



STUDY PROTOCOL

**REVISED** Stage 1 Registered Report: How subtle linguistic cues prevent unethical behaviors [version 4; peer review: 2 approved]

Wen Guo <sup>1</sup>, Huanxu Liu <sup>1</sup>, Jingwen Yang <sup>1</sup>, Yuqi Mo<sup>1</sup>, Can Zhong <sup>1</sup>, Yuki Yamada <sup>2</sup>

<sup>1</sup>Graduate School of Human-Environment Studies, Kyushu University, Fukuoka, Japan  
<sup>2</sup>Faculty of Arts and Science, Kyushu University, Fukuoka, Japan

**v4** First published: 22 Aug 2019, 8:1482 (<https://doi.org/10.12688/f1000research.20183.1>)  
 Second version: 24 Dec 2019, 8:1482 (<https://doi.org/10.12688/f1000research.20183.2>)  
 Third version: 02 Mar 2020, 8:1482 (<https://doi.org/10.12688/f1000research.20183.3>)  
 Latest published: 12 Mar 2020, 8:1482 (<https://doi.org/10.12688/f1000research.20183.4>)

**Abstract**

Different ways of description can easily influence people’s evaluations and behaviors. A previous study by Bryan and colleagues suggested that subtle linguistic differences in ethical reminders can differentially prevent readers’ unethical behavior. The present study aims to replicate the previous finding in the Japanese context, additionally exploring the influence of unfamiliar instruction words that capture participants’ attention. In two experiments, which are planned to be conducted online, participants are asked to make 10 coin-tosses and report the number of “heads” results, indicating the amount of money that they can earn. We will manipulate instructions (“Don’t cheat” vs. “Don’t be a cheater” vs. baseline as a control) for each participant group, including nearly 270 participants (Experiment 1). Next, we will conduct an extended experiment with an additional task in which more attention is directed toward the text (Experiment 2). Through these registered experiments, we examine the credibility of the previous finding that type of instruction affects the occurrence of unethical behaviors.

**Keywords**

cheating, persuasion, attention, moral, self construal, labeling



This article is included in the **Preclinical Reproducibility and Robustness** gateway.

**Open Peer Review**

Reviewer Status

	Invited Reviewers	
	1	2
<b>version 4</b> (revision) 12 Mar 2020		
<b>version 3</b> (revision) 02 Mar 2020		report
<b>version 2</b> (revision) 24 Dec 2019	report	report
<b>version 1</b> 22 Aug 2019	report	

- 1 Michiko Asano**, Rikkyo University, Saitama, Japan
- 2 Sergio Cervera-Torres** <sup>1</sup>, Leibniz-Institute für Wissensmedien (IWM), Tübingen, Germany

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding author:** Yuki Yamada ([yamadayuk@gmail.com](mailto:yamadayuk@gmail.com))

**Author roles:** **Guo W:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Liu H:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Yang J:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Mo Y:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Zhong C:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Yamada Y:** Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Original Draft Preparation

**Competing interests:** No competing interests were disclosed.

**Grant information:** This research was supported by JSPS KAKENHI (15H05709, 16H01866, 17H00875, 18H04199, and 18K12015).  
*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2020 Guo W *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Guo W, Liu H, Yang J *et al.* **Stage 1 Registered Report: How subtle linguistic cues prevent unethical behaviors [version 4; peer review: 2 approved]** F1000Research 2020, 8:1482 (<https://doi.org/10.12688/f1000research.20183.4>)

**First published:** 22 Aug 2019, 8:1482 (<https://doi.org/10.12688/f1000research.20183.1>)

**REVISED Amendments from Version 3**

Based on Reviewer 2's Comments, some minor changes were made to improve readability and accuracy.

**Any further responses from the reviewers can be found at the end of the article**

## Introduction

When people behave dishonestly, they usually downplay the seriousness of the dishonest act (e.g., [Monin & Jordan, 2009](#); [Steele, 1988](#)), weakening the link between the dishonesty and one's self-identity (e.g., [Bandura, 1999](#)) to avoid the correspondent inference ([Jones & Nisbett, 1972](#); [Ross, 1977](#)) that one is the kind of person who behaves dishonestly. According to self-concept maintenance theory, individuals in general strive to create and maintain an image of themselves as good and ethical people ([Markus & Wurf, 1987](#); [Mazar et al., 2008](#)).

In general, we believe that highlighting a self-identity word will prevent unethical behaviors to some degree. According to [Blasi \(1984\)](#), a moral person is one for whom moral categories and moral notions are central, essential, and important to self-understanding. Morals cut deeply to the core of what and who such people are as individuals. However, one study revealed that highly constructed self-identities are associated with more unethical behaviors ([Cojuharencu et al., 2012](#)).

Regarding ethical behavior, a moral-character model has been proposed, where moral character consists of motivation, ability, and identity elements ([Cohen & Morse, 2014](#)). Moral identity here refers to being disposed toward valuing morality and wanting to view oneself as a moral person. This disposition should be considered when attempting to understand why people who behave unethically tend to apply a variety of strategies to weaken the behavior–identity link ([Bandura, 1999](#)). The use of “euphemistic labeling” to describe one's attributes and weaken the link regarding language should also be included in this disposition.

Different ways of description can easily influence people's evaluation and judgment about something, even if they have a wealth of previously established knowledge ([Fausey & Boroditsky, 2010](#)). For instance, using a transitive verb (agentive description, e.g., “Timberlake ripped the costume”) to describe an accident makes participants significantly more likely to blame the actor compared to the same description with the words changed to an intransitive verb (nonagentive description, e.g., “The costume ripped”). Another study found that, for children aged 5–7 years old, when a noun label was employed to describe a character (e.g., “She is a carrot-eater”) rather than a verbal predicate (e.g., “She eats carrots whenever she can”), their judgment about those characteristics would be more stable over time ([Gelman & Heyman, 1999](#)). The same phenomenon has been demonstrated regarding self-perception ([Walton & Banaji, 2004](#)). It is possible that language has some effect in this category ([Gelman et al., 2000](#)) because when nouns are used to refer to something, one may have a deeper understanding of

it, which is noted to “enable inductive inferences” ([Gelman & O'Reilly, 1988](#)).

Once the subtle description is used to refer to oneself, a noun label may have a stronger effect. [Bryan et al. \(2011\)](#) found that more people would choose to vote if they heard the words “be a voter” rather than “to vote” on the day before election day. Additionally, research showed that, compared to “helping,” “being a helper” encouraged more children to conduct kind behaviors toward others ([Bryan et al., 2014](#)). However, subsequent research found that although “being a helper” can lead to more kind behaviors initially, once there is a setback, the backlash may also be stronger accordingly ([Foster-Hanson et al., 2018](#)). The reason underlying this phenomenon is as follows: as category labels, nouns bear a strong link to identity and may lead to self-doubt once one fails.

According to [Bryan et al. \(2011\)](#), the effect of noun expression comes from a motivation-driven process. When a noun is involved with a positive identity such as “voter” and “helper,” people simply see themselves as voters or helpers and they produce more correlated behaviors; When the noun is involved with undesirable (negative) identities, however, these kind of words should cause people to avoid correlated behaviors.

In social psychology, experiments of priming of unethical behaviors and its subsequent prevention typically involve money or time ([Gino & Mogilner, 2014](#); [Gino & Pierce, 2009](#); [Mogilner & Asker, 2009](#); [Vohs et al., 2006](#)). A mere exposure to money is associated with unethical outcomes ([Kouchaki et al., 2013](#)). In [Gino et al.'s](#) experiment (2014), participants were asked to complete a scrambled-sentences task using some money-related words or time-related words; results showed that priming time (rather than money) makes people behave more ethically.

In contrast, another experiment by [Bryan et al. \(2013\)](#) allowed experimenters to prevent unethical behaviors through semantic priming. They manipulated the task's instructions by changing the use of verbs (“Don't cheat”) to noun labels (“Don't be a cheater”) to inhibit participants from engaging in unethical behaviors. The self-identity related group (“don't be a cheater”) had significantly lower proportion of unethical cheating behaviors.

