Counterfactual Thinking and the First Instinct Fallacy

Justin Kruger & Derrick Wirtz

University of Illinois at Urbana-Champaign

Dale T. Miller

Stanford University

In press: *Journal of Personality and Social Psychology*

Corresponding author:

Justin Kruger

University of Illinois at Urbana-Champaign

603 East Daniel Street

Champaign, IL 61820

Telephone: 217-244-8296

E-mail: jkruger@uiuc.edu

Abstract

Most people believe that they should avoid changing their answer when taking multiple-choice tests. Virtually all research on this topic, however, suggests that this strategy is ill-founded: most answer changes are from incorrect to correct, and people who change their answers usually improve their test scores. Why do people believe in this strategy if the data so strongly refute it? We argue that the belief is in part a product of counterfactual thinking. Changing an answer when one should have stuck with one's original answer leads to more "if only…" self-recriminations than does sticking with one's first instinct when one should have switched. As a consequence, instances of the former are more memorable than instances of the latter. This differential availability provides individuals with compelling (albeit illusory) personal evidence for the wisdom of always following their first instinct, with sub-optimal test scores the result.

Abstract character-count (with spaces): 953

"Exercise great caution if you decide to change an answer. Experience indicates that many students who change answers change to the wrong answer."
—Kaplan (p. 3.7, 1999)

When taking multiple-choice tests it is often the case that one answer seems correct initially, but upon further reflection another answer seems correct. In such situations is it better to switch your answer—or to stick with your "first instinct?"

Most people endorse the strategy advocated in the test-preparation guide above: As a general rule it is best to stick with one's first instinct when taking multiple-choice tests. In surveys of college students, for instance, approximately three out of four students believe that answer-changing usually lowers test scores (Mathews, 1929; Foote & Belinky, 1972; Lynch & Smith, 1975; Ballance, 1977; Smith, White, & Coop, 1979). Many college instructors hold a similar view. In one survey of faculty at Texas A&M University (including 23 from the College of Education), the majority (55%) believed that changing one's initial answer probably would lower test scores, whereas only 16% believed that answer changing would improve a student's score (Benjamin, Cavell, & Shallenberger, 1984).

The vast majority of over 70 years of research on answer-changing, however, seriously questions the validity of this belief and the utility of the "always stick with your first instinct" test-taking strategy. The majority of answer changes are from incorrect to correct, and most people who change their answers usually improve their test scores (Archer & Pippert, 1962; Bath, 1967; Clark, 1962; Copeland, 1972; Crocker & Benson, 1980; Davis, 1975; Foote & Belinky, 1972; Jarrett, 1948; Johnston, 1975; Lamson, 1935; Lehman, 1928; Lowe & Crawford, 1929; Lynch & Smith, 1975; Mathews, 1929; Pascale, 1974; Range, Anderson, & Wesley, 1982; Reile & Briggs, 1952; Reiling & Taylor, 1972; Schwarz, McMorris, & DeMers, 1991; Sitton, Adams, & Anderson, 1980; Smith et al., 1979; Vidler, 1980; Vispoel, 1998). This is true regardless of whether the test is multiple-choice or true-false, achievement or aptitude, timed or un-timed, computer or pencil-and-paper. In fact, the evidence so strongly counters the belief and strategy that one comprehensive review found that in not one of 33 studies were test-takers hurt, on average, by changing their answers (Benjamin, Cavell, & Shallenberger, 1984).

Why do people believe in the strategy of always sticking with their first instinct if the data so strongly refute it? We propose that the belief can be traced (in part) to a memory bias produced by counterfactual thinking (Miller & Taylor, 1995). The frustrating self-recriminations that follow the change of a right answer to a wrong answer make these instances more memorable, and hence seemingly more common, than either those (actually more common) instances where people changed a wrong answer to a right answer or those instances where people failed to change a wrong answer to a right answer. This prediction follows from the more general finding that events preceded by actions are more easily imagined otherwise and, as a consequence, generate stronger affect than events preceded by inactions (Kahneman & Tversky, 1982; Miller & Taylor, 1995; Miller, Turnbull & McFarland, 1990; Roese, Hur, & Pennington, 1999; Zeelenberg, van den Bos, van Dijk, & Pieters, 2002; but see also Gilovich & Medvec, 1994)—an effect that may be particularly pronounced in academic, as opposed to interpersonal, contexts (Mandel, 2003).

In summary, we propose that an error that results from the change of a correct answer to a wrong answer seems like an error that almost did not happen, and as such, seems like an error that should not have happened. The frustration associated with this conclusion serves to make this type of error more available in memory (Gilovich, Medvec, & Chen, 1995; Miller & Taylor, 1995) and hence seemingly more frequent than it is (cf. Jacowitz & Kahneman, 1995; Quattrone, 1982; Slovic & Lichtenstein, 1971; Tversky & Kahneman, 1974). In short, the advice provided by Kaplan in the quote that began this section is half-right: Experience *does* indicate that answer-changing is a poor strategy. But experience, in this case at least, is misleading.

As an analogy, consider the observation that one should avoid changing lines in the grocery store (or lanes on the highway), because to do so is inevitably followed by one's original line speeding up and one's new line slowing down. Are the gods punishing us for our impulsiveness? Perhaps, but another explanation is that changing lines when one shouldn't have is more frustrating and memorable than is failing to change lines when one should have. In much the same way, we argue—and for much the same reason—changing a correct answer to an incorrect answer is likely to be more frustrating and memorable than is failing to change an incorrect answer to a correct answer.

We conducted 4 studies to test these hypotheses. First, we compared the anticipated and actual outcome of sticking versus switching among a group of 1561 test-takers, to see whether people do indeed overestimate the effectiveness of sticking with their first instinct (Study 1). Studies 2 through 4 were designed to test the counterfactual thinking interpretation of this fallacy. Study 2 was designed to examine whether switching from the correct answer to an incorrect answer is more irksome than is failing to switch an incorrect answer to the correct answer. Study 3 was designed to see whether this hedonic asymmetry translates into a memory asymmetry, such that sticking with one's first instinct is remembered as being a better strategy than it in fact is. Finally, Study 4 was designed to link the effects demonstrated in the first three studies, by testing whether the belief in the veracity of first instincts is mediated by the heightened frustration and accessibility of changing from the correct answer to an incorrect answer versus failing to change from an incorrect answer to the correct answer.

