STUDY PROTOCOL

Stage 1 Registered Report: How subtle linguistic cues prevent unethical behaviors [version 1; peer review: 1 approved with reservations]

Wen Guo ¹, Huanxu Liu ¹, Jingwen Yang ¹, Yuqi Mo¹, Can Zhong ¹, Yuki Yamada ²

¹Graduate School of Human-Environment Studies, Kyushu University, Fukuoka, Japan

²Faculty of Arts and Science, Kyushu University, Fukuoka, Japan

V1 **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 evaluation and behaviors. A previous study suggested that subtle linguistic differences in ethical reminder instructions can cause differences in preventing readers' unethical behavior. The present study aims to replicate the previous finding by Bryan and his colleagues (2013) in the Japanese context, additionally exploring the influence of unfamiliar instructions 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," indicating the amount of money that could be earned. We will manipulate instructions ("Don't cheat" vs. "Don't be a cheater" vs. no instruction 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

Open Peer Review

Approval Status

	1	2
version 4 (revision) 12 Mar 2020		
version 3 (revision) 02 Mar 2020		view
version 2 (revision) 24 Dec 2019	view	view
version 1 22 Aug 2019	view	

1. **Michiko Asano** , Rikkyo University,
Saitama, Japan

2. **Sergio Cervera-Torres** , 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.

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

This article is included in the **Social Psychology** gateway.

Corresponding author: Yuki Yamada (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: © 2019 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 1; peer review: 1 approved with reservations]** F1000Research 2019, 8:1482 <https://doi.org/10.12688/f1000research.20183.1>

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

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 he or she is the kind of person who behaves dishonestly. According to self-concept maintenance theory, individuals strive to create and maintain an image of themselves as good and ethical people ([Markus & Wurf, 1987](#); [Mazar et al., 2008](#)).

Meanwhile, different ways of description can easily influence people's evaluation and judgment about something, even though 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 the 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 is 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](#)).

When a word description is linked with unethical behaviors and used to categorize who someone is, people act more conservatively. [Bryan et al. \(2013\)](#) study showed that changing the term from "Don't cheat" to "Don't be a cheater" can decrease unethical behavior. They manipulated an instruction (e.g., "Please don't be a cheater" or "Please don't cheat") to inhibit participants from conducting unethical behaviors. As a result, the "Don't be a cheater" group had a significant decrease in unethical cheating behaviors. In another experiment, [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 backfire 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 he/she fails. 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 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 all kinds of strategies to weaken the behavior-identity link ([Bandura, 1999](#)), including "euphemistic labeling" regarding language.

According to [Bryan et al. \(2011\)](#), the effect of noun expression is a motivation-driven process. When the noun is involved with positive identities, such as "voter" and "helper," people produce more correlated behaviors; when the noun is involved with negative identities such as "cheater," however, people produce fewer correlated behaviors.

In general, we believe that highlighting a self-identity word will decrease unethical behaviors; for example, 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 they are as people. However, one study revealed that highly constructed self-identities are associated with more unethical behaviors ([Cojuharencu et al., 2012](#)). There are therefore still many unsolved problems about the relationship between the concept of self-identity and unethical behaviors.

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

First, the participants of Experiment 1 in [Bryan et al. \(2013\)](#) were asked to think of a number from 1 to 10. If the number was even, they would be paid \$5; if it was odd, there was no reward. The authors found that "participants who were instructed to generate a random number typically show a strong bias toward odd numbers," yet an accurate number could not be verified at the end of experiment. Thus, we abandoned the method of Experiment 1 because it contains much uncertainty. Compared with Experiment 2, which just constituted by two condition: "cheater" and "cheating", a baseline group was included in Experiment 3 which made Experiment 3 more complete in its design. Moreover, we found that the effect size of Experiment 3 was small ($f = 0.302$). With an effect size equaling 0.302 in G*Power (significance level $\alpha = 0.05$, power level $1-\beta = 0.95$), 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 decrease unethical behaviors. We anticipate that the results will remain the same as for the original article, in which utilizing the self-relevant noun "cheater" curbed cheating behaviors significantly more than the verb "cheating."

