BRIEF REPORT

Cheaters claim they knew the answers all along

Matthew L. Stanley [1] · Alexandria R. Stone [1] · Elizabeth J. Marsh [1]

Abstract
Cheating has become commonplace in academia and beyond. Yet, almost everyone views themselves favorably, believing that they are honest, trustworthy, and of high integrity. We investigate one possible explanation for this apparent discrepancy between people's actions and their favorable self-concepts: People who cheat on tests believe that they knew the answers all along. We found consistent correlational evidence across three studies that, for those particular cases in which participants likely cheated, they were more likely to report that they knew the answers all along. Experimentally, we then found that participants were more likely to later claim that they knew the answers all along after having the opportunity to cheat to find the correct answers – relative to exposure to the correct answers without the opportunity to cheat. These findings provide new insights into relationships between memory, metacognition, and the self-concept.

Keywords Cheating · Moral psychology · Memory · Metacognition · Hindsight bias

Introduction

Cheating scandals have become commonplace in recent years. For example, in 2012, roughly 125 Harvard students were investigated for cheating in a government class; in 2014, 92 Air Force Officers were suspended for cheating on a missile exam; in 2019, parents allegedly paid for the correction of their children's SAT test answers (Operation Varsity Blues). These high-profile cases are unusual in that cheating was caught. The vast majority of cheating cases likely go uncovered and unnoticed, given that some surveys show that more than 90% of undergraduates admit to such behavior (Evans & Craig, 1990; McCabe & Trevino, 1997; McCabe, Trevino, & Butterfield, 2001; McKibban & Burdsal, 2013; Stuber-McEwen, Wiseley, & Hoggatt, 2009; Williams, Tanner, Beard, & Hale, 2012).

Despite the pervasiveness of cheating, people are strongly motivated to enhance and protect their favorable self-concepts, believing that they truly exemplify positive traits and virtues like honesty, trustworthiness, and integrity (Alicke & Sedikides, 2009; Aquino & Reed, 2002; Stanley & De Brigard, 2019; Stanley, Henne, & De Brigard, 2019a; Wojciszke, 2005). Such positive traits and virtues are particularly central to constructions of personal identity; they play a prominent role in how people define who they are, who they wish to be, and how they attempt to present themselves to others (Aquino & Reed, 2002; Stanley & De Brigard, 2019; Strohminger, Knobe, & Newman, 2017). The pervasiveness of cheating stands in stark contrast to people's widespread beliefs that they embody positive traits and virtues like honesty, trustworthiness, and integrity. Some literature suggests that a motivated forgetting mechanism explains this apparent discrepancy between people's actions and their favorable self-concepts. In other words, one possibility is that people selectively forget information about their past failures and improprieties (Kouchaki & Gino, 2016). Supporting this explanation, some research suggests that, when consumers desire a product but learn that it was produced in an unethical way (e.g., with child labor), they later show willfully ignorant memory for how that product was produced (Reczek et al., 2017). Similarly, participants who cheated on a task strategically forgot the content of an honor code meant to bring awareness to honesty standards (Shu et al., 2011). Forgetting is one way to eliminate concrete evidence challenging the possibility that we are not as morally good as we believe

✉ Matthew L. Stanley
matthew.stanley@duke.edu

[1] Department of Psychology and Neuroscience, Duke University, 417 Chapel Drive, Durham, NC 27708, USA

ourselves to be. However, the experimental evidence for this account is not clearly generalizable or consistent, as the findings do not always replicate (Stanley, Yang, & De Brigard, 2018; Huang, Stanley, & De Brigard, 2020).

We suspect that there are many different ways in which memory processes play roles in maintaining and protecting a favorable self-concept (Stanley & De Brigard, 2019), and we investigate another explanation for why people can cheat without damaging their favorable self-concepts. The explanation we investigate here involves systematic alterations in how past events are recollected and re-evaluated in light of what has transpired since the events occurred. Converging lines of evidence suggest that people often have difficulty retrospectively determining what they knew prior to acquiring new information (e.g., Fischoff, 1975; Hasher, Attig, & Alba, 1981; Hawkins & Hastie, 1990; Roese & Vohs, 2012; Wood, 1978). In seminal research on the hindsight bias (Fischoff, 1975), for example, participants read historical scenarios and answered multiple-choice questions about several possible outcomes. After being informed of the actual outcomes, participants were asked to indicate which outcome they would have guessed, if they had not been provided the correct answer. Participants were consistently biased toward reporting they would have selected the correct answer all along (see also Guilbault, Bryant, Brockway, & Posavac, 2004). Similar effects occur across a variety of domains, including: economic decisions, political elections, and sports events (Christensen-Szalanski & Willham, 1991; Guilbault, Bryant, Brockway, & Posavac, 2004; Hawkins & Hastie, 1990). Building on early work on the hindsight bias, other research has found that, after attempting to answer general knowledge questions and then receiving the correct answers to the questions, people often come to believe that they just knew the correct answers all along (i.e., a "knew-it-all-along effect"; Arnold & Lindsay, 2007; Fischoff, 1977; Hasher et al., 1981; Jacoby & Kelley, 1987; Metcalfe & Finn, 2011; Wood, 1978).