Study 1: The Eraser Study

Our first study was designed to see whether test-takers overestimate the effectiveness of sticking with their first instincts. To find out whether this was the case, we obtained the exams of 1561 students enrolled in the Fall 2000 introductory psychology course (PSYCH 100) at the University of Illinois at Urbana-Champaign. For each problem on each test, independent coders made note of any answers that had been changed by checking for eraser marks. We then compared the outcome of these changes with students' *intuitions* about the outcome of these changes, to see whether answer-changing is believed to be an unwise course of action even among the very students who benefit from it.

Method

Participants. Participants were all 1561 University of Illinois students (49% men, 51% women) who took the Fall 2000 PSYCH 100 multiple-choice midterm exam.

Procedure. Several weeks after the midterm, independent judges (after a thorough training session) recorded the number of answers on each exam that appeared to have been changed from a previous answer (or answers) by checking for eraser marks, many of which—despite 80 years of research and development in the field of eraser technology—were plainly visible. Judges also noted whether the change was from wrong to right, right to wrong, or wrong to wrong. In the case of multiple eraser marks on a single problem, judges were instructed to consider the change to be right-to-wrong if *any* of the erased answers was the correct answer. Because our predictions were that answer changes from wrong to right would outnumber changes from right to wrong, this coding protocol constituted a conservative test of our hypothesis.

Fifty-one of the test-takers were randomly selected to provide their intuitions about the outcome of this study.[1] Specifically, they were told that the exams from their class would be checked for eraser marks, and that the result of each change would be recorded (i.e., wrong-to-right, right-to-wrong, or wrong-to-wrong). They were then asked to indicate the proportion of changes that they expected in each category, as well as the proportion of answer-changers that would probably benefit from answer changing, be hurt by answer changing, or neither.

Students' general beliefs about the utility of answer changing were also assessed. Specifically, we asked:

> When taking multiple-choice tests, it is sometimes the case that one answer seems correct at first, but upon further reflection another answer seems correct. As a general rule is it better to stick with your initial "first instinct"—or to change your answer if another choice seems better? In other words, which answer is probably more likely to be correct?

Participants then endorsed one of 3 statements: (1) original answer more likely to be correct, (2) new answer more likely to be correct, or (3) neither answer more likely to be correct than the other. The order in which the question options were presented was counterbalanced across participants.

Results

Neither gender nor the order in which either the questions or question options were presented qualified the basic results of this or any of the subsequent studies reported in this manuscript, and thus will not be discussed further.

Our predictions were that the majority of answer-changes would be from wrong to right, but that students' intuitions would be just the opposite. As Table 1 reveals, our predictions were confirmed. Of the 3291 answers that were changed, 1690 (51%) of the changes were from wrong to right, 838 (25%) were from right to wrong, and 763 (23%) were from wrong to wrong, with changes from wrong to right outnumbering changes from right to wrong by a margin of over 2-to-1. Students' intuitions, in contrast, were just the opposite. As Table 1 shows, participants expected answer changes from right-to-wrong to outnumber both those from wrong to right, paired $t(50) = 1.80$, $p < .08$, $d = .25$, and those from wrong-to-wrong, paired $t(50) = 4.56$, $p < .001$, $d = .64$. Comparing students' expected patterns with the actual patterns, revealed, as predicted, that participants underestimated the proportion of changes from wrong to right, one-sample $t(50) = 7.89$, $p < .001$, and overestimated the proportion of changes from right to wrong, one-sample $t(50) = 6.54$, $p < .001$. Participants did not over- or underestimate the number of changes that were from wrong to wrong, one-sample $t(50) < 1$.

In light of these data, it should come as no surprise that test-takers who changed their answers usually improved their test score. Of the 1561 test takers, 1231 (79%) changed one or more answers. Of these, 666 (54%) were helped by answer-changing (that is, changes from wrong to right outnumbered changes from right to wrong), 233 (19%) were hurt by answer-changing (changes from right to wrong outnumbered changes from wrong to right), and 332 (27%) were neither helped nor hurt. Here, too, participants' intuitions were off the mark. As Table 2 reveals, participants underestimated the proportion of their peers that would benefit from answer-changing (predicted = 33% vs. actual = 54%, one-sample $t(50) = 9.79$, $p < .001$), and overestimated the proportion of their peers that would be hurt by going against their first instinct, predicted = 38% vs. actual = 19%, one-sample $t(50) = 7.09$, $p < .001$, with no significant difference found, nor predicted, between the anticipated and actual proportion of test-takers whose answer changing would neither improve nor reduce scores, predicted = 29% vs. actual = 27%, one-sample $t(50) < 1$.

Students' beliefs about the general effectiveness of sticking with one's first instinct revealed a similar pattern. The vast majority of participants (75%) believed that when deciding between one's original answer and another answer, one's original answer is more likely to be correct—even if the new answer seems better. Only 12 (24%), in contrast, believed that the new answer was more likely to be correct, and one individual (2%) did not think one was more likely to be correct than the other.

Discussion

The results of Study 1 corroborate previous work showing that answer changes from wrong to right outnumber changes from right to wrong, and that people who change their answers generally benefit from doing so.[2] What is more, the data suggest that test-takers—even those who themselves have benefited from answer-changing—have just the opposite intuition.

One obvious limitation of Study 1 is that erased answers are not necessarily first instincts. Although an erased answer is undeniably an answer that has been selected and then rejected, it is not necessarily the answer that was *first* selected. It is possible that an individual might, for instance, decide on one answer, but prior to selecting it mark another answer that he or she then erases in favor of the original answer. In such a case, the individual's first instinct is in fact the selected answer, not the erased answer. As well, by looking for visible eraser marks, we obviously miss those that are *in*visible, and thus our data undoubtedly underestimate the true number of answers changed by students. Although neither of these limitations can explain the gap between the observed and intuited effect of answer-changing, we believe this issue warrants further attention, which we devote to it in Study 3.

Study 2: Should I Stay or Should I Go?

Why do people overestimate the effectiveness of sticking with their first instinct? Our account is that the belief is in part a product of counterfactual thinking: Because the decision to change a correct answer to an incorrect answer leads to more "if only…" self-recriminations than does the decision *not* to change an incorrect answer, instances of the former are more memorable than are instances of the latter. As a consequence, although answer-changing generally helped the students we surveyed, they remembered it as having hurt them.

Study 2 was designed to test the first part of this assertion: that switching a correct answer to an incorrect answer is more regrettable than is failing to switch from an incorrect answer to a correct answer. Participants imagined that they were taking an important multiple-choice test and estimated how much regret they would experience after switching from the correct answer to an incorrect answer versus failing to switch from an incorrect answer to the correct answer. Despite the equivalence of the outcome, we expected the former to be associated with more regret than the latter.