In the present study, we aim to replicate Experiment 3 of [Bryan et al. \(2013\)](#), for the following reasons:

First, the participants in Experiment 1 in [Bryan et al. \(2013\)](#) were asked to think of a number from 1 to 10. If the number was even, they were paid \$5; if it was odd, there was no reward. [Bryan et al. \(2013\)](#) paid for even numbers because it has been reported that participants typically show a strong bias toward odd numbers in a random number generation task ([Kubovy & Psootka, 1976](#)), but this oddness bias had not been confirmed for betting behaviors. Furthermore, an even or odd number participants think of is just imaginary, occurring in one's inside world, not an external real event; hence, it is difficult

to use it as an index of falsification. An index used for cheating should emphasize that participants' reports can differ from the fact. Thus, we abandoned the method of [Bryan et al. \(2013\)](#) Experiment 1. In their Experiments 2 and 3, they used a coin-tossing task: participants were asked to toss a coin and receive a reward corresponding to the result of their coin flips. We choose this method for our experiment because tossing a coin induces a real external event, which is more objective and operable, and hence it is better than thinking of a number to measure cheating behavior. In addition, compared with Experiment 2 in the original study, which just used two conditions, "cheater" and "cheating", a baseline group was included in Experiment 3, which made Experiment 3 more complete in its design—an approach we will follow also.

Moreover, we found that the effect size in Experiment 3 was small ( $f = 0.302$  in G\*Power (significance level  $\alpha = 0.05$ , power level  $1-\beta = 0.95$ ), meaning that Experiment 3 required at least 174 participants; in fact, only 99 people joined the original research. From this, we suppose that the effect size in Experiment 3 was overvalued.

According to the above review, high levels of self-identity and the willingness of individuals to maintain a positive self-view should prevent unethical behaviors. We predict that the self-relevant noun "cheater" will curb cheating behaviors more significantly than the verb "cheating" and the baseline condition (in which there is no reminder in the instruction).

## Methods

### Experiment 1

Our experiment will be conducted online in a private and impersonal way, which means that participants will not meet or be expected to meet the experimenters. We aim to replicate Experiment 3 of [Bryan et al. \(2013\)](#), in which there are three conditions: "cheater," "cheating," and "baseline"; in the baseline condition, a reminder about cheating will not be mentioned.

**Participants.** Participants will be users of the [Yahoo! Crowdsourcing Service](#) in Japan. Participants are required to meet the *a priori* criterion that they are native Japanese speakers. We plan to conduct a pilot test to determine the shortest time in which one could reasonably participate in the experiment in good faith. This pilot test is detailed in a later section (*Outlier extraction*). Participants will be excluded if they complete the experiment faster than the pilot test time. Repeat participation will be prevented.

**Procedure.** At the beginning of the experiment, demographic information of participants' age and gender will be collected. Online instructions will indicate that a recent controversial article has claimed to report the first scientific evidence for paranormal phenomena ([Bem, 2011](#)); this is the same cover story used in the original study.

We ask the participants to find a coin at home and, while trying to influence the outcome of each toss with their minds, flip the coin 10 times, making it land on "heads" as often as

possible. They will be asked to ensure that they are "properly motivated." For every toss landing on "heads," they will receive 10 T-points ( $\approx 10$  yen) as a reward. The instructions indicate that participants will receive a reward corresponding to the result of their coin flips. Noting that the average Yahoo! Crowdsourcing user receives two T-points for one task, the reward we set is five times that amount. Even those participants who report a chance level of five "heads" will receive 25 times more points than for a normal task. This reward should be large enough for some participants to engage in cheating behavior.

Participants will be randomly assigned the three conditions. All three conditions will be posted on the crowdsourcing website, and a person who participates in one of the conditions cannot participate in the others. This results in passive randomization. In the "cheater" and "cheating" conditions, the manipulation is embedded in the instructions. We referred to the instructions in the original experiment ([Bryan et al., 2013](#)) and translated it into Japanese, as follows:

NOTE: Please don't [cheat/be a cheater] and report that one or more of your coin flips landed heads when it really landed tails! Even a small [amount of cheating/number of cheaters] would undermine the study, making it appear that psychokinesis is real.

[嘘をつかない/嘘つきにならない]でください。コイントスの結果、実際にはコインの裏が出たのに表が出たと報告しないでください。

たとえわずかでも[嘘があると/嘘つきがいると], この研究を台無しにし、念力が実在するよう見せてしまいます。

The manipulation is also embedded in the instructions on the next page, where participants log the outcomes of their 10 coin-flips. At the top of the page, a message is displayed in large red words: "嘘をつかないでください/嘘つきにならないでください" This means "PLEASE DON'T [CHEAT/BE A CHEATER]" as in the original experiment ([Bryan et al., 2013](#)).

In the baseline condition, the instructions are the same as above, except that the cheating message is not mentioned.

**Power analysis and sample size.** Because Experiment 3 of [Bryan et al. \(2013\)](#) did not report the effect size,  $\eta^2$ , first, we calculated the effect size of the analysis of variance (ANOVA) result from the  $F$  and  $df$  values. [Bryan et al. \(2013\)](#) reported the statistics of their one-way ANOVA as  $F(2, 96) = 4.38$ ,  $p = .015$ . Hence, we calculated  $\eta^2$  based on [Cohen's \(1973\)](#) method, as  $\eta^2 = .0836$ . Then, we calculated the effect size,  $f$ , as follows:  $f = \sqrt{\eta^2 / (1 - \eta^2)} = 0.302$ . The small sample size may overestimate the effect size so, as a replication convention (e.g., [Nitta et al., 2018](#)), we halved the effect size of the original experiment, and used G\*Power 3.1.9.3 ([Faul et al., 2009](#)) to conduct a power analysis (i.e., to 0.151). In G\*Power, we set the significance level  $\alpha = 0.05$ , power level  $1-\beta = 0.95$ , and effect size  $f = 0.151$ . According to the conditions of the original experiment, we will divide the participants into three groups. The required total sample size is 681, with

227 participants in each group; therefore, we will try to recruit at least 681 participants, and data collection will not exceed 810 participants. This stopping rule is set because it is difficult for us to limit the number of participants to exactly 681, due to the characteristics of the simultaneous participatory online recruitment system; therefore, we will allow for up to 120% of the required sample size (i.e., 810). If more than 810 people participate in the experiment, we will select the data of the first 810 participants based on the time stamp and use this for the analysis. Also, we set the number of participants (max. 365 males and 445 females) to match the gender distribution of the original study (male: female = .45:.55).

**Data analyses.** In this study, the dependent variable is the mean number of “heads” reported. In the original experiment, a one-way ANOVA and *t*-test will be performed. Specifically, the ANOVA will be performed for analyzing the main effect of the three groups. A problem in the original study was that the authors did not report adjustments for any significance level in subsequent multiple comparisons. Therefore, in the present study, we will use a one-way ANOVA and Tukey’s method for the multiple comparisons. Additionally, in order to check the cheating in each group, the original study performed one-sample *t*-tests between the mean number of “heads” reported and the chance level (i.e., 50%). These analyses will be performed using jamovi (version 1.0.5). The original results are summarized in [Table 1](#).

Moreover, as the dependent variable is based on the counts of “heads” reported and that the 10 coin tosses are nested within each participant, a quasi-Poisson or Poisson regression will be used for exploratory analyses. In the (quasi-)Poisson model, the variance is assumed to be the mean multiplied by a dispersion parameter (Ma *et al.*, 2014). Dispersion parameters with a value greater than one indicate that overdispersion exists; in this case, quasi-Poisson regression will be performed. Thus, which analysis to use depends on the result of variance and the mean of “head” counts. We will first test the original hypothesis. Then, information of gender and age will be added as predictors to establish a regression model.

**Outlier extraction.** For our online experiment, we will establish a minimum completion time (MCT) for inclusion in the final

sample by asking five colleagues who are unfamiliar with this experiment to complete the experiments as fast as possible, then calculating the mean completion time. Specifically, each colleague will perform a coin toss ten times; after each toss they will record the result on the experiment website. This pilot test will not include the attempt to motivate psychokinesis and will measure only the required time of the coin toss and recording. Bryan *et al.* (2013) also used the MCT as an extraction criterion. We will exclude those participants who complete it faster than the MCT, because they may rush through the experiment and fail to complete it in good faith.