Returning to our focus on cheating, we suggest that a knew-it-all-along effect might offer one way for people to explain away their cheating behavior in a manner that protects their favorable self-concepts. That is, cheating might not be as distasteful (and might even be justified) if people believe that they already knew the answers. In this sense, cheating might be considered accidental or unnecessary, and thus, not a real threat to a favorable self-concept. Furthermore, to the extent that this illusion is protective of one's identity, the knew-it-all-along effect accompanying cheating may be larger than would be expected if people simply received the answers without cheating (as feedback, for example).

In three studies we tested the hypothesis that cheating is associated with amplified estimates of prior knowledge. These studies depended on two procedural choices. First, we conducted extensive pre-testing to identify questions from geography, history, science, and sports for which *no* participants knew the correct answers. The goal was to identify items for which participants likely cheated if they answered them correctly. Second, cheating was made possible in some conditions by providing the correct answer upside-down at the bottom of the screen, in the same way that magazine quizzes often provide answers for self-scoring. In Study 1, we found that participants reported higher estimates of prior knowledge for questions they correctly answered than for questions they did not – even though no participants correctly answered any of the questions in our pre-tests (this effect replicated in Studies 2 and 3). Studies 2 and 3 included control conditions in which participants did not have the opportunity to cheat but were exposed to the correct answers before making judgments about whether they knew the correct answers prior to the experimental session. The control conditions were included to demonstrate that cheating inflated estimates of prior knowledge, beyond what was expected from a general knew-it-all-along effect. In both Study 2 and Study 3, we found that participants assigned to the experimental condition who had the opportunity to cheat exhibited a larger knew-it-all-along effect relative to participants in the control conditions who were exposed to the correct answers but had no opportunity to cheat in the study.

All studies were pre-registered (https://osf.io/vqhwp/). For all studies we report all exclusion criteria, all conditions included, and all independent and dependent measures.

Study 1

Materials and method

Participants A total of 150 American residents with at least 50 completed HITs and an approval rating above 90% voluntarily participated in this study on Amazon's Mechanical Turk (AMT) for monetary compensation. Three participants were excluded from analyses for failing at least one of the two attention checks (see below for details); data were thus analyzed with the remaining 147 individuals (M_{age} = 36.13 years, SD = 9.89, age range = [19–63], 80 males, 66 females). To determine our sample size in Study 1, we sought to match or exceed the sample sizes of seminal research on the knew-it-all-along effect (Hasher et al., 1981; Hawkins & Hastie, 1990; Wood, 1978). In all studies reported herein, data were analyzed only after the required sample size target was met, and the sample size was determined prior to data collection in each study (see pre-registrations). In Studies 1 and 2, we recruited participants through AMT to obtain more representative samples of the US population than traditional convenience samples obtained through undergraduate participant pools. In Study 3, we collected a nationally representative sample through Lucid. The Duke University Campus Institutional

Review Board approved all procedures for all studies reported in this article.

Materials Stimuli consisted of 24 general knowledge questions that systematically differed as a function of difficulty (some of these stimuli have been used in published work from Tauber et al., 2013, and from Wang et al., 2016). The questions came from several domains, including: geography, history, science, and sports.

Pre-testing was conducted to ensure that participants on AMT were unlikely to know the answers to 12 target questions (i.e., the difficult questions) and were likely to know the answers to 12 filler questions (i.e., the easy questions). To this end, a separate sample of 100 participants (five participants were excluded for not answering all questions) were asked to produce answers to a large set of questions. We adapted instructions from a large-scale norming study on general knowledge statements (Tauber et al., 2013; the full instructions are available in the Online Supplemental Material). Briefly, participants were instructed to answer general knowledge questions and to search memory for the correct answer. Participants were also told that there would be no penalty for guessing. For example, the following question was presented to participants: *What was the first genomically sequenced fish?* The correct response to this question is the *pufferfish*. As in Tauber et al. (2013), these responses were scored with leniency for misspellings. No participants correctly answered any of the 12 difficult questions (e.g., "*Which nation was the first to ratify the United Nations charter in 1945?*"), and at least 70% of participants correctly recalled the answers to each of the 12 easy questions (e.g., "*How many legs does a spider have?*"). All 24 questions and answers are provided in the Online Supplemental Material. The easy questions were included only to help conceal the aims of the study; only the difficult questions were used to answer our research questions.

Procedure After providing informed consent, participants were told that they would be presented with many different general knowledge questions of varying difficulty, and they were instructed to answer each question by typing it in. These 24 questions were presented one at a time in a randomized order; this comprised the *testing phase* of the study. Prior to beginning the testing phase, participants were told that the correct answer to each question would be presented in small print (size 10 font, Times New Roman) and upside-down in the bottom-right corner of the screen. A screen shot of an example trial is provided in Online Supplemental Material. When answering the questions, participants were explicitly instructed *not* to look at the correct answers on the screen. They were told that it would be cheating to look at the correct answers on the screen. Participants were also explicitly told that it would be cheating to use outside resources (e.g., other people, the internet, or books). We tested participants on these

instructions with the following true/false question: *When answering the trivia questions, it is cheating to look at the correct answers on the screen or to use outside resources (e.g., other people, the internet, or books).* Participants who failed this first attention check (by responding "false") were excluded from our analyses. This attention check was meant to ensure that the included participants understood it would be cheating to look at the correct answers on the screen or to use outside resources.