Method

Participants. 23 University of Illinois students earned course credit in exchange for their participation.

Procedure. Participants read the following scenario, modeled after one employed by Kahneman & Tversky (1982):

> Imagine that you are taking a very important multiple-choice exam. On two particular problems (Problem #1 and Problem #2), you are considering changing your initial "first instinct" to another answer. You agonize over the decision, going back and forth between thinking that you are going to switch your answer and thinking that you will stick with your original answer. You ultimately decide on two different courses of action: to change your answer on Problem #1 and to stick with your original answer on Problem #2. Suppose that both decisions turn out badly: Your original answer to Problem #1 was correct and thus by switching your answer you get Problem #1 wrong, and the other answer you were considering for Problem #2 was correct and thus by sticking with your original answer you get Problem #2 wrong.[3]

Participants were then asked "which problem would you regret missing most?," "which problem would be more likely to cause you to feel foolish for missing?," and "which problem would be more likely to cause you to think that you 'should have known better?'" After each question, participants checked one of three options: (1) Problem 1, (2) Problem 2, or (3) Neither.

Results & Discussion

Our prediction was that participants would expect to experience more regret when switching from the correct answer to an incorrect answer than when failing to switch from an incorrect to the correct answer. As Table 3 reveals, our predictions were confirmed. The majority of participants reported that they would regret missing Problem #1 more than they would regret missing Problem #2, whereas not a single participant thought the converse, which a binomial test revealed was significant, $p < .05$. As well, participants were 5 times as likely to feel foolish for missing Problem #1 than for missing problem #2, $p < .02$, and 3 times more likely to kick themselves because they "should have known better," $p = .12$.

Study 3: The standardized test

The results of Study 2 suggest that changing a correct answer to an incorrect answer leads to more regret than does failing to change an incorrect answer to the correct answer. Our contention is that this difference in regret leads to a difference in memory; that instances in which one has gone against one's first instinct and gotten the problem wrong are more memorable than are instances in which one has stuck with one's first instinct and gotten the problem wrong. Study 3 was designed to put this assertion to a direct test.

Participants took a portion of either the Scholastic Aptitude Test (SAT) or the Graduate Record Exam (GRE) in which they were instructed to narrow their selection to two answers if they could not decide on a single answer (indicating which was their first instinct). Participants then selected a final answer for each problem and, after the test was complete, received detailed performance feedback from the experimenter. Several weeks later, participants were asked to recall the number of times they stuck with their first instinct and got the problem right versus wrong, as well as to recall the number of times they went against their first instinct and got the problem right versus wrong. If switching when one should have stuck is more memorable than failing to switch when one should have, participants should overestimate how often they got a problem wrong as a consequence of switching and underestimate how often they got a problem wrong as a consequence of sticking. To put it another way, we predicted that participants would remember the outcome of sticking with their first instinct as being *better* than it actually was, and would remember the outcome of going against their first instinct as being *worse* than it really was.

Method

Participants. 27 University of Illinois undergraduates (14 women, 13 men) earned course credit in exchange for their participation.

Procedure. Participants were given a 30-minute multiple-choice exam with items taken from either the Scholastic Aptitude Test (SAT) or the Graduate Record Exam (GRE). They were told to answer the questions just as if they were taking a real exam, with one exception: If on any problem they could not decide on a single answer, they were instructed to narrow the problem down to two answers, indicating which of the two was their first instinct. The experimenter explained that by "first instincts" he (the experimenter was always male) did not mean the answer the participant necessarily thought was most likely to be correct, but rather, the answer the participant *first* thought was correct.

Once participants completed the exam, they were instructed to go back and make a final choice on the problems they had narrowed down to two selections. The experimenter then told each participant the correct answer to each problem.

Four to six weeks after participants completed the exam, those who narrowed at least one problem down to two answers (n = 21) were contacted via email and asked to complete a follow-up questionnaire. The questionnaire reminded participants of the number of problems they had narrowed down to two answers (which ranged from 1 to 19 per participant), and asked them to estimate the number of times they stuck with their first instinct and got the problem right, the number of times they stuck with their first instinct and got the problem wrong, the number of times they went against their first instinct and got the problem right, and the number of times they went against their first instinct and got the problem wrong. All but two participants responded, yielding a final sample of 19 participants (10 women, 9 men).

Results

Our first question was whether test-takers tended to avoid going against their first instinct, and whether their test scores were hurt because of it. To find out whether this was the case, for each participant we computed the number of problems in which participants stuck with their first instinct and got the problem right, stuck with their first instinct and got the problem wrong, went against their first instinct and got the problem right, and went against their first instinct and got the problem wrong.

Consequences of Switching vs. Sticking. A 2 (stay vs. switch) X 2 (right vs. wrong) fully within-subject Analysis of Variance (ANOVA) revealed two main effects. Participants stuck with their first answer more often than they went against it, Ms = 6.21 vs. 1.26, $F(1,18) = 11.09$, $p < .001$, $\eta^2 = .38$, and got the problem wrong more often than they got it right, Ms = 4.84 vs. 2.63, $F(1,18) = 5.84$, $p = .027$, $\eta^2 = .25$. As well, the analysis revealed a significant interaction, $F(1,18) = 4.99$, $p = .038$, $\eta^2 = .22$, indicating that participants were more likely to get the problem wrong if they stuck with their first instinct than if they went against it. These data, more directly than any other research of which we are aware, show that people stick with their first instinct more often than they should.

Memory for consequences of switching vs. sticking. Our primary prediction was that participants would remember the outcome of sticking with their first instinct as being *better* than it actually was, and would remember the outcome of switching as being *worse* than it really was. To find out whether this was the case, we compared the actual and remembered number of problems answered correctly versus incorrectly in a 2 (actual vs. remembered) X 2 (stay vs. switch) X 2 (correct vs. incorrect) fully within-subject ANOVA.