**Experiment 2**

This experiment is employed as an extended, conceptual replication of Experiment 3 in the original study (Bryan *et al.*, 2013). Our Experiment 2 is only performed when the results of Experiment 1 successfully replicate those of the original experiment. In the original experiment, the numbers of heads claimed in the “cheater” condition was significantly lower than that in the “cheating” and baseline conditions, but no difference was found between the “cheating” and baseline conditions. Here we cannot easily interpret the non-significant results based on self-identity alone. We aim to test whether lower levels of attention to the instruction in the “cheating” condition reduced the effectiveness of preventing dishonest behaviors in our Experiment 1. Thus, we conduct Experiment 2, adding a “cheating” with task condition in which we use tasks concerning an instruction to ensure that participants’ attention is captured (e.g., Folk *et al.*, 1992; Folk *et al.*, 2002). When we translated the instruction into Japanese, we felt the unfamiliarity of a “cheater” condition in a Japanese language situation. Participants in our experiment may find that the reminder “don’t be a cheater” commands extra attention because of this sense of deviation. Therefore, even if the result of the original experiment is completely reproduced in our Experiment 1, it will not fully support the finding of the original experiment, as the reason for the possible different dishonest behavior rates between the “cheating” and “cheater” conditions in our Experiment 1 may be that the participants in the “cheating” group paid relatively less attention to the instruction; for this reason, “cheating” may have worked weakly as a moral reminder in this condition. Because the experiments are conducted online, it is difficult to ensure that the participants have actually seen and understood the instruction; in addition, it is also possible that the participants ignored the instructions of Experiment 1 due to satisficing, (e.g., Chandler *et al.*, 2014; Oppenheimer *et al.*, 2009; Sasaki & Yamada, 2019), further diminishing the effect of the unattended reminder (i.e., “cheating”). In this Experiment 2 we address these attention-related effects.

Noticeably, the main difference between our Experiment 1 and the original Experiment 3 lies in the different language used in the instruction. Thus, if our Experiment 1 is a successful replication, we will then choose to focus on the expression used in the Japanese instruction, rather than the English instruction of the original Experiment 3.

To support this approach, we conducted a preliminary experiment, asking participants to evaluate their familiarity

**Table 1. Results of Experiment 3 of Bryan *et al.* (2013).**

Analysis types		Reported <i>p</i> -value	Degree of freedom	Effect size
Main effect: three groups		.015	96	<i>f</i> = 0.302
<i>t</i> -test	“cheating” vs “cheater”	.013	96	<i>d</i> = 0.71
	“cheater” vs baseline	.004	96	<i>d</i> = 0.66
	“cheating” vs baseline	> .80	96	<i>d</i> = 0.05
	“cheating” vs “chance”	< .0005	36	<i>d</i> = 0.79
	baseline vs “chance”	< .0005	35	<i>d</i> = 0.78
	“cheater” vs “chance”	> .30	25	<i>d</i> = 0.19

with certain expressions in Japanese. The expressions “Don’t cheat” and “Don’t be a cheater” were translated into Japanese, and native speakers evaluated their familiarity with them (1: not familiar to 5: very familiar) via an Internet survey on Yahoo! Crowdsourcing. The protocol of this experiment was registered on the Open Science Framework (Guo *et al.*, 2019). The results showed that the familiarity rating score in the “cheater” condition was significantly lower than that in the “cheating” condition,  $t(64) = 6.73$ ,  $p < .001$ , Cohen’s  $d = 0.834$ . Hence, we conjecture that the anticipated difference in the results between the “cheating” and “cheater” conditions in Experiment 1 may partly occur due to differences in attention paid to the instruction, instead of the preservation of a positive self-image proposed by the previous study (Bryan *et al.*, 2013). This means that part of the effect of the “cheater” condition is due to the unfamiliar expression, which attracts people’s attention then plays a role in preventing them from conducting unethical behavior. See *Extended data* for details about this experiment.

In our Experiment 2, we will manipulate the way in which participants see the instructions to explore the differences between the “cheating” and baseline conditions. Experiment 2 comprises three conditions: “cheating,” “cheating” with task, and baseline. We predict that the “cheating” with task condition will be more effective in curbing unethical behaviors than the “cheating” and baseline conditions, because the task will arouse more attention. While the instruction in the “cheating” condition will be in large red capital letters, this should entail no significant difference compared with baseline.

**Procedure.** The procedure for Experiment 2 is identical to that of Experiment 1, except for important differences in two aspects. In Experiment 2, we will focus on whether the participants read the instructions as diligently as we expect. First, we will delete the original “cheater” condition and add another “cheating” condition (i.e., “cheating” with task condition). Second, in the “cheating” with task condition, we will add a task page in which participants are asked to choose the exact expression (i.e., “Don’t cheat”) that appeared on the screen from three sample sentences. We will remind participants of this task in advance to ensure they read the instructions carefully.

**Power analysis and participants.** Because the power analysis of Experiment 2 is the same as in Experiment 1, we intend to recruit participants in the same way as Experiment 1. The minimum completion time will also be established for participants to be included in the final sample. This exclusion standard is similar to that in Experiment 1.

**Data analyses.** In Experiment 2, the dependent variable is the mean number of “heads” reported. We will still use a one-way ANOVA and Tukey’s method for the multiple comparisons. To check the cheating rate in each group, a one-sample  $t$ -test between the mean number of “heads” reported and the chance level (50%) will be analyzed. The data of participants who failed to provide the right answer to the attention task will not be used for further analysis. Another analysis by a (quasi-)Poisson regression model will also be performed to explore the contribution factors of cheating counts.

### Study timeline

Currently, the online experiments for participants to conduct the coin-toss task are under construction. After Stage 1 acceptance, our colleagues will be asked to complete the pilot test to calculate the MCT. Then, we will post our experiments on the Yahoo! Crowdsourcing Service to recruit participants. We are supposed to complete the experiments and subsequent analysis within two months.

### Ethical approval and consent to participate

The present study received approval from the psychological research ethics committee of the Faculty of Human-Environment Studies at Kyushu University (approval number: 2019-004). Completion of experiments by participants will be regarded as consent to participate; they will also have the right to withdraw from the experiment at any time without providing a reason. In addition, we will protect participants’ personal information. Because this study will be conducted online, even if participants engage in cheating behaviors, we cannot identify them or meet the participants face-to-face.

### Data availability

#### Underlying data

No underlying data are associated with this article.

#### Extended data

Open Science Framework: How subtle linguistic cues prevent unethical behaviors, <https://doi.org/10.17605/OSF.IO/68FVK> (Guo *et al.*, 2019).

This project contains the following extended data:

- Protocol for the pilot study conducted for Experiment 2.
- Data collected for the pilot study conducted for Experiment 2.

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

## References

- Bandura A: **Moral disengagement in the perpetration of inhumanities.** *Pers Soc Psychol Rev.* 1999; 3(3): 193–209.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bem DJ: **Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect.** *J Pers Soc Psychol.* 2011; 100(3): 407–25.  
[PubMed Abstract](#) | [Publisher Full Text](#)

- Blasi A: **Moral identity: Its role in moral functioning.** *Morality, Moral Behavior, and Moral Development.* 1984; 128–139.
- Bryan CJ, Adams GS, Monin B: **When cheating would make you a cheater: implicating the self prevents unethical behavior.** *J Exp Psychol Gen.* 2013; 142(4): 1001–5.  
[PubMed Abstract](#) | [Publisher Full Text](#)