After attempting to answer all 24 questions, participants completed a brief, unrelated, distractor task. In a subsequent block, participants were then presented with the same 24 questions seen in the testing phase of the experiment, one at a time in a randomized order. This time, the correct answer was presented with each question. In this phase, participants reported whether they knew the answer to each question prior to the beginning of the study (1 = *this was new to me*, 7 = *I actually knew this all along*). This prior knowledge measure was adapted from Metcalfe and Finn (2011). Participants were instructed to answer each question honestly, told that their responses would be confidential, and assured that there would be no negative repercussions no matter how they responded.

At the end, participants completed an attention check: "Do you feel that you paid attention, avoided distractions, and took the survey seriously?" They responded by selecting one of the following: (1) no, I was distracted; (2) no, I had trouble paying attention; (3) no, I didn't take the study seriously; (4) no, something else affected my participation negatively; or (5) yes. We assured participants that their responses would not affect payment or eligibility for future studies. Only those participants who selected (5) were included in the analyses. This same attention-check question has been used previously in published research (Stanley, Marsh, & Kay, 2020; Stanley, Yin, & Sinnott-Armstrong, 2019b). Upon completion, participants were monetarily compensated for their time.

Statistical analyses Data were analyzed using R with the 'lme4' software package (Bates, Maechler, Bolker, & Walker, 2015). Data were fit to a linear mixed-effects model (LMEM) with the 'optimix' optimizer. Significance for fixed effects was assessed using Satterthwaite approximations to degrees of freedom, and 95% confidence intervals (CIs) around beta-values were computed using bootstrapping (*n* simulations = 1,000).[1] The alpha level for all statistical tests was set at .05.

Results

Overall, participants correctly answered 14% of the difficult questions, suggesting that participants cheated on roughly 14% of the questions given that the expected answer rate

---

[1] 95% CIs around beta-values offer, on our view, the best available indication of effect size for LMER models.

was zero based on our pre-testing. A histogram depicting counts for the number of items on which participants responded correctly is available in Online Supplementary Material. We tested our hypothesis that, for those cases in which participants likely cheated to answer the question (i.e., answered correctly even though no participants in the pre-test answered correctly), they would be more likely to report that they actually knew the answers all along (prior to beginning the study). To this end, we computed a LMEM with the knew-it-all-along judgments as the outcome variable. Whether or not participants reported the correct answer to each of the difficult questions was modeled as a binary fixed factor (yes or no). Participant and item were included as crossed random effects (random intercepts only). For questions for which participants provided the correct answer, relative to questions for which participants did not provide the correct answer, participants were more likely to report that they knew the answer all along ($b = 3.20$, $SE = .10$, $t = 31.53$, $p < .001$, 95% CI = [2.99, 3.42]). As an exploratory analysis, the same pattern of findings was obtained when we computed the proportion of questions on which each participant cheated and the average of the knew-it-all-along judgments for each participant, and then correlated these resultant values: $r = .69$, $p < .001$. Average knew-it-all-along judgments were 4.40 for cases in which participants reported the correct answer, and average knew-it-all-along judgments were 1.28 for cases in which participants did not provide the correct answer.

Study 2

Study 1 provides correlational evidence consistent with our hypothesis that, for those particular cases in which participants likely cheated (i.e., provided the correct answer to questions that *no* participants answered correctly in the pre-test), they were more likely to report that they actually knew the answers prior to beginning the study. In Study 2, we offer experimental evidence to show that this effect is not simply a typical knew-it-all-along effect. To this end, we instituted a control condition in which participants received feedback on their answers but were *not* given the opportunity to cheat. Knew-it-all-along judgments made by participants in this control condition were compared to those made by participants in the critical experimental condition in which they had the opportunity to cheat.

Materials and method

Participants A total of 150 American residents with at least 50 completed HITs and an approval rating above 90% voluntarily participated in this study on Amazon's Mechanical Turk (AMT) for monetary compensation. No participants failed either of the attention checks, so data were analyzed with all

150 individuals ($M_{age} = 35.34$ years, $SD = 10.81$, age range = [18–70], 79 males, 69 females).

Materials The 24 questions (12 easy, 12 difficult) from Study 1 were also used in Study 2 (see Online Supplemental Material for all questions and answers).

Procedure After providing informed consent, participants were randomly assigned to one of two conditions in a between-subjects fashion. In the experimental condition (the cheating condition), the procedure was identical to the procedure experienced by all participants in Study 1. The critical differences between the experimental and control conditions were in the *testing phase*. Participants in the cheating condition were provided with the correct answers to the questions in small print and upside-down in the bottom-right corner of the screen (as in Study 1), but participants in the control condition were *not* presented with the correct answers on the same screen as the questions. Instead, participants in the control condition attempted to answer each question, and then after submitting their response, they were shown the correct answer to the question on the *following* page. Thus, participants in the control condition had no opportunity to cheat on any of the questions, but they received similar (or even more) exposure to the answers as did the participants in the cheating condition. Otherwise, the control condition was identical to the cheating condition.