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. Online instructions will explain that a recent controversial article has claimed to report the first scientific evidence for paranormal phenomena.

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. 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,” they will receive 25 times more points than for a normal task. This reward should be large enough for some participants to conduct a cheating behavior.

Participants will be randomly assigned into the three conditions. Three conditions are posted on the crowdsourcing website, but a person who participated in one of the conditions cannot participate in the other conditions. 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.

[嘘つきにならない/嘘をつかない]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に見せてしまいます。

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: “嘘つきにならないでください/嘘をつかないでください” Which was “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 of η^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 η^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. Data collection will not exceed 810 participants. This stopping rule is set because it is difficult for us to control the number of participants to just 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 it did not adjust 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](#).

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 record on the website that will be used in the experiment. This pilot test will not include the work that “influence the outcome of each toss with their ‘properly motivated’” and measures 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 this MCT, because they may rush through the experiment and fail to complete it in good faith.

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	$f = 0.302$
t-test	"cheating" vs "cheater"	.013	96	$d = 0.71$
	"cheater" vs baseline	.004	96	$d = 0.66$
	"cheating" vs baseline	> .80	96	$d = 0.05$
	"cheating" vs "chance"	< .0005	36	$d = 0.79$
	baseline vs "chance"	< .0005	35	$d = 0.78$
	"cheater" vs "chance"	> .30	25	$d = 0.19$

Experiment 2

This experiment will be employed as an extended, conceptual replication of Experiment 3 in the original study (Bryan *et al.*, 2013). If our Experiment 1 successfully replicates the previous 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 (e.g., Folk *et al.*, 1992; Folk *et al.*, 2002). There are two main reasons why we wish to conduct Experiment 2. First, some studies have revealed that a moral reminder may be an effective way to decrease unethical behaviors (Belle & Cantarelli, 2017). While there was no difference in the rate of cheating between the baseline and the "Don't cheat" groups in Experiment 3 of Bryan *et al.* (2013), their results suggest that the ethical reminder in the "Don't cheat" condition did not work. Because those experiments were conducted online, it is difficult to ensure whether the participants had actually seen and understood the instruction. Otherwise, it is also possible that the participants ignored the original experiment's instruction due to satisficing (e.g., Chandler *et al.*, 2014; Oppenheimer *et al.*, 2009; Sasaki & Yamada, 2019). This may be the reason that there was no significant difference between the cheating and baseline conditions. The second reason to conduct Experiment 2 is that the significance of the cheater condition may occur because of excessive attention being paid to the instruction. 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 of the Japanese instruction, rather than the English instruction of the original Experiment 3.

To confirm this assumption, we conducted a preliminary experiment, 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 and native speakers evaluated their familiarity (1: not familiar to 5: very familiar) with them via an Internet survey on the Yahoo! Crowdsourcing Service. 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 manipulated the way in which participants saw 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 baseline condition, because the task will arouse more attention. While the instruction in the cheating condition will be in large red capital letters, this should have 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. Second, in the new "cheating" 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. The exclusion standard is similar to 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.

Study timeline

Currently, 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 will be taken as consent to participate from participants. Participants 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).

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.

References

- Bandura A: **Moral disengagement in the perpetration of inhumanities.** *Pers Soc Psychol Rev.* 1999; 3(3): 193–209.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Belle N, Cantarelli P: **What Causes Unethical Behavior? A Meta-Analysis to Set an Agenda for Public Administration Research.** *Public Adm Rev.* 2017; 77(3): 327–339.
[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](#)
- Cojoharencu 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, Eggh 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](#)
- 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.
- 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](#)
- Monin B, Jordan A: **The dynamic moral self: A social psychological perspective.**

Personality, Identity, and Character: Explorations in Moral Psychology. 2009; 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: **Is crowdsourcing suitable for investigating visual perception? A case of contrast sensitivity**. *Manuscript Submitted for Publication*. 2019.

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](#)

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 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)¹ 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" (嘘つきにならないで), compared to that of "don't cheat" (嘘をつかないで), 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)¹ 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 & Psotka, 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" (嘘つきにならないで), compared to that of "don't cheat" (嘘をつかないで), 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 Computer-Mediated 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

F1000Research