- Bryan CJ, Master A, Walton GM: **"Helping" versus "being a helper": invoking the self to increase helping in young children.** *Child Dev.* 2014; **85**(5): 1836–1842.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bryan CJ, Walton GM, Rogers T, et al.: **Motivating voter turnout by invoking the self.** *Proc Natl Acad Sci U S A.* 2011; **108**(31): 12653–12656.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chandler J, Mueller P, Paolacci G: **Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers.** *Behav Res Methods.* 2014; **46**(1): 112–130.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cohen J: **Eta-squared and partial eta-squared in fixed factor ANOVA designs.** *Educ Psychol Meas.* 1973; **33**(1): 107–112.  
[Publisher Full Text](#)
- Cohen TR, Morse L: **Moral character: What it is and what it does.** *Res Organ Behav.* 2014; **34**: 43–61.  
[Publisher Full Text](#)
- Cojuharencu I, Shteynberg G, Gelfand M, et al.: **Self-construal and unethical behavior.** *J Bus Ethics.* 2012; **109**(4): 447–461.  
[Publisher Full Text](#)
- Faul F, Erdfelder E, Buchner A, et al.: **Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses.** *Behav Res Methods.* 2009; **41**(4): 1149–1160.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Fausey CM, Boroditsky L: **Subtle linguistic cues influence perceived blame and financial liability.** *Psychon Bull Rev.* 2010; **17**(5): 644–650.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Folk CL, Leber AB, Egeth HE: **Made you blink! Contingent attentional capture produces a spatial blink.** *Percept Psychophys.* 2002; **64**(5): 741–753.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Folk CL, Remington RW, Johnston JC: **Involuntary covert orienting is contingent on attentional control settings.** *J Exp Psychol Hum Percept Perform.* 1992; **18**(4): 1030–44.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Foster-Hanson E, Cimpian A, Leshin RA, et al.: **Asking Children to "Be Helpers" Can Backfire After Setbacks.** *Child Dev.* 2018; 1–13.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gelman SA, Heyman GD: **Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories.** *Psychol Sci.* 1999; **10**(6): 489–493.  
[Publisher Full Text](#)
- Gelman SA, Hollander M, Star J, et al.: **The role of language in the construction of kinds.** *Psychol Learn Motiv.* 2000; **39**: 201–263.  
[Publisher Full Text](#)
- Gelman SA, O'Reilly AW: **Children's inductive inferences within superordinate categories: the role of language and category structure.** *Child Dev.* 1988; **59**(4): 876–887.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Gino F, Mogilner C: **Time, money, and morality.** *Psychol Sci.* 2014; **25**(2): 414–421.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Gino F, Pierce L: **The abundance effect: Unethical behavior in the presence of wealth.** *Organizational Behavior and Human Decision Processes.* 2009; **109**(2): 142–155.  
[Publisher Full Text](#)
- Guo W, Liu H, Yang J, et al.: **How subtle linguistic cues prevent unethical behaviors.** 2019.  
<http://www.doi.org/10.17605/OSF.IO/68FVK>
- Jones E, Nisbett R: **The actor and the observer: Divergent perceptions of the causes of behavior.** *Attribution: Perceiving the Causes of Behavior.* 1972; 79–94.
- Kouchaki M, Smith-Crowe K, Brief A, et al.: **Seeing green: Mere exposure to money triggers a business decision frame and unethical outcomes.** *Organizational Behavior and Human Decision Process.* 2013; **121**(1): 53–61.  
[Publisher Full Text](#)
- Kubovy M, Psotka J: **The predominance of seven and the apparent spontaneity of numerical choices.** *J Exp Psychol Hum Percept Perform.* 1976; **2**(2): 291–294.  
[Publisher Full Text](#)
- Ma L, Yan X, Qiao W: **A quasi-Poisson approach on modeling accident hazard index for urban road segments.** *Discrete dynamics in nature and society.* 2014; **2014**(3): 8.  
[Publisher Full Text](#)
- Markus H, Wurf E: **The dynamic self-concept: A social psychological perspective.** *Annu Rev Psychol.* 1987; **38**(1): 299–337.  
[Publisher Full Text](#)
- Mazar N, Amir O, Ariely D: **The dishonesty of honest people: A theory of self-concept maintenance.** *J Mark Res.* 2008; **45**(6): 633–644.  
[Publisher Full Text](#)
- Mogilner C, Aaker J: **"The time vs. money effect": Shifting product attitudes and decisions through personal connection.** *Journal of Consumer Research.* 2009; **36**(2): 277–291.  
[Publisher Full Text](#)
- Monin B, Jordan A: **The dynamic moral self: A social psychological perspective.** *Personality, Identity, and Character: Explorations in Moral Psychology.* 2009; **7**: 341–354.  
[Publisher Full Text](#)
- Nitta H, Tomita H, Zhang Y, et al.: **Disgust and the rubber hand illusion: a registered replication report of Jalal, Krishnakumar, and Ramachandran (2015).** *Cogn Res Princ Implic.* 2018; **3**(1): 15.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Oppenheimer DM, Meyvis T, Davidenko N: **Instructional manipulation checks: Detecting satisficing to increase statistical power.** *J Exp Soc Psychol.* 2009; **45**(4): 867–872.  
[Publisher Full Text](#)
- Ross L: **The intuitive psychologist and his shortcomings: Distortions in the attribution process.** In: *Adv Exp Soc Psychol.* 1977; **10**: 173–220.  
[Publisher Full Text](#)
- Sasaki K, Yamada Y: **Crowdsourcing visual perception experiments: a case of contrast threshold.** *PeerJ.* 2019; **7**: e8339.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Steele CM: **The psychology of self-affirmation: Sustaining the integrity of the self.** In: *Adv Exp Soc Psychol.* 1988; **21**: 261–302.  
[Publisher Full Text](#)
- Vohs K, Mead N, Goode M: **The psychological consequences of money.** *Science.* 2006; **314**(5802): 1154–1156.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Walton GM, Banaji MR: **Being what you say: The effect of essentialist linguistic labels on preferences.** *Soc Cogn.* 2004; **22**(2): 193–213.  
[Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status:  

Version 3

Reviewer Report 03 March 2020

<https://doi.org/10.5256/f1000research.24963.r60776>

© 2020 Cervera-Torres S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Sergio Cervera-Torres** 

Leibniz-Institute für Wissensmedien (IWM), Tübingen, Germany

The report has very much improved and the authors have answered all my concerns in a reasonable way.

Perhaps, I have a couple of minor points to outline regarding Experiment 2:

- a) I think that the rationale behind Experiment 2 can be reorganized a bit to understand better the motivation of the "cheating with task" condition. The explanation that you offer to one of my questions (Linguistic issue; point 3) is very clarifying and could be used as a guideline in this regard.

In a similar vein,

- b) If I understand well, both, the "cheating" and "cheating with task" conditions will/should not differ with regard to the baseline. Please, make that point more explicit in the text as you did in your answer.

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Cognitive Psychology. Clinical Psychology. Emotion processing

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 10 Mar 2020

**Yuki Yamada**, Kyushu University, Fukuoka, Japan

Thank you for your further comments to improve the quality of this manuscript. We will respond to the individual comments below.

*a) I think that the rationale behind Experiment 2 can be reorganized a bit to understand better the motivation of the "cheating with task" condition. The explanation that you offer to one of my*



questions (Linguistic issue; point 3) is very and could be used as a guideline in this regard.

**Reply:**

Thank you for your kind suggestion. We have added related sentences to the introduction of Experiment 2 for clarity.

*In a similar vein,*

*b) If I understand well, both, the "cheating" and "cheating with task" conditions will/should not differ with regard to the baseline. Please, make that point more explicit in the text as you did in your answer.*

**Reply:**

Our previous answer was related to one of the possible results that could come about if there is no difference among the three conditions in Experiment 2; we wanted to show how to exclude the influence of attention bias.

In actuality, the pilot experiment showed that there was a significant difference in the familiarity of expressions in the Japanese reminders, so the dishonest behavior rate may be influenced by different levels of attention. For this reason, we predicted that the "cheating" with task condition would induce significantly lower numbers of heads claimed than the "cheating" and baseline conditions. We have clarified this in the manuscript.

**Competing Interests:** No competing interests were disclosed.

---

**Version 2**

Reviewer Report 18 February 2020

<https://doi.org/10.5256/f1000research.24062.r58686>

© 2020 Asano M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Michiko Asano**

Department of Psychology, College of Contemporary Psychology, Rikkyo University, Saitama, Japan

I feel that the proposal has been substantially improved. The authors have satisfactorily addressed all the issues I had raised. However, I have some minor comments/suggestions on the revised manuscript, which I address below.

**Minor comments:**

1. Data analysis of the "cheating with task" condition in Experiment 2:  
If you are going to delete the data of participants who fail to give the right answer to the attention task, this information should be provided in the main text.

2. The Japanese translation of the instructions in Experiment 1:  
The orders of the [/] (don't be a cheater/don't cheat) and ["/"] "PLEASE DON'T [BE A CHEATER/CHEAT]" need to be reversed so that they would be consistent with the corresponding English descriptions.
3. I agree with the comment by Dr. Sergio Cervera-Torres on the readability of the introduction. The storyline can be improved.
4. The fourth paragraph of Experiment 2:  
"In our Experiment 2, we manipulated the way in which participants saw..."  
-> "In our Experiment 2, we will manipulate the way in which participants see..."?

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Cognitive psychology.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 26 Feb 2020

**Yuki Yamada**, Kyushu University, Fukuoka, Japan

Thank you very much for approving this manuscript. Here, we will respond to your comments individually.

*I feel that the proposal has been substantially improved. The authors have satisfactorily addressed all the issues I had raised. However, I have some minor comments/suggestions on the revised manuscript, which I address below.*

**Reply:**

We appreciate your positive evaluation and further suggestions. We have further revised the manuscript, taking into account your kind comments.

*1. Data analysis of the "cheating with task" condition in Experiment 2:*

*If you are going to delete the data of participants who fail to give the right answer to the attention task, this information should be provided in the main text.*

**Reply:**

Thank you for your advice. We have added the relevant information in our manuscript as follows: "The data of participants who failed to provide the right answer to the attention task will not be used for further analysis."