Statistical analyses The same statistical software and packages used in Study 1 were also used in Study 2.

Results

Manipulation check First, a non-parametric independent-samples Mann-Whitney U test revealed that participants in the cheating condition ($n = 73$, $Mean = .17$, $SD = .31$, $Median$, $= .00$, $IQR = .17$) answered more of the difficult questions correctly than the participants in the control condition ($n = 77$, $Mean = .02$, $SD = .04$, $Median$, $= .00$, $IQR = .00$; $p < .001$). This suggests that more participants cheated in the cheating condition than in the control condition.[2] See Online Supplementary Material for a graphical depiction of the proportion of correct answers to the difficult questions as a function of condition.

Next, we attempted to replicate our findings from Study 1. To this end, for only those participants assigned to the cheating

---

[2] Based on these descriptive statistics, it is worth noting that participants in control condition almost never answered the difficult questions correctly. This accords with what we expected based on pre-testing. These descriptive statistics further suggest that participants in the control condition did not cheat by searching for the answers on the internet. We suspect this is because searching for the answers on the internet requires time and effort, and far more time and effort than just glancing at the bottom of the screen as in the cheating condition. As such, our control condition is sufficient to prevent cheating, at least in the vast majority of cases.

condition, we computed a LMEM with knew-it-all-along judgments as the outcome variable. Whether or not participants reported the correct answer to each of the difficult questions was modeled as a binary fixed factor (yes or no). For questions for which participants provided the correct answer, relative to questions for which participants did not provide the correct answer, participants were more likely to report that they knew the answer all along ($b = 2.96$, $SE = .15$, $t = 19.83$, $p < .001$, 95% CI = [2.66, 3.25]). As an exploratory analysis, the same pattern of findings was obtained when we computed the proportion of questions participants got correct (i.e., likely cheated) and the average of the knew-it-all-along judgments for each participant, and then correlated the resultant values: $r = .80$, $p < .001$. Average knew-it-all-along judgments were 4.54 for cases in which participants provided the correct answer, and average knew-it-all-along judgments were 1.30 for cases in which participants did not provide the correct answer.

Finally, we tested our second hypothesis that participants in the cheating condition would be more likely to report that they knew the answers all along (prior to beginning the study) – relative to participants in the control condition. To this end, we computed a LMEM with condition (cheating vs. control) as a binary fixed factor and with knew-it-all-along judgments as the outcome variable. Participant and item were included as crossed random effects in the model. The results indicate that participants in the cheating condition, relative to participants in the control condition, reported higher knew-it-all-along judgments ($b = .63$, $SE = .17$, $t = 3.76$, $p < .001$, 95% CI = [.33, .95]). Average knew-it-all-along judgments were 1.89 for participants in the cheating condition, and average knew-it-all-along judgments were 1.25 for participants in the control condition. Figure 1 graphically depicts these results.

Importantly, an additional study presented in the Online Supplemental Material replicates the findings from Study 2 with a different set of general knowledge questions.

Study 3

Study 3 further investigates whether cheating amplifies the knew-it-all-along effect while accounting for two possible alternative explanations that could have produced the observed effects in Study 2. One possibility is that cheaters in Study 2 consciously lied when making their judgments of prior knowledge judgments, so that the experimenters would not detect their earlier cheating and subject them to negative repercussions (e.g., not paying them, blocking them from participating in future studies). To address this possibility, we explicitly told participants in Study 3 that their estimates of prior knowledge would *not* affect their payment or their eligibility for future studies. Participants were tested to ensure they understood this instruction, and the analyses included only those participants who reported understanding that their responses
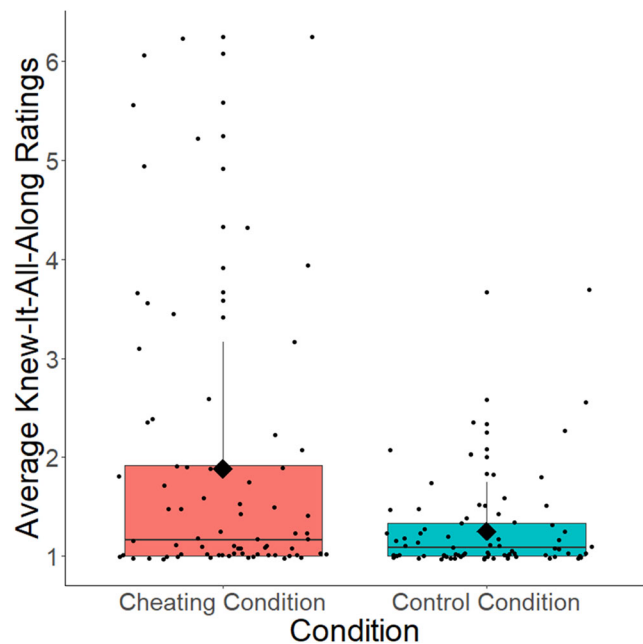

Fig. 1 Boxplots for average knew-it-all-along ratings (across the 12 difficult questions) split by condition (cheating vs. control) in Study 2. Each participant's average knew-it-all-along rating is represented by the black dots, and these dots were jittered for visualization purposes. The averages across participants of average knew-it-all-along ratings are represented with the black diamonds

would not have any negative repercussions. To preview, even after these exclusions, we still found that participants in the cheating condition exhibited higher knew-it-all-along judgments than participants in the control condition.