This analysis revealed several effects of interest. First, there were main effects for all three independent variables: Participants remembered sticking with their first instinct (and actually stuck with their first instinct) more often than they went against it,

$F(1,18) = 8.30$, $p = .010$, $\eta^2 = .32$, and participants remembered getting, and in fact got, more problems wrong than right, $F(1,18) = 5.24$, $p = .034$, $\eta^2 = .23$. Curiously, there was also a slight tendency for participants to overestimate the number of times they narrowed a problem down to two answers, despite the fact that participants were reminded of this number just prior to completing the questionnaire, $F(1,18) = 7.68$, $p = .013$, $\eta^2 = .30$. As well, there was a significant 2 (actual vs. remembered) X 2 (stay vs. switch) interaction, $F(1,18) = 5.44$, $p = .031$, $\eta^2 = .23$, suggesting that the actual difference in the number of instances in which participants stuck with their original answer and went against it was greater than the remembered difference. Neither the 2 (actual vs. remembered) X 2 (correct vs. incorrect) interaction nor the 2 (stay vs. switch) X 2 (correct vs. incorrect) interactions were significant, Fs < 1.3, η^2s < .07.

Of greatest importance, we also obtained the expected 3-way interaction, $F(1,18) = 7.98$, $p = .011$, $\eta^2 = .31$, indicating that the rift between actual and remembered outcomes depended on whether participants stuck with their first instinct or switched. As Figure 1 reveals, whereas participants tended to underestimate how often they stuck with their first instinct and got the problem wrong by an average of 1.03 problems, $t(18) = -1.62$, $p = .122$, $\eta^2 = .37$, they *overestimated* how often they went against their first instinct and got the problem wrong by an average of 1.61 problems, $t(18) = 2.94$, $p = .009$, $\eta^2 = .68$. As a result, participants tended to remember the outcome of going against their first instinct as being *worse* than it really was, as evidenced by a marginally significant 2 (switching frequency: actual vs. remembered) X 2 (correct vs. incorrect) two-way interaction, $F(1,18) = 3.79$, $p = .067$, $\eta^2 = .17$. In contrast, participants remembered the outcome of sticking with their first instinct as being *better* than it actually was, as evidenced by a significant 2 (sticking frequency: actual vs. remembered) X 2 (correct vs. incorrect) two-way interaction, $F(1,18) = 4.45$, $p = .049$, $\eta^2 = .20$.

Discussion

The results of Study 3 suggest that getting a problem wrong as a result of going against one's first instinct is more memorable than getting a problem wrong because of failing to go against one's first instinct. As Figure 1 reveals, participants showed a marked memory bias such that they overestimated how often they switched their answer and got the problem wrong, but underestimated how often they stuck with their first instinct and got the problem wrong. As a result, sticking with one's first instinct was remembered as being a better strategy than it in fact was.

Study 4: Who Wants to Be a Millionaire?

Thus far we have shown that switching a correct answer to an incorrect answer is more regrettable (Study 2) and memorable (Study 3) than is failing to switch from an incorrect answer to the correct answer. We have also shown that people tend to overestimate the effectiveness of sticking with their first instinct (Study 1). Our contention is that the former causes the latter. That is, it is *because* switching when one should have stuck is more frustrating and memorable than is sticking when one should have switched that people come to believe that it is best to stick with one's first instinct.

Our fourth and final study was designed to test this hypothesis, using a procedure inspired by the popular television show *Who Wants to Be a Millionaire?* In the show, contestants are asked a series of trivia questions, and those who correctly answer 15 in a row win a million dollars. Along the way, contestants may use up to three "lifelines," or hints, offered by the audience, host, or by "phoning a friend." Often, contestants indicate their first instinct prior to soliciting the hint, and, when the lifeline suggests a different answer, must choose whether to go against their first instinct when selecting their "final answer."

Part of the appeal of the show, such that it is, comes from the tension produced in precisely such dilemmas. What is the contestant to do? On the one hand, the hints are often reliable, and contestants typically solicit them only when they are unsure of the correct answer to begin with. On the other hand, the regret produced by switching an answer when one should have stuck with one's original answer is enough to make the misfortune of having missed the question seem almost tragic. Indeed, it is not uncommon for the host to ask contestants to explicitly disclose their first instincts, perhaps in anticipation of the counterfactual tension it can create.

Study 4 was designed to capitalize on this feature of the show in an effort to test the link between counterfactual thinking and the belief that one should stick with one's first instinct. Participants watched a video of a contestant playing a variant of *Millionaire*. For each question, the contestant indicated his or her first instinct, and then had to decide whether to stick with his or her answer or to switch to another answer suggested by a lifeline. To enhance involvement, participants were asked to imagine that the contestant was his or her teammate, and that for each dollar he or she lost, the participant also lost. Participants watched as their teammate's (and by extension, their own) initial winnings of $12,000 dwindled down to $2,000 after a series of incorrect answers.

Exactly *how* the participants lost money, however, varied by condition: In one condition, participants lost because their teammate repeatedly *stuck* with his or her first instinct when he or she should have switched, and in the other condition participants lost because their teammate repeatedly went *against* his or her first instinct when he or she should have stuck.

We predicted, first, that switching when one should have stuck would be more frustrating and memorable than sticking when one should have switched, consistent with the results of Studies 2 & 3. We also predicted that participants would be more critical of the strategy of going against one's first instinct than with the strategy of sticking with it, consistent with the results of Study 1. Of key importance, we predicted that this criticism would be mediated by the differential frustration and availability of regrettable switches versus regrettable sticks, consistent with our causal account.

Method

Participants. 68 University of Illinois students (37 women, 31 men) participated on a volunteer basis.

Stimuli. A computerized "video" was created using Microsoft Powerpoint. The video was designed to look like it was displaying the responses of a contestant playing a computerized version of *Who Wants to be a Millionaire?* in real time. A question first appeared on the top of the screen, followed by 4 possible answers. After a brief delay, the duration of which varied from question to question in an effort to mimic the behavior of real-life contestants, the answer the contestant ostensibly chose lit up. Next, two of the remaining answer choices were eliminated, a la the "50/50" lifeline in *Millionaire*. This left only two answers, one of which was correct. A graphic indicated that the computer was waiting for the contestant to decide whether to stick with his or her first answer or to switch to the one remaining answer, which was accompanied by various sound clips from the show, including the dramatic music and the host, Regis Philbin, uttering his famous catch-phrase, "is that your *final* answer?"

Depending on condition, the contestant either chose to stick with his or her first answer or to switch to the other answer. Next, the computer revealed the correct answer. This repeated for a total of 20 questions, with a total duration of approximately 20m. In one version of the video, the contestant always stuck, and in another, he or she always switched. The outcome of the two strategies, however, was identical: in each case, the contestant answered exactly half of the questions correctly. To further ensure that the only difference between conditions was the contestant's decision whether to stick or to switch, the order in which the contestant answered correctly versus incorrectly was constant across conditions (differing orders, we reasoned, might lead to differing memories). Again, various sound effects were used in an effort to punctuate the drama of the situation (and to keep participants alert).