*2. The Japanese translation of the instructions in Experiment 1:*

*The orders of the [/] (don't be a cheater/don't cheat) and ["/"] "PLEASE DON'T [BE A CHEATER/CHEAT]" need to be reversed so that they would be consistent with the corresponding English descriptions.*

**Reply:**

Thank you very much for such attention to detail. We have adjusted the order of the reminders in our manuscript to correspond to the original text.

*3. I agree with the comment by Dr. Sergio Cervera-Torres on the readability of the introduction. The*

storyline can be improved.

**Reply:**

We have corrected the logical flow of the introduction.

**4. The fourth paragraph of Experiment 2:**

“In our Experiment 2, we manipulated the way in which participants saw...”

-> “In our Experiment 2, we will manipulate the way in which participants see...”?

**Reply:**

As a registered report, you are correct, this should be written using the future tense. We have changed the tense according to your suggestion.

**Competing Interests:** The authors declare no competing interests.

Reviewer Report 06 February 2020

<https://doi.org/10.5256/f1000research.24062.r58909>

© 2020 Cervera-Torres S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Sergio Cervera-Torres** 

Leibniz-Institute für Wissensmedien (IWM), Tübingen, Germany

This study aims at replicating, in a Japanese sample, Experiment 3 by Bryan et al. (2013), which investigates whether and how subtle cheating-related ethical reminders (don't cheat/be a cheater) subsequently prevent cheating behavior. In general, I find the proposal interesting, informative, and worthy to be further conducted. However, I have some concerns/suggestions after carefully reading the report.

**Minor aspects concerning the introduction:**

1. I find the storyline a bit difficult to read. As a suggestion, and if the authors consider it appropriate, I propose reordering some pieces of information. After the first paragraph, I would introduce [According to Blasi (1984)... till (Cojuharenco *et al.*)] followed by [a moral character model... till “euphemistic labeling” regarding language]. Then, [different ways of description...till (Gelman *et al.*)] followed first by the studies regarding ethical behaviour and [According to Bryan *et al.* (2011)...correlated behaviors]. Finally, I would introduce Bryan *et al.* (2013), which is the main study of reference in the current proposal. I think that the motivation of the study could be reinforced beyond the methodological arguments. For example, are there other studies investigating such prevention of unethical behaviour linked to semantic priming? Has this question been investigated in a different cultural setting at all?
2. Please amend the term “decrease” and use “prevent” “reflects less proportion” or something similar. Decrease or increase denotes change. This is not the main point of the study but rather testing the hypothesis that the “cheater” condition will significantly reflect less ratio of cheating behaviour. In the abstract, amend also “no instruction” as control and use baseline control or non-related cheating instruction or something similar. Participants in this group will also have

instructions.

3. Does your hypothesis skip the baseline condition for any particular reason? (e.g., “cheater” curbing unethical behavior *more* than both “cheating” and “baseline”).

#### **Methodology section:**

1. I understand that you will use ANOVAS to compare your results with those by Brian *et al.* However, I think that the study will be methodologically more sound if, in addition, you perform an alternative analysis. Considering that your DV is based on counts and that the 10 trials are nested within each participant (therefore probably correlated), I propose a Poisson or quasi-Poisson regression from a generalized estimating equations approach to investigate differences in proportions and Odds Ratio (see Ruiz Fernández, Kastner, Cervera-Torres, Müller, & Gerjets (2019)). You can include gender, age or, another demographic predictor such as job status (working/not working). The analysis can be performed first without such predictors to stick with the original hypothesis and then with the predictors.
2. In the procedure section, I would change “stimulus” I the first lines with “cover story”. In the data analyses section, the term “original” is confusing because it refers to you experiment and Bryans’. Unless you are sure that they didn’t, I suggest stating that [a problem in the original study was that the authors did not report adjustments for any significance level...]

#### **Linguistic issue:**

The study by Bryan *et al.* (2013) and the proposed replication relay on the core assumption that instructions compelling self-identity (“don’t be a cheater”) prevent cheating behavior due to “increased” self-identity activation/social-desirability bias. You suspect that the direct translation of the original English instructions “don’t be a cheater” into Japanese might be perceived as rather unfamiliar. In other words, your expected effects might be potentially due to (a) the remainder “cheater” activates self-identity more than the remainder “cheat” and/or (b) the remainder “cheater” promotes extra attentional salience due to unfamiliarity, which, as the authors state, may be problematic to fully support the initial theoretical assumptions. I have some questions in this regard.

1. Please, could you clarify whether your preliminary pilot study testing the familiarity of the expressions is based on direct translations from English or analogous expressions in Japanese? It is a bit confusing as it is written in the text “...asking participants to evaluate their familiarity with certain expressions in Japanese. The expressions *Don’t cheat* and *Don’t be a cheater* were translated into Japanese”. In my opinion, using genuine Japanese expressions instead of direct translations should be adequate for a conceptual (cultural) replication.
2. I am wondering if you could find a way to examine whether familiarity predicts/moderates the expected effects. It could be the case that participants grasp the meaning of the expression even if they find it relatively unfamiliar. For example, like a loose idea, a baseline task could be created where participants can rate the familiarity of a series of words/expressions in which the ethical reminders or related words of interest are included. Or, in addition, a multiple-choice task could be designed to reinforce the meaning of the expressions while at the same time potentially reducing unfamiliarity (e.g., in your opinion, which meaning/description do you think that fits better with this expression). If the authors consider that this question is beyond the scope of the study, it would be necessary at least to make this point clear in the discussion.

3. Beyond familiarity, why do you plan to exclude the “cheater” condition in Experiment 2? If you assume that the “cheating” condition captures less attention than the “cheater” one and design a “cheating with task” condition to compensate somehow the attentional bias, should you not compare this new condition with the “cheaters”? I think that you would be able to potentially provide stronger arguments supporting that the attentional bias is less likely to explain the expected findings in Experiment 1.

**Have the authors pre-specified sufficient outcome-neutral tests for ensuring that the results obtained can test the stated hypotheses, including positive controls and quality checks?**

Partly

**Is the rationale for, and objectives of, the study clearly described?**

Partly

**Is the study design appropriate for the research question?**

Partly

**Are sufficient details of the methods provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

No

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Cognitive Psychology. Clinical Psychology. Emotion processing

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 26 Feb 2020

**Yuki Yamada**, Kyushu University, Fukuoka, Japan

Thank you for your effort and time in reviewing this manuscript. We will respond individually to your comments as follows:

*This study aims at replicating, in a Japanese sample, Experiment 3 by Bryan et al. (2013), which investigates whether and how subtle cheating-related ethical reminders (don't cheat/be a cheater) subsequently prevent cheating behavior. In general, I find the proposal interesting, informative, and worthy to be further conducted. However, I have some concerns/suggestions after carefully reading the report.*

**Reply:**

We appreciate the use of your valuable time for reviewing this manuscript. We are really pleased that you accurately understand this research and have given it a positive evaluation. We have substantially revised the manuscript to take your kind comments into account.

**Minor aspects concerning the introduction:**

1. *I find the storyline a bit difficult to read. As a suggestion, and if the authors consider it appropriate, I propose reordering some pieces of information. After the first paragraph, I would introduce [According to Blasi (1984)... till (Cojuharenco et al.)] followed by [a moral character model... till “euphemistic labeling” regarding language]. Then, [different ways of description...till (Gelman et al.)] followed first by the studies regarding ethical behaviour and [According to Bryan et al. (2011)...correlated behaviors]. Finally, I would introduce Bryan et al. (2013), which is the main study of reference in the current proposal.*

**Reply:**

Thank you for your insightful comment. We have adjusted the order of the introduction in accordance with your suggestion.

First, we introduced how dishonest acts can be weakened, followed by an explanation of the relationship between dishonest behavior and self-identity based on the words that are being used. We then elaborated how a subtle difference in description affects people’s judgments or behaviors, especially when these words are connected to an individual’s identity. We further introduced Bryan et al.’s experiment, noting its purpose as well as the reasons we wanted to replicate this research. Regarding the priming effects that you mentioned, as far as we know, similar research that uses semantic priming to prevent unethical behavior has not been conducted. Instead, previous research on priming for unethical behavior has mainly focused on monetary priming (Gino & Mogilner, 2014; Gino & Pierce, 2009; Vohs et al., 2006). A mere exposure to the money construct is associated with unethical outcomes (Kouchaki et al., 2013). Other studies have explored an inhibition effect of time priming on cheating behavior (Gino & Mogilner, 2014; Mogilner & Asker, 2009). In Gino et al.’s experiment (2014), participants were asked to complete a scrambled-sentences task using some money-related words or time-related words; results showed that priming time (rather than money) induces people to behave more ethically. Although the primer stimuli used were different from those in our experiment (time vs. language), both experiments show how to prevent unethical behaviors by encouraging participants to be more conscious of self-identity so as to maintain a positive self-image.