A second possibility is that a difficulty in source monitoring drove the effects in Study 2, as opposed to the motivational account that we hypothesized. That is, estimating one's prior knowledge requires one to discriminate what (if anything) one learned in the experiment from what one knew beforehand. The design of Study 2 may have made this decision more difficult for participants in the cheating condition than for participants in the control condition. In the cheating condition, the answers appeared on the same screen as the general knowledge questions, and cheating involved typing in the answer as when answering the easy filler questions without cheating. In contrast, the control condition used a traditional feedback design, with participants first attempting to answer the question followed by answer feedback on a separate screen. This means that the correct answer was slightly delayed in time, and when participants did not know the answer, they did not type it in. While it would still be interesting if cheating produces conditions that lead to difficult source judgments, this would be a cognitive explanation rather than a motivational explanation. To address this concern, Study 3 equated the difficulty of the source judgments across the cheating and control conditions, to rule out that explanation for our results. To preview, we still found that participants

in the cheating condition exhibited higher knew-it-all-along judgments than participants in the control condition, consistent with our hypothesized motivational account.

Materials and method

Participants A total of 603 American residents were recruited through Lucid. For a recent analysis of the Lucid platform for participant recruitment for experimental research, see Coppock and McClellan (2019). We aimed to recruit 600 participants through Lucid with the expectation that we would end up with roughly 2.5 times as many participants as in Study 2, after exclusions. 191 participants failed at least one attention check, so data were analyzed with 412 individuals (M_{age} = 48.62 years, SD = 17.31, age range = [18-82], 167 males, 235 females). The number of participants who failed each individual attention check are indicated below (note that some participants failed multiple attention checks, so adding up number of participants who failed each attention check will result in a number greater than the number of participants recruited).

Materials The 24 questions (12 easy, 12 difficult) from the previous two studies were also used in Study 3 (see Online Supplemental Material for all questions and answers).

Method After providing informed consent, participants were randomly assigned, in a between-subjects fashion, to the cheating condition or to the control condition. The cheating condition in Study 3 was the same as in Study 2. However, the control condition in Study 3 was changed so that control participants also saw the correct answers at the bottom of the screen – but critically, control participants were told they could look at the answer at the bottom of the screen if they could not generate the answer themselves. Thus, participants in both conditions saw exactly the same screens; what differed was whether they were told that looking at the answers was allowed versus prohibited. Participants in both conditions were explicitly told that it would be cheating to use outside resources (e.g., other people, the internet, or books).

We tested participants on these instructions. Participants in the cheating condition were presented with the following true/false question: *When answering the trivia questions, it is cheating to look at the correct answers on the screen or to use outside resources (e.g., other people, the internet, or books).* Participants in the cheating condition who failed this attention check (by responding "false") were excluded from our analyses (14 participants failed this attention check). Participants in the control condition were presented with the following true/false question: *When answering the trivia questions, it is not cheating to look at the correct answers on the screen, but it is cheating to use outside resources (for example: other people, the internet, or books).* Participants in the control condition who failed this attention check (by

responding "false") were excluded from our analyses (41 participants failed this attention check). Participants in both conditions received feedback about their responses. These attention checks were meant to ensure that the included participants understood what did and did not constitute cheating in the study. These attention checks were again presented immediately after participants finished the testing phase of the study. Again, participants who failed these attention checks were excluded from our analyses (19 participants in the cheating condition failed the attention check, and 19 participants in the control condition failed the attention check). And again, participants received feedback about their responses.

Participants then completed a brief, unrelated distractor task. Then, participants were instructed that they would be presented with the same 24 questions seen earlier in the experiment, one at a time in a randomized order, and that each question would be paired with its correct answer. In this phase, and as in the previous studies, participants were instructed to report whether they knew the answer to each question prior to the beginning of the study (1 = *this was new to me*, 7 = *I actually knew this all along*). We explicitly told participants that how they answered the questions would not affect their payment or eligibility for future studies, that there would be no negative repercussions regardless of how they answered, and that their responses would be confidential. Two attention-check questions were presented before participants made any prior knowledge judgments. First, participants were presented with the following true/false question: *In this next part, I am indicating whether I knew the answer to each question before I started this study.*[3] Participants were then presented with this second true/false question: *Your answers to the following questions will not affect your payment or eligibility for future studies. No matter how you answer, we will not penalize you and you will be able to participate in the future studies from our lab.* Participants who failed either of these attention checks (by responding "false") were excluded from our analyses (32 participants failed the first attention check, and 36 participants failed the second attention check). Participants received feedback about their responses to the attention-check questions.

After making prior knowledge judgments for all 24 questions, participants answered demographics questions and were presented with the final attention-check question that we provided in the previous studies. The 63 participants who failed this final attention check were also excluded from our analyses.

Statistical analyses The same statistical software and packages used the previous studies were also used in Study 3.