A corner of the screen displayed the contestant's earnings. Rather than starting from 0 and merely earning money for each correct answer, the contestant started with $12,000 and not only gained money ($500) for correct answers, but lost money for incorrect ones ($1,500). This was done, once again, to punctuate the drama of the situation.

Procedure. Participants were recruited individually by separate experimenters as part of a research methods class project. On arrival to the lab, participants were told that they were to watch a video of a contestant playing a variant of *Who Wants to Be a Millionaire?* After learning the rules of the game, the participant was asked to imagine that the contestant was his or her teammate in an actual game show, and that for every dollar he or she earned (and lost), he or she also earned (and lost). Participants' task, they were told, was simply to record the outcome of each question (i.e., correct or incorrect), and to rate how frustrated they felt on a scale from –5 (*extremely frustrated*) to +5 (*extremely pleased*).

After the video was complete, the experimenter probed the participant's memory of the contestant's pattern of outcomes by asking the participant to indicate the extent to which the contestant was helped or hurt by his or her answer changing strategy on a scale from –3 (*greatly hurt*) to 0 (*neither helped nor hurt*) to +3 (*greatly helped*). In addition, the experimenter elicited the participant's evaluation of the contestant's overall answer changing strategy on a scale from –3 (*poor strategy*) to 0 (*neither poor nor good strategy*) to +3 (*good strategy*). Finally, all participants were thanked, debriefed, and dismissed.

Results & Discussion

Our first set of analyses focused on the differences in negative affect, memory, and test-taking strategy evaluation as a function of the decision to switch when one should have stuck versus stick when one should have switched. We predicted that switching when one should have stuck would be more frustrating and memorable than sticking when one should have switched, consistent with the results of Studies 2 & 3, and also that participants would be more critical of the strategy of going against one's first instinct than with the strategy of sticking with it, consistent with the results of Study 1.

All predictions were confirmed. First, after averaging the frustration ratings across the 10 incorrectly answered questions we found that participants in the switch condition were considerably more frustrated by the contestants' incorrect answers than were participants in the stick condition, *M*s = -3.18 vs. –1.08, respectively, $t(66) = 8.65$, $p < .001$, $\eta^2 = .23$. This finding complements that of Study 2, but with a non-hypothetical measure of frustration—an important distinction in light of recent work showing the occasional difficulty people have anticipating the regret they will experience (Gilbert, Morewedge, Risen, & Wilson, 2004).

Second, participants remembered their teammate having better luck when he or she stuck with his or her first instinct than when he or she did not—despite the fact that the actual outcome was equivalent across conditions, *M*s = 0.34 vs. 1.28, respectively, t(52) = 3.43, $p = .020$, $\eta^2 = .10$. Finally, participants were more critical of their teammate's test taking strategy when he or she switched when he or she should have stuck, $M = -1.32$, than when he or she stuck when he or she should have switched, $M = -0.07$, $t(52) = 4.58$, $p = .016$, $\eta^2 = .11$. Again, this was true despite the fact that the teammate's performance was constant across conditions.

Perhaps most important, path analyses fully supported our proposed causal model, namely, that switching when one stuck produces more negative affect, which in turn produces a memory bias, which in turn gives rise to the first instinct fallacy (FIF). First, as already mentioned, there was a significant relationship between the decision of whether to switch versus stick and frustration, $\beta = .48$, $p < .001$. But we also found that affect was associated with memory, $\beta = .45$, $p = .001$, and that memory was associated with participants' overall evaluation of the wisdom of their teammate's test-taking strategy, $\beta = .56$, $p < .001$. Rounding out the path analyses, we found that the link between switching and memory ($\beta = .32$, $p = .020$) was reduced to nonsignificance when the frustration measure was entered into the regression ($\beta = .18$, $p = .209$), which a Sobel (1982) test revealed was a significant reduction, $z = 2.41$, $p = .016$. As well, we found that the link between affect and the first instinct fallacy ($\beta = .35$, $p = .020$) was also reduced to non-significance when that path's proposed mediator, memory, was entered into the regression, $\beta = .12$, $p = .348$, which a Sobel (1982) test again revealed was a significant reduction, $z = 2.67$, $p = .008$. A summary of these results is depicted in Figure 2.

Importantly, we found no evidence for any other causal paths. It might be argued that, rather than memory mediating the link between affect and the first instinct fallacy, it is affect that mediates the relationship between memory and the first-instinct fallacy. Or, perhaps rather than the first instinct fallacy being a product of biased memories, it is the belief in the first instinct fallacy that engenders the memory bias. Although we do not doubt that such paths are possible, follow-up analyses revealed that they cannot account for the results of Study 4. For instance, when we repeated the above analyses with the order of the proposed mediators reversed, we found that although there was a significant relationship between switching and memory, memory and frustration, and frustration and the first instinct fallacy, there was no evidence for mediation (specifically, neither the path between the decision to switch vs. stick and frustration nor the path between the memory bias and fallacy was reduced when the appropriate mediators were entered into the regression). We similarly failed to find evidence for mediation when testing the path from switching, affect, the first instinct fallacy, and memory, or the path from switching, memory, the first instinct fallacy, and affect. Taken together, these results provide strong support for our proposed causal account of the first instinct fallacy.

General Discussion

The present research corroborates previous work showing that people overestimate the effectiveness of sticking with their first instinct when taking multiple choice tests. Participants in Study 1 believed that test-takers who changed their answers would be hurt by doing so, when in fact they were helped. This was true despite the fact that these intuitions were provided by the very students who themselves benefited from changing their answers.

This research also provides the most direct evidence to date that individuals are loath to change their answers, even at a measurable cost to their performance. Participants in Study 3 were far more likely to stick with their first instinct than to go against it, and as a consequence, achieved lower test scores than if they had changed their answers more often.

Finally, this research is the first to provide and substantiate an explanation for this "first instinct fallacy." The key assumption in this account is that changing a correct answer to an incorrect answer engenders more "if only" self-recriminations than does failing to change an incorrect answer to a correct answer. Psychologically at least, some errors are worse than others. This asymmetry in negative affect, in turn, results in an asymmetry in memory accessibility, which provides the individual with "evidence" that sticking with one's first instinct is a better strategy than it in fact is.