Moreover, we have additional information regarding your question of whether similar studies have been performed in other cultures. Tomer and Eyal also replicated Bryan et al.’s (2013) experiment in Israel (Savir & Gamliel, 2019). All of their participants were native Hebrew speakers. Although the experiments were conducted in a completely different culture, instructions and reminders were still shown in English. They concluded that positive-valence words relating to the self (e.g., “Be an honest person”) have a similar effect on preventing unethical behavior as a negative-valence reminder (“Don’t be a cheater”).

Since the introduction of the experiment results above may help readers to understand the purpose and rationale of our replication better, we decided to substantially reorganize these parts of the manuscript. We apologize that we were unable to introduce all the individual minor changes because they are so many, but please see the text for details. Thank you very much for your suggestion.

2. *Please amend the term “decrease” and use “prevent” “reflects less proportion” or something similar. Decrease or increase denotes change. This is not the main point of the study but rather testing the hypothesis that the “cheater” condition will significantly reflect less ratio of cheating behaviour. In the abstract, amend also “no instruction” as control and use baseline control or non-related cheating instruction or something similar. Participants in this group will also have instructions.*

**Reply:**

Thank you for your suggestions. We amended the words “no instruction” to “baseline” in the abstract. In addition, we replaced the word “decrease” with “prevent” in the introduction as you proposed.

*3. Does your hypothesis skip the baseline condition for any particular reason? (e.g., “cheater” curbing unethical behavior more than both “cheating” and “baseline”).*

**Reply:**

As in our introduction, we mainly discussed how different types of descriptions, especially different expressions of the same meaning in terms of a noun or a verb, influence people’s behaviors. Our hypothesis now covers the baseline condition as follows: “We predict that the self-relevant noun “cheater” will curb cheating behaviors more significantly than the verb “cheating” and the baseline condition (in which there is no reminder in the instruction).” We have added these sentences to the manuscript.

**Methodology section:**

*1. I understand that you will use ANOVAS to compare your results with those by Brian et al. However, I think that the study will be methodologically more sound if, in addition, you perform an alternative analysis. Considering that your DV is based on counts and that the 10 trials are nested within each participant (therefore probably correlated), I propose a Poisson or from a generalized estimating equations approach to investigate differences in proportions and Odds Ratio (see Ruiz Fernández, Kastner, Cervera-Torres, Müller, & Gerjets (2019)). You can include gender, age or, another demographic predictor such as job status (working/not working). The analysis can be performed first without such predictors to and then with the predictors.*

**Reply:**

Thank you for the helpful suggestion. We completely agree with you and decided to add a Poisson or quasi-Poisson regression as exploratory analyses, as the quasi-Poisson regression model is capable of considering overdispersed data, which is a common characteristic in cheating counts. In the Poisson model, the variance is assumed to be the mean multiplied by a dispersion parameter (Ma et al., 2014). The dispersion parameters values that are greater than one indicated that overdispersion exists; in this case, we will use a quasi-Poisson regression for the analysis. Thus, our decision on the analysis will depend on the variance and mean of counts that participants report for “heads.”

In addition, we have two experiments in this study; Experiment 2 will be implemented depending on the results of ANOVA in Experiment 1, not on those of a regression analysis. Thus, the results of the regression analysis will not influence the implementation of Experiment 2. If Experiment 2 is performed, we will also perform a (quasi-)Poisson regression analysis in the results section of Experiment 2. The results of such analyses will be used for discussion in the “General Discussion” part.

We have added these analysis plans and criteria to the revised manuscript.

*2. In the procedure section, I would change “stimulus” I the first lines with “cover story”. In the data analyses section, the term “original” is confusing because it refers to you experiment and Bryans’. Unless you are sure that they didn’t, I suggest stating that [a problem in the original study was that the authors did not report adjustments for any significance level...]*

**Reply:**

Thank you for your suggestion, we have amended the words to “cover story.” If you felt confusion about what the term “original” was referring to, it is likely that other readers may face the same confusion. Hence, we decided to follow your suggestions and have modified the corresponding

parts of the revised manuscript.

**Linguistic issue:**

*The study by Bryan et al. (2013) and the proposed replication rely on the core assumption that instructions compelling self-identity (“don’t be a cheater”) prevent cheating behavior due to “increased” self-identity activation/social-desirability bias. You suspect that the direct translation of the original English instructions “don’t be a cheater” into Japanese might be perceived as rather unfamiliar. In other words, your expected effects might be potentially due to (a) the remainder “cheater” activates self-identity more than the remainder “cheat” and/or (b) the remainder “cheater” promotes extra attentional salience due to unfamiliarity, which, as the authors state, may be problematic to fully support the initial theoretical assumptions. I have some questions in this regard.*

*1. Please, could you clarify whether your preliminary pilot study testing the familiarity of the expressions is based on direct translations from English or analogous expressions in Japanese? It is a bit confusing as it is written in the text “...asking participants to evaluate their familiarity with certain expressions in Japanese. The expressions Don’t cheat and Don’t be a cheater were translated into Japanese”. In my opinion, using genuine Japanese expressions instead of direct translations should be adequate for a conceptual (cultural) replication.*

**Reply:**

Our expression is a direct translation from English. We chose to translate the original English instruction to Japanese directly not because we were hoping to see the difference in familiarity with the two kinds of expressions, but, due to the grammatical features in Japanese, an analogous translation would be difficult. In Japanese, there is a tendency for the subject to be omitted and the corresponding noun connected to that verb to be used less. Thus, if we were to adopt an analogous translation that sounds more familiar, little difference (including whether a verb or a noun is used) would remain between the two kinds of instruction, and the main independent variable in the experiment would disappear.

*2. I am wondering if you could find a way to examine whether familiarity predicts/moderates the expected effects. It could be the case that participants grasp the meaning of the expression even if they find it relatively unfamiliar. For example, like a loose idea, a baseline task could be created where participants can rate the familiarity of a series of words/expressions in which the ethical reminders or related words of interest are included. Or, in addition, a multiple-choice task could be designed to reinforce the meaning of the expressions while at the same time potentially reducing unfamiliarity (e.g., in your opinion, which meaning/description do you think that fits better with this expression). If the authors consider that this question is beyond the scope of the study, it would be necessary at least to make this point clear in the discussion.*

**Reply:**

We do not assume that participants are unable to fully understand the instruction because of unfamiliarity with it. In fact, the Japanese reminder, though not commonly used, can still be understood well by native speakers. We aim to emphasize that, since the unfamiliarity of reminders attracts more attention, the participants are more likely to take note of a less common phrase. Since “don’t cheat” is a very common expression (similar to the slogan “Smoking causes lung cancer” on a cigarette case), participants may ignore it, thus reducing its effect as a moral reminder.

However, your suggestion is very meaningful, and we include a more detailed and clear explanation based on the results of our experiment in the discussion.



3. *Beyond familiarity, why do you plan to exclude the “cheater” condition in Experiment 2? If you assume that the “cheating” condition captures less attention than the “cheater” one and design a “cheating with task” condition to compensate somehow the attentional bias, should you not compare this new condition with the “cheaters”?*

**Reply:**

Experiment 2 is designed and will be conducted only when the results of our Experiment 1 completely replicate the results of the original experiment. In the original experiment, the dishonesty rate for the “don’t be a cheater” condition was significantly lower than that for the “don’t cheat” condition and the baseline, but there was no difference in the dishonesty rate between the “don’t cheat” condition and the baseline. We cannot explain this lack of a significant difference, and we want to analyze the effect of attention (in other words, exclude or confirm whether this happened due to participants’ attention) in our experiment in Japanese. When participants pay enough attention to the reminder “don’t cheat,” is there still no difference in the dishonesty rate between the cheating group and the baseline group? If the results in our Experiment 1 are the same as those in the original experiment, we aim to explore whether lower levels of attention to the cheating condition reduced the effectiveness of preventing dishonest behaviors in our Experiment 1. For this reason, we added a “cheating with task” condition to ensure that the participants actually pay enough attention to the reminder. If there is still no difference in the dishonesty rate for the “cheating with task” condition compared to the baseline, then we can say that the different rate of cheating between “don’t be a cheater” and “don’t cheat” is not because of attention bias.

In this way, your thoughtful and constructive comments have greatly improved our manuscript. We express our sincere gratitude again. We would appreciate it if you approved this manuscript.