---

[3] Note that we added this attention-check question after we pre-registered this study. This was the only deviation from the pre-registration. We came to believe this attention check would be important for ensuring that participants understood how they were supposed to make prior knowledge judgments.

Results and discussion

Descriptively, the average proportion of difficult questions answered correctly by participants in the cheating condition was .22 ($n = 225$, $SD = .35$, $Median$, $= .00$, IQR $= .29$), and the average proportion of difficult questions answered correctly by participants in the control condition was .79 ($n = 187$, $SD = .34$, $Median$, $= 1.00$, IQR $= .25$). See Online Supplementary Material for a graphical depiction of the proportion of correct answers to the difficult questions as a function of condition.

We first attempted to replicate our finding from Studies 1 and 2 that participants reported higher estimates of prior knowledge for questions they correctly answered than for questions they did not – even though no participants correctly answered any of the questions in our pre-tests. To this end, for only those participants assigned to the cheating condition, we computed a LMEM with knew-it-all-along judgments as the outcome variable. Whether or not participants reported the correct answer to each of the difficult questions was modeled as a binary fixed factor (yes or no). For questions for which participants provided the correct answer, relative to questions for which participants did not provide the correct answer, participants were more likely to report that they knew the answer all along ($b = 2.13$, $SE = .09$, $t = 24.01$, $p < .001$, 95% CI = [1.96, 2.32]). As an exploratory analysis, for participants in the cheating condition, the same pattern of findings was obtained when we computed the proportion of questions participants got correct (i.e., likely cheated) and the average of the knew-it-all-along judgments for each participant, and then correlated the resultant values: $r = .70$, $p < .001$. For participants in the cheating condition, average knew-it-all-along judgments were 4.79 for cases in which participants provided the correct answer, and average knew-it-all-along judgments were 1.65 for cases in which participants did not provide the correct answer. In contrast, for participants in the control condition, there was no statistically significant relationship between the proportion of questions participants got correct and average knew-it-all-along judgments ($r = .06$, $p = .43$). For participants in the control condition, average knew-it-all-along judgments were 2.02 for cases in which participants provided the correct answer, and average knew-it-all-along judgments were 1.53 for cases in which participants did not provide the correct answer.

Next, we tested our second hypothesis that participants in the cheating condition would be more likely to report that they knew the answers all along (prior to beginning the study) – relative to participants in the control condition. To this end, we computed a LMEM with condition (cheating vs. control) as a binary fixed factor and with knew-it-all-along judgments as the outcome variable. Participant and item were included as crossed random effects in the model.

The results indicate that participants in the cheating condition, relative to participants in the control condition, reported higher knew-it-all-along judgments ($b = .43$, $SE = .16$, $t = 2.77$, $p = .006$, 95% CI = [.13, .74]). Average knew-it-all-along judgments were 2.35 for participants in the cheating condition, and average knew-it-all-along judgments were 1.91 for participants in the control condition. Figure 2 graphically depicts these results.

Overall, the results from Study 3 support our hypothesis that cheating amplifies the knew-it-all-along effect. The control condition instituted in Study 3 eliminated one possible alternative explanation for our results: by using the exact same procedures across conditions (other than the instructions about cheating), we equated the two conditions in source monitoring difficulty. Furthermore, our extensive instructions and the new instructional attention check in Study 3 should help to alleviate the concern that cheaters consciously lied when making prior knowledge judgments to prevent the experimenters from subjecting them to negative repercussions (e.g., not paying them, blocking them from participating in future studies). Of course, we cannot entirely rule out the possibility that cheaters consciously lied about their prior knowledge to save face with the experimenter – but we believe it is unlikely that participants were concerned with a nameless experimenter's opinion of them in a brief online study, especially since they knew their compensation was safe.
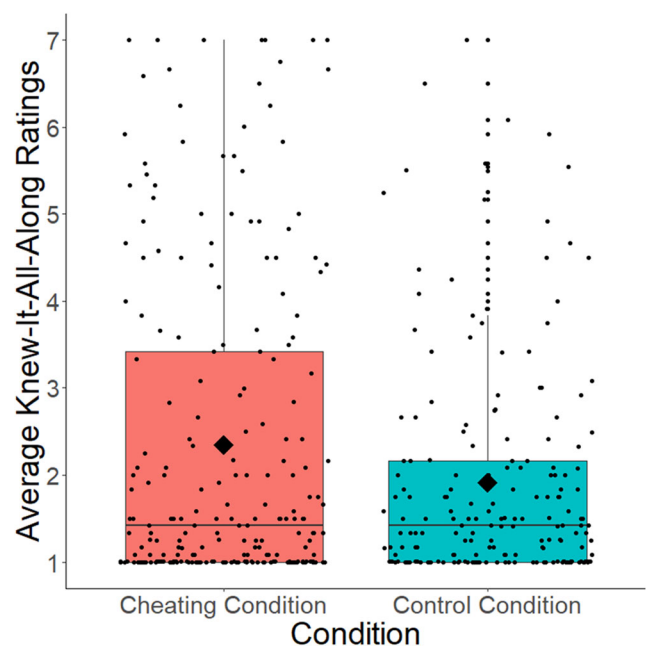