Studies 2 through 4 provide strong support for this analysis. In Study 2 & 4, we found that despite the logical equivalence of the outcome, switching when one should have stuck produced more regret and frustration than did sticking when one should have switched. Follow-up research from our lab suggests that it also produces more spontaneous counterfactual thoughts. For instance, in a conceptual replication of Study 2, we asked 28 students to once again imagine that they were taking an important multiple choice test. As in that study, participants were asked to imagine that on two problems they were torn between two answers: their first instinct and an alternate answer. Both questions, they were told, they had answered incorrectly, but whereas on one problem this was because they stuck with their first instinct, on the other problem it was because they went against it. Participants were then asked to indicate which problem, if any, they would feel "more foolish for missing," as well as the reason for their answer.

As in Study 2, despite the equivalence of the outcome, considerably more participants indicated that they would feel more foolish for missing the problem in which they went against their first instinct (18 out of 28, or 64%) than the problem in which they stuck with their first instinct (9 out of 28, or 32%). (The remaining 1 participant did not feel one way or the other.) Perhaps more important, there was also a considerable difference in the spontaneous mention of counterfactuals—that is, something that did not happen that they wished had happened or something that did happen that they wished had not happened. Among those who felt that switching when one should have stuck would engender more frustration, 61% cited a counterfactual as the reason for their answer (e.g., "if only I stuck with my original answer, I would have gotten the problem right"). Not a single participant, in contrast, did so when describing the reason they felt foolish for sticking when they should have switched. These data, taken together with the results of Studies 2 and 4, support our contention that switching when one should have stuck engenders more counterfactual regret than does sticking when one should have switched.

This was true despite the fact that in Study 4, at least, the decision to stick with one's first instinct was almost certainly the wrong one. The trivia questions in that experiment were extremely difficult (e.g., "When was the first national park in a Canadian province other than Alberta or British Columbia established?"), and as such, the first instincts of the undergraduate contestant were presumably little more than guesses. If so, then the alternate answer is exactly three times more likely to be correct than the contestant's original answer.[5]

We also found that switching when one should have stuck is more memorable than is sticking when one should have switched. Participants in Study 4, for instance, remembered the outcome of their virtual teammate's game show performance as being worse when he switched when he should have stuck than when he stuck when he should have switched, despite the fact that the actual outcome was constant across conditions. Importantly, this result was not limited to participants' reactions to the test-taking behavior of someone else: Participants taking a standardized multiple-choice test in Study 3, for instance, showed an analogous bias. Specifically, they tended to overestimate the number of times they switched their original answer to an incorrect answer and tended to underestimate how often they stuck with their original answer when it was incorrect.

In yet another (unpublished) study, we also found that participants were better able to recall details such as question content and answer choices of specific questions when the test-taker switched when he or she should have stuck than when he or she stuck when he or she should have switched. Specifically, 57 participants watched a video of a student from their school participating in a mock "college bowl" in a design not unlike Study 4. That is, participants watched as a contestant ostensibly representing their school was asked a series of trivia questions. For each problem, after revealing her first instinct the contestant learned the response of a previous contestant and was given the opportunity to either stick with her first instinct or to switch to the previous contestant's answer. As in Study 4, on some problems, she stuck with her first instinct, and on others, she went against it. Also as in Study 4, sometimes this paid off and other times it did not.

Exactly one week later, participants returned to the laboratory for the second phase of the experiment in which both their question and answer recognition was tested. Specifically, participants completed a questionnaire in which 80 trivia questions were presented, half of which they were told (correctly) were asked of the contestant in the video and half of which were not. Their task was simply to indicate, for each question, whether they thought the question was or was not asked of the contestant. Next, participants completed a measure of answer recognition. Specifically, the questionnaire presented the 40 questions asked of the contestant, each one followed by 8 possible answers—half of which were presented to the contestant and half of which were not. Again, participants' task was to simply indicate whether the stimulus was present in the original video. We next computed two measures of recognition for each of the 40 questions, one a binomial index of question recognition (i.e., whether the participant correctly indicated that the question was present), and the second an index of answer recognition, which was the total number of answers correctly identified as present and absent (for a theoretical range of 0 to 8). As predicted, participants showed slightly better answer-recognition for the questions in which the contestant switched when she should have stuck ($M = 6.28$) than those in which she stuck when she should have switched ($M = 5.97$), $t(55) = 1.84$, $p = .071$, $d = .25$. The same tended to be true of question-recognition, but the difference was not statistically significant, $t(55) = 0.95$. Averaging across the two indices of recognition, however, yielded a marginally significant effect, $t(55) = 1.80$, $p = .077$, $d = .24$. Taken together, these data, along with the results of Studies 3 & 4, support our contention that switching when one should have stuck is more memorable than is sticking when one should have switched.

Perhaps most important, Study 4 linked each of these processes together. As Figure 2 shows, path analyses revealed that the belief that one should stick with one's first instinct was mediated by the differential frustration, and—in turn, accessibility—associated with switching when one should have stuck versus sticking when one should have switched. This, too, is a result corroborated by supplemental research. For instance, in a follow-up to the "college bowl" study described above, we presented 81 participants with a video in which a contestant once again answered a series of trivia questions. As in that study, the contestant had to decide whether to stick with her first instinct or to change her answer to a new answer. Also as in that study, sometimes she stuck with her first instinct and sometimes she did not, and sometimes this paid off and sometimes it did not. In addition to once again measuring participants' memory for the contestants behavior (this time with a simple frequency measure in which participants estimated how often the contestant stuck when she should have switched, switched when she should have stuck, and so on), we also measured participants affective reaction to the contestants behavior. Specifically, after each question participants rated how frustrated they were with the contestant's decision on a scale from -5 (*extremely frustrated*) to +5 (*extremely pleased*).

As expected, we observed a significant relationship between affect and memory—independent of the contestant's behavior *and* the outcome of that behavior. Specifically, the more frustrating participants found the contestant's decision to switch when one should have stuck to be, the greater the number of such questions they recalled seeing, $r = .24$, $p = .035$. Taken together, these data provide not only a sufficient explanation for the widely-held first instinct fallacy, but are among the first to document the widely assumed—but seldom tested—link between regret and memory (Taylor, 1989, 1991; Taylor & Miller, 1995).