**References**

- Gino, F., & Mogilner, C. (2014). Time, money, and morality. *Psychological Science, 25*(2), 414-421.
- Gino, F., & Pierce, L. (2009). The abundance effect: Unethical behavior in the presence of wealth. *Organizational Behavior and Human Decision Processes, 109*(2), 142-155.
- Kouchaki, M., Smith-Crowe, K., Brief, A. P., & Sousa, C. (2013). Seeing green: Mere exposure to money triggers a business decision frame and unethical outcomes. *Organizational Behavior and Human Decision Processes, 121*(1), 53-61.
- Ma, L., Yan, X., & Qiao, W. (2014). A quasi-Poisson approach on modeling accident hazard index for urban road segments. *Discrete dynamics in nature and society, 2014*.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633-644.
- Mogilner, C., & Aaker, J. (2009). “The time vs. money effect”: Shifting product attitudes and decisions through personal connection. *Journal of Consumer Research, 36*(2), 277-291.
- Savir, T., & Gamliel, E. (2019). To be an honest person or not to be a cheater: Replicating the effect of messages relating to the self on unethical behaviour. *International Journal of Psychology, 54*(5), 650-658.
- Vohs, K. D., Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *Science, 314*(5802), 1154-1156.

**Competing Interests:** The authors declare no competing interests.

---

**Version 1**

Reviewer Report 24 October 2019

<https://doi.org/10.5256/f1000research.22170.r55680>

© 2019 Asano M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Michiko Asano**

Department of Psychology, College of Contemporary Psychology, Rikkyo University, Saitama, Japan

**Summary**

This study is designed to examine an interesting topic of whether and how subtle linguistic cues may prevent unethical behaviors. The proposal is generally well written. However, I have some concerns, which I address below.

1. Introduction: It is not clear to me why the authors decided to replicate Bryan *et al.* (2013)<sup>1</sup> among many other studies. I understand from the proposal the reason why they chose Experiment 3 of Bryan *et al.* (2013) over Experiment 1 or 2. I also understand that the sample size in Experiment 3 of Bryan *et al.* was not sufficient and thus the results may require replication with a larger sample. However, I do not understand the importance of replicating their results in addressing the research question of whether and how subtle linguistic cues prevent unethical behaviors. The authors should elaborate on this issue to clarify the significance of their study (e.g., "It is important to replicate Bryan *et al.*'s study as it has a significant impact on this field." or "The protocol of Experiment 3 is useful and applicable to future studies. Therefore, the robustness of their findings should be tested carefully).
2. Please state in the second to last paragraph of the Introduction that, unlike Experiment 1, Bryan *et al.* (2013) used a coin-tossing task in Experiments 2 and 3. Otherwise, this paragraph is hard to follow.
3. The procedure of Experiment 1 in the current study is not described sufficiently. Please state that the cover story (examination of a paranormal phenomenon) is exactly the same as that of Experiment 3 of Bryan *et al.* (2013).

It is also necessary to specify how the authors will explain to their participants about the relationship between the coin flips and the amount of reward. Bryan *et al.* (2013) used the following description: "The instructions acknowledged that the laws of probability dictate that people would, on average, make \$5, although some would 'make as much as \$10 just by chance' and others would 'make as little as \$0'".

If the procedure is identical to that of Experiment 3 of Bryan *et al.* (2013) except for the use of the Japanese language, stating so would help readers understand the protocol.

4. The authors list the following two points as the reason why they plan to direct participants' attention toward the ethical reminder in Experiment 2. (a) First, there was no significant difference in the rate

of cheating between “don’t cheat” and the baseline (i.e., no ethical reminder) conditions in the study by Bryan *et al.* The authors conjecture that participants in Bryan *et al.*’s did not pay sufficient attention to the ethical reminder. (b) Second, Their preliminary experiment has revealed that the Japanese translation of the expression “don’t be a cheater” (), compared to that of “don’t cheat” (), is less familiar to Japanese speakers and thus may attract more attention. Such difference in familiarity may produce results similar to Experiment 3 of Bryan *et al.* (2013), irrespective of whether the expression bears a strong link to self-identity or to the action. To address these concerns, the authors are planning to add “cheating (‘don’t cheat’) with task” condition, in which they test whether participants paid attention to the ethical reminder.

As for (a), I am not sure whether the lack of difference in cheating rates between “don’t cheat” and the baseline conditions in Bryan *et al.*’s study necessarily means that the participants did not pay attention to the ethical reminder. The cheating rate was significantly lower in “don’t be a cheater” condition than in the “don’t cheat” condition in their study. Doesn’t this mean that the participants paid sufficient attention to the ethical reminder?

As for (b), I agree with the authors that ethical reminders that attract more attention may prevent unethical behaviors more strongly. However, I do not understand what hypothesis the authors will test using the protocol of Experiment 2. Are they trying to show that it is attention to ethical reminders, rather than the linguistic expression linked to self-identity, that prevents unethical behaviors? If so, how do they interpret the results of Bryan *et al.* (2013)?

5. I do not understand the task of the “cheating with task” condition in Experiment 2. Will the authors use this task to classify participants into those who paid attention to the ethical reminder and those who did not, and analyze the cheating rate separately for these two groups? Or is the task just for bringing participants’ attention to the ethical reminder before the coin-tossing task?

## References

1. Bryan CJ, Adams GS, Monin B: When cheating would make you a cheater: implicating the self prevents unethical behavior. *J Exp Psychol Gen.* 2013; **142** (4): 1001-5 [PubMed Abstract](#) | [Publisher Full Text](#)

**Have the authors pre-specified sufficient outcome-neutral tests for ensuring that the results obtained can test the stated hypotheses, including positive controls and quality checks?**

Partly

**Is the rationale for, and objectives of, the study clearly described?**

Partly

**Is the study design appropriate for the research question?**

Partly

**Are sufficient details of the methods provided to allow replication by others?**

Partly

**Are the datasets clearly presented in a useable and accessible format?**

No

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Cognitive psychology.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 18 Dec 2019

**Yuki Yamada**, Kyushu University, Fukuoka, Japan

Thank you very much for taking the time and effort to check our manuscript. we also appreciate your helpful comments. Individual answers are given below.

*1. Introduction: It is not clear to me why the authors decided to replicate Bryan et al. (2013)<sup>1</sup> among many other studies. I understand from the proposal the reason why they chose Experiment 3 of Bryan et al. (2013) over Experiment 1 or 2. I also understand that the sample size in Experiment 3 of Bryan et al. was not sufficient and thus the results may require replication with a larger sample. However, I do not understand the importance of replicating their results in addressing the research question of whether and how subtle linguistic cues prevent unethical behaviors. The authors should elaborate on this issue to clarify the significance of their study (e.g., "It is important to replicate Bryan et al.'s study as it has a significant impact on this field." or "The protocol of Experiment 3 is useful and applicable to future studies. Therefore, the robustness of their findings should be tested carefully).*

**Reply:**

In the present replication study, we are planning to test whether instruction has any effect on the outcome of the experiments. Psychological researchers use instruction before an experiment to explain the aims of the experiment and what participants should do in the experiment. However, up to now it has remained unclear whether these instructions affect the performance of participants. In psychological experiments, one of the problems that affects the results is expectancy effects or participants' laziness. Laziness is a particularly frequent phenomenon, seeing participants try to finish the experiment as easily as possible just to obtain rewards. Sometimes there is even cheating. There is much evidence suggesting a relationship between reminders and behaviors in the field of social psychology (e.g., Johns, Schmader, & Martens, 2005; Ling, Beenen et al., 2005; Bryan, Master, & Walton, 2014; Bryan, Walton, Rogers, & Dweck, 2011). In particular, Bryan et al. (2013) reported for the first time that a subtle difference in instruction affects cheating. For the reasons above, the results of Bryan et al. are really important for conducting psychological research. Therefore, it is necessary to confirm the reliability of their finding. In fact, another lab has already tried to replicate Bryan et al., although they were not pre-registered (Savir, 2019).

In Experiment 1 of Bryan et al., participants were asked to think of a number from 1 to 10; if the number was even, they would be paid \$5, and otherwise, they would gain no reward. We abandoned this method because it leads to much uncertainty. First, Bryan et al. cited previous research results indicating that participants instructed to generate a random number typically show a strong bias toward odd numbers (Kubovy & Pstotka, 1976), so they intentionally paid for even numbers. However, this oddness bias has not been confirmed for betting behaviors. Secondly, thinking of a number is an imaginary event occurring in one's inside world, not a real external event. Hence, it is suboptimal as an index of falsification, as the index used for cheating should emphasize that the cheating participants' reports differ from the fact. Therefore, in our experiments,

we use the coin-tossing task in Experiments 2 and 3 of Bryan et al. as a method of measuring cheating behavior (falsifying the coin-toss result). Furthermore, compared with Experiment 2 of Bryan et al. which just consisted of two conditions, a baseline group was included in Experiment 3, which made it more complete in design. Therefore, we decided to replicate Experiment 3 of Bryan et al. in our study.