Fig. 2 Boxplots for average knew-it-all-along ratings (across the 12 difficult questions) split by condition (cheating vs. control) in Study 3. Each participant's average knew-it-all-along rating is represented by the black dots, and these dots were jittered for visualization purposes. The averages across participants of average knew-it-all-along ratings are represented with the black diamonds

General discussion

In three studies, we investigated whether cheating on a general knowledge test is associated with increased belief of having known the answers prior to the experimental session. In all three studies, we found correlational evidence that, for those particular cases in which participants provided the correct answer, they were more likely to report that they actually knew the answers all along (prior to beginning the study). In Studies 2 and 3, we found that this effect is not simply a typical knew-it-all-along effect; participants in control conditions who received the correct answers, but were not given the same opportunity to cheat, showed a significantly smaller knew-it-all-along effect than participants in the experimental condition who had the opportunity to cheat. These experimental findings are particularly surprising given that participants only cheated on a *minority* of questions in the cheating condition (see Online Supplemental Material for details); this minority was enough to obtain higher average knew-it-all-along judgments in the cheating condition relative to the control condition in which participants were provided with the correct answer to *all* questions before making knew-it-all-along judgments. Overall, our results are consistent with the idea that people are motivated to believe that they already knew the answers to the questions they cheated on. Cheating poses no threat to one's favorable self-concept if one already "knew" the information.[4]

Our findings contribute to a nascent but growing literature on how memory can support people's self-concepts and allow them to believe they are morally good despite committing moral transgressions frequently and repeatedly. Prior research has shown that people forget damaging details about their moral transgressions (Kouchaki & Gino, 2016; Reczek et al., 2017; Shu et al., 2011), and other research suggests that people recall their transgressions in a way that distances themselves from their more egregious past transgressions while perceiving moral improvement over time (Stanley et al., 2017, 2019a; Stanley & De Brigard, 2019). Our work shows a metacognitive contribution. Our findings broaden the scope of the roles that memory and metacognition play in protecting and fostering a morally good self-concept.

Our results fit well with the larger literature, lending direct evidence to a mechanism that was postulated (but not explicitly tested) in a paper investigating self-deception (Chance, Norton, Gino & Ariely, 2011; see also, Chance, Gino, Norton, & Ariely, 2015). The procedure in that paper was similar to the initial testing phase in our experiment: Participants took a general knowledge test, with a subset of participants able to view the answer key while taking the test (although looking at the

answer key was not explicitly labeled as cheating). Participants then predicted their future performance on a similar test. Critically, participants who had been given the opportunity to use an answer key on the initial test made higher estimates for their performance on the subsequent, similar test. This work did not directly test whether a knew-it-all-along effect drove the inflated estimates, but the authors posited it as a possible explanation for their results (Chance et al., 2011).

Our findings also provide insight into the relationship between the knew-it-all-along effect, hindsight bias, and the self. Considerable research has shown that the hindsight bias is *reduced* or *eliminated* following negative (as opposed to positive) outcomes that are of direct relevance to the self (Hölzl, Kirchler, & Rodler, 2002; Louie, 1999; Louie, Curren, & Harich, 2000; Mark, Boburka, Eyssell, Cohen, & Mellor, 2003; Mark & Mellor, 1991; Pezzo, 2011; Pezzo & Pezzo, 2007). For example, when an individual is laid off from a job, they might believe that they could not have possibly seen it coming (Mark & Mellor, 1991). After a negative self-relevant outcome, people who strategically come to believe that they could not have foreseen that outcome occurring can absolve themselves from blame and wrongdoing in a way that protects their favorable self-concepts (Pezzo, 2011; Roese & Vohs, 2012). In contrast, our results show an *increase* in knew-it-all-along judgments as a function of threat to a favorable self-concept (i.e., having cheated relative to not cheating). This might suggest that, at least in some cases, we can boost the knew-it-all-along effect to absolve ourselves of blame and wrongdoing in a way that protects our favorable self-concepts.

Cheating behavior has become commonplace and widespread (McKibban & Burdsal, 2013; Williams, Tanner, Beard, & Hale, 2012). Committing these transgressions can be deleterious to individuals, groups, universities, and society-at-large, so it is important to better understand why people so frequently commit them. Our findings provide evidence for a new mechanism that might help in understanding why people can so frequency commit these transgressions, even while they continue to believe they are morally good, honest, trustworthy, and of high integrity.

Open practices statement All experiments reported in this article were formally pre-registered. De-identified data from all experiments are publicly available on OSF (https://osf.io/vqhwp/). All materials for all studies are provided in the Online Supplemental Material associated with this article.

Compliance with ethical standards

Declaration of competing interests The authors declared that they have no competing interests with respect to the publication of this article.

---

[4] Because we recruited our participants from AMT and Lucid, we believe that we can make stronger generalizations about the role of cheating in boosting the knew-it-all-along effect than we could have with convenience samples comprised of undergraduate students.

References

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*(1), 1-48.

Aquino, K., & Reed, I. I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423-1440.

Arnold, M. M., & Lindsay, D. S. (2007). "I remember/know/guess that I knew it all along!": Subjective experience versus objective measures of the knew-it-all-along effect. *Memory & Cognition*, *35*(8), 1854-1868.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effect models using lme4. *Journal of Statistical Software, 67*, 1-48.