This is not to say that process depicted in Figure 2 is the only one by which people overestimate the efficaciousness of sticking with their first instinct. For instance, although we found no link between action and memory that was not mediated by affect, it is entirely possible that there may be a direct link between action and memory as well. After all, actions tend to be more salient than inactions, which might engender a memory bias independent of affect. Relatedly, it could also be the case that people expend more cognitive effort in situations in which they decide to switch versus stick, which again might cause a memory bias unrelated to affect. Although neither of these explanations are consistent mediational results of Study 4 (in which we found no support for any alternative mediational paths), they remain viable explanations for the results of the "college bowl" studies described above. Finally, as the opening quote from Kaplan suggests, the belief is sometimes presented as official test-taking dogma, and that fact alone may explain why so many of us believe that it is best to stick with our first instincts (although one might reasonably ask why the fallacy became dogma in the first place, a question the present work may help answer).

There may be still other reasons people stick with their first instinct other than an erroneous belief in its likely success. For example, the fear of feeling foolish, even without the accompanying belief that sticking with one's first instinct will boost one's score, could render people reticent to change answers. That is, to avoid kicking themselves later test-takers could set a sub-optimally high threshold for changing answers that they suspect are wrong, as we observed in Study 3. Even though anticipatory regret would not seem by itself to explain why people overestimate the effectiveness of that strategy, it is possible that people invoke the theory in an effort to justify their reluctance to change their answer (Keith Markman, 2003, personal communication). It is also possible that changing one's answer may seem a sign of inconsistency, which individuals are generally inclined to avoid (Cialdini, 1993; Kelman & Baron, 1968).[6] In other words, although our analysis of the first instinct fallacy is decidedly information-processing in nature, there may be other, more motivational factors that play a role as well.

One implication of our biased-memory account of the first instinct fallacy is that the belief in sticking with one's first instinct ought to be strengthened as one's experience with multiple-choice tests increases. If the strategy of sticking with one's first instinct is borne of the hedonic asymmetry between regrettable switches and less regrettable failures to switches, then (paradoxically) the greater one's exposure to multiple choice tests, the better the strategy of sticking with one's first instinct is likely to seem.

Some support for this thesis came from a cross-sectional follow-up study we conducted using the belief survey used in Study 1, with test-takers drawn from either an experienced sample (Juniors and Seniors drawn from an advanced psychology course, n = 83) or a less-experienced sample (Freshman and Sophomores drawn from an introductory psychology course, n = 126). As in Study 1, we found that the vast majority (71%) believed that when deciding between one's original answer and another answer, one's original answer is more likely to be correct, compared with the 17% who believed that one's new answer is more likely to be correct. However, this difference was bigger among the experienced group (78% vs. 11%) than it was among the less-experienced group (65% vs. 21%), $X^2 = 4.43$, $df = 1$, $p < .05$. Although hardly definitive on the issue, these data are at least consistent with the notion that the first instinct fallacy persists—indeed, strengthens—in the face of mounting personal evidence to the contrary.

The most obvious implication of this research, however, is that test-takers should be warned that sticking with one's first instinct is an ill-advised strategy. Even with such exhortations, however, people may be reluctant to switch as often as they should. Students who have been explicitly instructed as to the invalidity of the first instinct theory are no more likely to change their answers—nor receive higher test scores—than test-takers who do not receive instruction (McMorris, DeMers, & Schwarz, 1987; McMorris & Weidema, 1986; Prinsell, Ramsey, & Ramsey, 1994; Schwarz, McMorris, & DeMers, 1991). The (misleading) personal evidence garnered from a lifetime of test-taking may be difficult for test-takers to overcome.

References

Archer, N. S. & Pippert, R. (1962). Don't change the answer! An expose of the perennial myth that first choices are always the correct ones. *Clearing House, 37*, 39-41.

Ballance, C. T. (1977). Students' expectations and their answer-changing behavior. *Psychological Reports, 41*, 163-166.

Bath, J. A. (1967). Mental ability of residents in a state school for dependent and neglected youth. *Psychological Reports*, 20(2), 469-470.

Benjamin, L.T., Cavell, T. A., & Shallenberger, W.R. (1984). Staying with initial answers on objective tests: Is it a myth? *Teaching of Psychology*, 11(3), 133-141

Clark, C. A. (1962). Should students change answers on objective tests? *Chicago Schools Journal, 43*, 382-385.

Cialdini, R. B. (1993). Influence: The psychology of persuasion. New York: Quill, William Morrow.

Copeland, D. A. (1972). Should chemistry students change answers on multiple-choice tests? *Journal of Chemistry Education, 49*, 258.

Crocker, L., & Benson, J. (1980). Does answer-changing affect test quality? *Measurement and Evaluation in Guidance, 12*, 233-239.

Davis, R. E. (1975). Changing examination answers: And educational myth? *Journal of Medical Education, 50*, 685-687.

Diaconis, P. & Zabell, S. L. (1988). Updating subjective probability. In D. E. Bell, H. Raiffa, et al. (Eds.), *Decision making: Descriptive, normative, and prescriptive interactions* (pp. 266-283). New York: Cambridge University Press.

Foote, R. & Belinky, C. (1972). It pays to switch? Consequences of changing answers on multiple-choice examinations. *Psychological Reports, 31*, 667-673.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science, 15*, 346-350.

Gilovich, T. & Medvec, V. H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology, 67*, 357-365.

Gilovich, T., Medvec, V. H., & Chen, S. (1995). Commission, omission, and dissonance reduction: Coping with regret in the "Monty Hall" problem. *Personality & Social Psychology Bulletin, 21*, 182-190.

Granberg, D. & Brown, T. A. (1995). The Monty Hall dilemma. *Personality & Social Psychology Bulletin, 21*, 711-723.

Ichikawa, S. (1989). The role of isomorphic schematic representation in the comprehension of counterintuitive Bayesian problems. *Journal of Mathematical Behavior,* 8(3), 269-281.

Jacowitz, K.E. & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality & Social Psychology Bulletin, 21*, 1161-1166.

Jarrett, R. F. (1948). The extra-change nature of changes in students' responses to objective test items. *Journal of General Psychology, 38*, 243-250.

Johnston, J. J. (1975). Sticking with first responses on multiple-choice exams: For better or worse? *Teaching of Psychology, 2*, 178-179.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American, 246*, 160-173.

Kelman, H. C., & Baron, R. M. (1968). Inconsistency as a psychological signal. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, P. H. Tannenbaum (Eds.). *Theories of Cognitive Consistency: A Source Book (pp. 331-336)*. Chicago: Rand McNally.

Lamson, E. E. (1935). What happens when the second judgment is recorded in a true-false test? *Journal of Educational Psychology*, 26, 223-227.