We amended the introduction to make these points clearer. Thank you for your suggestions, which have helped us clarify the purpose of the present replication.

*2. Please state in the second to last paragraph of the Introduction that, unlike Experiment 1, Bryan et al. (2013) used a coin-tossing task in Experiments 2 and 3. Otherwise, this paragraph is hard to follow.*

**Reply:**

As a response to this suggestion, we added the following sentences: "In [Bryan et al.'s] Experiments 2 and 3, they used a coin-tossing task: participants were asked to toss a coin and receive a reward corresponding to the result of their coin flips. We choose this method for our experiment because tossing a coin induces a real external event, which is more objective and operable, and hence it is better than thinking of a number to measure cheating behavior."

*3. The procedure of Experiment 1 in the current study is not described sufficiently. Please state that the cover story (examination of a paranormal phenomenon) is exactly the same as that of Experiment 3 of Bryan et al. (2013).*

*It is also necessary to specify how the authors will explain to their participants about the relationship between the coin flips and the amount of reward. Bryan et al. (2013) used the following description: "The instructions acknowledged that the laws of probability dictate that people would, on average, make \$5, although some would 'make as much as \$10 just by chance' and others would 'make as little as \$0'".*

*If the procedure is identical to that of Experiment 3 of Bryan et al. (2013) except for the use of the Japanese language, stating so would help readers understand the protocol.*

**Reply:**

In the original experiment, Bryan et al. referred to the article by Bem (2011, which had received considerable media attention) to describe a recent discovery of evidence for paranormal phenomena. We will use the same article for our instruction. As the details of the paranormal article are not given in Bryan et al.'s paper, we will abstract the content of Bem's article and translate it into Japanese and add this part to our instructions.

As for the relationship between the coin flips and reward, we will instruct participants that they will receive a reward corresponding to the result of their coin flips. In the Yahoo! Crowdsourcing Service, we gave a reward of about 5 times the average, to encourage cheating. However, the rewards that each person ultimately gets still depend on the result of their coin flips.

*4. The authors list the following two points as the reason why they plan to direct participants' attention toward the ethical reminder in Experiment 2. (a) First, there was no significant difference in the rate of cheating between "don't cheat" and the baseline (i.e., no ethical reminder) conditions in the study by Bryan et al. The authors conjecture that participants in Bryan et al.'s did not pay sufficient attention to the ethical reminder. (b) Second, Their preliminary experiment has revealed that the Japanese translation of the expression "don't be a cheater" ( ), compared to that of "don't cheat" ( ), is less familiar to Japanese speakers and thus may attract more attention. Such*

*difference in familiarity may produce results similar to Experiment 3 of Bryan et al. (2013), irrespective of whether the expression bears a strong link to self-identity or to the action. To address these concerns, the authors are planning to add “cheating (‘don’t cheat’) with task” condition, in which they test whether participants paid attention to the ethical reminder.*

*As for (a), I am not sure whether the lack of difference in cheating rates between “don’t cheat” and the baseline conditions in Bryan et al.’s study necessarily means that the participants did not pay attention to the ethical reminder. The cheating rate was significantly lower in “don’t be a cheater” condition than in the “don’t cheat” condition in their study. Doesn’t this mean that the participants paid sufficient attention to the ethical reminder?*

*As for (b), I agree with the authors that ethical reminders that attract more attention may prevent unethical behaviors more strongly. However, I do not understand what hypothesis the authors will test using the protocol of Experiment 2. Are they trying to show that it is attention to ethical reminders, rather than the linguistic expression linked to self-identity, that prevents unethical behaviors? If so, how do they interpret the results of Bryan et al. (2013)?*

**Reply:**

Thank you for your comments. We will answer your questions one by one.

*As for (a), I am not sure whether the lack of difference in cheating rates between “don’t cheat” and the baseline conditions in Bryan et al.’s study necessarily means that the participants did not pay attention to the ethical reminder. The cheating rate was significantly lower in “don’t be a cheater” condition than in the “don’t cheat” condition in their study. Doesn’t this mean that the participants paid sufficient attention to the ethical reminder?*

Possibly our description in the draft was not clear enough. In fact, we did not discuss participants’ attention in the original experiment. However, in our replication experiment, we realized that the attention of the participants to the reminder may become a problem, and aimed to explore this. As your comment reflects, the existing evidence is not sufficient to address this problem, nor, of course, can our experiments clarify the problem of attention among the participants in the original experiment. Based on result of our Experiment 2, participants’ attention to reminders in the original experiment will be discussed in the *General discussion* section of the paper (which is not included in the Stage 1 manuscript). Hence, we removed reason (a) in our draft and instead described the reason why our Experiment 2 had to highlight the issue of attention.

We made the following changes to the manuscript: “When we translated the instruction into Japanese, we felt the unfamiliarity of a “cheater” condition in a Japanese language situation. Participants in our experiment may find that the reminder “not to be a cheater” captures extra attention because of this sense of deviation. Therefore, even if the result of the original experiment is completely reproduced in our Experiment 1, it will not fully support the finding of the original experiment, as the reason for the possible different dishonest behavior rates between the “cheating” and “cheater” conditions in our Experiment 1 may be that the participants in the “cheating” group paid relatively less attention to the instruction, so that “cheating” weakly worked as a moral reminder in this condition. Because the experiments are conducted online, it is difficult to ensure that the participants have actually seen and understood the instruction; in addition, it is also possible that the participants ignored the instruction of our Experiment 1 due to satisficing, (e.g., Chandler et al., 2014; Oppenheimer et al., 2009; Sasaki & Yamada, 2019), further diminishing the effect of the unattended reminder (i.e., “cheating”).”

*As for (b), I agree with the authors that ethical reminders that attract more attention may prevent*

*unethical behaviors more strongly. However, I do not understand what hypothesis the authors will test using the protocol of Experiment 2. Are they trying to show that it is attention to ethical reminders, rather than the linguistic expression linked to self-identity, that prevents unethical behaviors? If so, how do they interpret the results of Bryan et al. (2013)?*

We will mainly discuss reason (b). We hypothesized that there is a difference in the dishonest behavior rate between the “cheating” with task condition and baseline in Experiment 2 based on the premise of the results of the preliminary experiment (i.e., the familiar expressions used in the two reminders are certainly different).

In order to convey the relationship between the preliminary experiment and the hypothesis more clearly, we made the following amendments to the manuscript: “Our Experiment 2 will only be performed when the results of Experiment 1 successfully replicate those of the original experiment. We will conduct Experiment 2, adding a “cheating” condition in which we use tasks concerning an instruction to ensure that participants’ attention is captured.” In this way, we want to find out if our experiments really support (or do not support) the results of the original experiment.

*5, I do not understand the task of the “cheating with task” condition in Experiment 2. Will the authors use this task to classify participants into those who paid attention to the ethical reminder and those who did not, and analyze the cheating rate separately for these two groups? Or is the task just for bringing participants’ attention to the ethical reminder before the coin-tossing task?*

**Reply:**

We intend to use this task only to make sure that participants pay enough attention to the ethical reminder. We will delete the data of participants who fail to give the right answer to the attention task.

**References**

- Bem, D. J. (2011). Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of personality and social psychology*, 100(3), 407.
- Bryan, C. J., Master, A., & Walton, G. M. (2014). “Helping” versus “being a helper”: Invoking the self to increase helping in young children. *Child Development*, 85(5), 1836-1842.
- Bryan, C. J., Walton, G. M., Rogers, T., & Dweck, C. S. (2011). Motivating voter turnout by invoking the self. *Proceedings of the National Academy of Sciences*, 108(31), 12653-12656.
- Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women’s math performance. *Psychological Science*, 16(3), 175-179.
- Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Li, X., ... & Resnick, P. (2005). Using social psychology to motivate contributions to online communities. *Journal of ComputerMediated Communication*, 10(4), 00-00.
- Savir, T., & Gamliel, E. (2019). To be an honest person or not to be a cheater: Replicating the effect of messages relating to the self on unethical behaviour. *International Journal of Psychology*, 54(5), 650-658.

In this way, your thoughtful and constructive comments have greatly improved our manuscript. We express our sincere gratitude again. We would appreciate it if you approved this manuscript.

**Competing Interests:** No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**