341****	**-0.11****
Older than 25	**0.617*****	**0.168*****	**0.602*****	**0.161*****	**0.335***	**0.11***
Financially comfortable	-0.003		**-0.29***	**-0.78***	-0.15	
High education level	-0.123		-0.07		-0.214	
Year (= 2020)	-0.204		-0.305		0.01	
Constant	**-0.78*****		-0.291		**0.965*****	
Log Likelihood	-164.99		-162.14		-197.81	
LR Chi2	24.19***		37.53***		20.49***	
Pseudo R2	0.0683		0.1037		0.0046	
Number of obs.	341		341		341	

(*), (**) and (***) stand for significance at the 10, 5 and 1% level, respectively.

We find results similar to those obtained in the ordered probit regression. First, informing respondents that the maximum limit is not reached reduces the probability to recommend the highest unethical grade for nearly 12% in scenario 1, 22.5% in scenario 2 and 18% in scenario 3. Second, older people opt more for the highest unethical grade for around 16% in scenarios 1 and 2 and 11% in scenario 3. In addition, the constant term is positive in scenario 3 but negative in scenarios 1 and 2. One of the plausible explanations can be related to the fact that polluting a river (public bad) is perceived as far more unethical than overcharging travel expenses and over-reporting working hours (private bads). Indeed, globally 70% of respondents find it totally unacceptable to dump illegally polluting substances into a river against 21% in scenario 1 and 22% in scenario 2. 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 their causes and recommend more severe judgments.

The main implication of our results is that perpetrators of wrongdoings can be tempted to use self-serving reference points to maintain a positive self-concept and obtain less severe judgements. Our findings extend the theory that many people cheat "a little" in order to maintain moral self-concepts by suggesting *they also do so to mitigate punishment if they get caught*. Indeed, perpetrators can keep a positive self-concept by emphasizing to themselves that they did not exploit the situation to the maximum extent. 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".

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 judgement 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 judgement 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.

5. Conclusion

Thanks to a survey experiment, we investigated the effect of mentioning a maximum limit on moral judgements. We observed that 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. We also found partial empirical support that the moral judgments are quite similar when the limit is indicated and exhausted compared to a limit-free situation. In other words, it is not the provision *per se* of the reference point that attenuates the moral judgement, but its mention *and* the fact that it has not been reached that are effective in attenuating moral judgement. Interestingly, this attenuating mechanism is likely to apply to a wide variety of unethical acts in the business realm. In addition of maintaining moral self-concepts, our findings extend the theory of "partial cheaters" by suggesting that they also cheat a little as a strategy to mitigate punishment if they get caught.

The way the reference point is created, framed and how far any departure from this reference point impacts judgement constitute insightful extensions. For instance, does the simple fact that the wrongdoer does not fully exploit the situation is enough to attenuate moral judgement? Or does the discrepancy between the achieved extent and the maximum extent matter? Moreover, once a reference point is created, it seems crucial to explore whether it has a long-lasting effect on the moral judgment of people's judgement.

Last but not least, we observed a different moral judgement level when the wrongdoing was related to human versus natural entities. Further investigations are needed to better understand the rationale behind the judgment discrepancy. Are judgements 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).

Our survey experiment has several limitations. We do not measure a moral judgement in a real-world situation. An appropriate lab or field experiment where judges are incentivized to punish wrongdoers

can significantly enrich the analysis. Moreover, our data does not allow us to examine how the gap size between the full extent and the observed wrongdoing level influences the mitigating effect. Computing this "elasticity" could usefully inform managers and policymakers on where to direct their attention and avoid a one-size-fits all approach. Rather than providing definitive replies, we consider our contribution as a stepping stone that paves the way to further research on this fascinating topic.

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.

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.

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

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

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

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.

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.

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), 1496-151.

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.

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, SM, 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.

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

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.

Appendix 1: Anonymous survey (Three versions: T0, T1, T2) - translated from French (For refereeing purposes only)

We are interested solely in your opinion. There are no good or bad answers. Thank you for responding as sincerely as possible

In the following hypothetical situations, suppose there is no economic or social sanction against the individual (or entity) concerned. For each scenario indicate, on a scale of 1 (very ethical) to 7 (very unethical), whether the behavior described above seems more or less ethically acceptable.

<u>**Scenario 1**</u>

Treatment T0: An employee found a way to overcharge his travel expenses to his business for higher refunds. He overcharges his travel expenses by 200 € / month.

Treatment T1: An employee found a way to overcharge his travel expenses to his business for higher refunds. He overcharges his travel expenses by 200 € / month, knowing that he could have safely overcharged up to 350 €.

Treatment T2: An employee found a way to overcharge his travel expenses to his business for higher refunds. He overcharges his travel expenses by 200 € / month, knowing that he could not have overcharged more than 200 €.

<u>**Scenario 2**</u>

Treatment T0: An employee has found a way to report more hours of work than he or she actually did and therefore earns a higher salary. He over-declares 10 hours / month.

Treatment T1: An employee has found a way to report more hours of work than he or she actually did and therefore earns a higher salary. He over-declares 10 hours / month, knowing that he could have safely over-declared until 15h

Treatment T2: An employee has found a way to report more hours of work than he or she actually did and therefore earns a higher salary. He over-declares 10 hours / month, knowing that he could not have over-declared more than 10 hours.

Scenario 3

Treatment T0: In order to make a financial profit, a company manager decides to unlawfully dump 100 liters of a polluting substance into a river near his production plant.

Treatment T1: In order to make a financial profit, a company manager decides to unlawfully dump 100 liters of a polluting substance into a river near his production plant, knowing that he could safely have poured up to 200 l of this substance.

Treatment T2: In order to make a financial profit, a company manager decides to unlawfully dump 100 liters of a polluting substance into a river near his production plant, knowing that he could not have poured more than 100 liters of this substance.