Chance, Z., Gino, F., Norton, M. I., & Ariely, D. (2015). The slow decay and quick revival of self-deception. *Frontiers in Psychology*, *6*, 1075.

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences*, *108*, 15655-15659.

Christensen-Szalanski, J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*(1), 147-168.

Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics*, *6*(1), 2053168018822174.

Evans, E. D., & Craig, D. (1990). Adolescent cognitions for academic cheating as a function of grade level and achievement status. *Journal of Adolescent Research, 5*, 325–345.

Fischoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288-299.

Fischoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 349-358.

Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic and Applied Social Psychology*, *26*(2-3), 103-117.

Hasher, L., Attig, M. S., & Alba, J. W. (1981). I knew it all along: Or, did I. *Journal of Verbal Learning and Verbal Behavior*, *20*(1), 86-96.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*(3), 311-327.

Hölzl, E., Kirchler, E., & Rodler, C. (2002). Hindsight bias in economic expectations: I knew all along what I want to hear. *Journal of Applied Psychology*, *87*(3), 437-443.

Huang, S., Stanley, M. L., & De Brigard, F. (2020). The phenomenology of remembering our moral transgressions. *Memory & Cognition.*

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin*, *13*(3), 314-336.

Kouchaki, M. & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences, 113*, 6166-6171.

Louie, T. A. (1999). Decision makers' hindsight bias after receiving favorable and unfavorable feedback. *Journal of Applied Psychology*, *84*(1), 29-41.

Louie, T. A., Curren, M. T., & Harich, K. R. (2000). "I knew we would win": Hindsight bias for favorable and unfavorable team decision outcomes. *Journal of Applied Psychology*, *85*(2), 264-272.

Mark, M. M., & Mellor, S. (1991). Effect of self-relevance of an event on hindsight bias: The foreseeability of a layoff. *Journal of Applied Psychology*, *76*(4), 569-577.

Mark, M., Reiter Boburka, R., Eyssell, K., Cohen, L., & Mellor, S. (2003). " I couldn't have seen it coming": The impact of negative self-relevant outcomes on retrospections about foreseeability. *Memory*, *11*(4-5), 443-454.

McCabe, D. L., & Trevino, L. K. (1997). Individual and contextual influences on academic dishonesty: a multicampus investigation. *Research in Higher Education, 38*, 379–396.

McCabe, D. L., Treviño, L. K., & Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research. *Ethics &Behavior*, *11*(3), 219-232.

McKibban, A. R. & Burdsal, C. A. (2013). Academic dishonesty: An in-depth investigation of assessing measurable constructs and a call for consistency in scholarship. *Journal of Academic Ethics, 11*, 185-197.

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 437-448.

Pezzo, M. V. (2011). Hindsight bias: A primer for motivational researchers. *Social and Personality Psychology Compass*, *5*(9), 665-678.

Pezzo, M. V., & Pezzo, S. P. (2007). Making sense of failure: A motivated model of hindsight bias. *Social Cognition*, *25*(1), 147-164.

Reczek, R. W., Irwin, J. R., Zane, D. M., & Ehrich, K. R. (2017). That's not how I remember it: Willfully ignorant memory for ethical product attribute information. *Journal of Consumer Research*, *45*(1), 185-207.

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science*, *7*(5), 411-426.

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin*, *37*(3), 330-349.

Stanley, M. L., & De Brigard, F. (2019). Moral memories and the belief in the good self. *Current Directions in Psychological Science*, *28*(4), 387-391.

Stanley, M. L., Henne, P., & De Brigard, F. (2019a). Remembering moral and immoral actions in constructing the self. *Memory & Cognition*, *47*(3), 441-454.

Stanley, M. L., Henne, P., Iyengar, V., Sinnott-Armstrong, W., & De Brigard, F. (2017). I'm not the person I used to be: The self and autobiographical memories of immoral actions. *Journal of Experimental Psychology: General*, *146*(6), 884-895.

Stanley, M. L., Marsh, E. J., & Kay, A. C. (2020). Structure-seeking as a psychological antecedent of beliefs about morality. *Journal of Experimental Psychology: General.*

Stanley, M. L., Yang, B. W., & De Brigard, F. (2018). No evidence for unethical amnesia for imagined actions: A failed replication and extension. *Memory & Cognition*, *46*(5), 787-795.

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

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, *12*(4), 551-560.

Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration*, *12*(3).

Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Sitzman, D. M. (2013). General knowledge norms: Updated and expanded

from the Nelson and Narens (1980) norms. *Behavior Research Methods*, *45*(4), 1115-1143.

Wang, W. C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On known unknowns: Fluency and the neural mechanisms of illusory truth. *Journal of Cognitive Neuroscience*, *28*(5), 739-746.

Williams, S., Tanner, M., Beard, J., & Hale, G. (2012). Academic integrity on college campuses. *International Journal for Educational Integrity, 8*(1), 9–24.

Wojciszke, B. (2005). Morality and competence in person-and self-perception. *European Review of Social Psychology*, *16*(1), 155-188.

Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(2), 345-353.