Lehman, H. C. (1928). Does it pay to change initial decisions in a true-false test? *School & Society, 28*, 456-458.

Lowe, M. L., & Crawford, C. C. (1929). First impression versus second thought in true-false tests. *Journal of Educational Psychology, 20*, 192-195.

Lynch, D. O., & Smith, B. C. (1975). Item response changes: Effects on test scores. *Measurement and Evaluation in Guidance, 7*, 220-224.

Mandel, D. R. (2003). Counterfactuals, emotions, and context. *Cognition and Emotion, 17*, 139-159.

Mathews, C. O. (1929). Erroneous first impressions on objective tests. *Journal of Educational Psychology, 20*, 280-286.

McMorris, R.F. & Weideman, A.H. (1986). Answer changing after instruction on answer changing. *Measurement & Evaluation in Counseling & Development, 19*, 93-101.

McMorris, R.F., DeMers, L. P., & Schwarz, S. P. (1987). Attitudes, behaviors, and reasons for changing responses following answer-changing instruction. *Journal of Educational Measurement, 24*, 131-143.

Miller, D. T., & Taylor, B. R. (1995). Counterfactual thought, regret, & superstition: How to avoid kicking yourself. In N. J. Roese & Olson, J. M. (Eds.), *What might have been: The social psychology of counterfactual thinking*. (pp. 305-332). Mahwah, NJ: Lawrence Erlbaum Associates.

Miller, D. T., Turnbull, W., & McFarland, C. (1990). Counterfactual thinking and social perception: Thinking about what might have been. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*. (Vol. 23, pp. 305-331).

Pascale, P.J. (1974). Changing initial answers on multiple-choice achievement tests. *Measurement & Evaluation in Guidance*, 6(4), 236-238.

Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology, 42*, 593-607.

Range, L.M., Anderson, H.N., & Wesley, A. L. (1982). Personality correlates of multiple choice answer-changing patterns. *Psychological Reports*, 51(2), 523-527.

Reile, P. J. & Briggs, L. J. (1952). Should students change their initial answers on objective-type tests? More evidence regarding an old problem. *Journal of Educational Psychology*, 43, 110-115.

Reiling, E. & Taylor, R. (1972). A new approach to the problem of changing initial responses to multiple choice questions. *Journal of Educational Measurement*, 9(1), 67-70.

Roese, N. J., Hur, T., & Pennington, G. L. (1999). Counterfactual thinking and regulatory focus: Implications for action versus inaction and sufficiency versus necessity. *Journal of Personality and Social Psychology, 77*, 1109-1120.

Schwarz, S. P., McMorris, R.F., & DeMers, L. P. (1991). Reasons for changing answers: An evaluation using personal interviews. *Journal of Educational Measurement,* 28(2), 163-171

Sitton, L.R., Adams, I.G., & Anderson, H.N. (1980). Personality correlates of students' patterns of changing answers on multiple-choice tests. *Psychological Reports,* 47(2), 655-660.

Slovic, P. & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior & Human Decision Processes*, 6(6), 649-744

Smith, M., White, K., & Coop, R. H. (1979). The effect of item type on the consequences of changing answers on multiple choice tests. *Journal of Educational Measurement*, 16 (3), 203-208.

Taylor, B. (1989). *Mental simulation and the anticipation of regret*. Unpublished master's thesis, Princeton University, Princeton, NJ.

Taylor, B. (1991). *The effects of counterfactual thought on affect, memory, and beliefs*. Unpublished doctoral dissertation, Princeton University, Princeton, NJ.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.

Vidler, D.C. (1980). Curiosity, academic performance, and class attendance. *Psychological Reports*, 47(2), 589-590.

Vispoel, W. P. (1998). Reviewing and changing answers on computer-adaptive and self-adaptive vocabulary tests. *Journal of Educational Measurement,* 35(4), 328-345.

vos Savant, M. (1990a, September 9). Ask Marilyn. *Parade, p.* 15.

vos Savant, M. (1990b, December 2). Ask Marilyn. *Parade, p.* 25.

vos Savant, M. (1991, February 17). Ask Marilyn. *Parade, p.* 12.

Zeelenberg, M., van den Bos, K., van Dijk, E., & Pieters, r. (2002). The inaction effect in the psychology of regret. *Journal of Personality and Social Psychology, 82*, 314-327.

Authors' Note

This research was supported by Research Grant 1-2-69853 from the University of Illinois Board of Trustees. We thank Heather Bauer, Erin Kelly, Maryn Comin, Effie Crambes, Desler Javier, Karen Koca, Jeff Lane, John Luburic, Steve Lowe, Nacera Mekki, Amy Nelson, Erica Portnoy, Sarah Quinlan, Sheruni Ratnabalasuriar, and Jennifer Siegel for their help in collecting and coding the data, particularly for their role in the arduous task of coding the PSYCH 100 exams (Study 1). We also thank Neal Roese and Ken Savitsky for commenting on an earlier draft of this manuscript, and Sandra Goss Lucas for her help obtaining the Psych 100 exams. Correspondence should be addressed to Justin Kruger at Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, Illinois 61820. Electronic mail may be sent to Kruger at jkruger@uiuc.edu.

Footnotes

[1] Nearly perfect random selection was possible because (1) all members of the population of interest (Fall 2000 University of Illinois Psych 100 students) were eligible for participation, (2) the 51 participants were selected via a computerized randomization procedure, and (3) no participant refused participation.

[2] As an additional test of this hypothesis, we also examined all 1250 University of Illinois PSYCH 100 exams that had been completed one year earlier in the Fall of 1999. Here, too, answer changes from wrong-to-right (1644 out of 3121 changes) outnumbered answer changes from right-to-wrong (763 out of 3121 changes) by over 2-to-1. As well, of the 973 students who changed one or more answers, the majority (63%) were helped by answer-changing, 185 (19%) were hurt by it, and 179 (18%) were neither helped nor hurt, consistent with the results of Study 1.

[3] In actuality, the problem numbers were counterbalanced, but for clarity of presentation problem 1 always refers to the problem in which the participant decides to change his or her answer.

[4] The degrees of freedom is reduced in some analyses because of missing data (not all participants answered all questions). The basic results are unchanged when participants with missing data are excluded from the analyses.

[5] For the solution to this once-controversial problem, see Diaconis & Zabell, 1986; Gilovich et al., 1995; Ichikawa, 1989; Selvin, 1975; Shaughnessy & Dick, 1991; vos Savant, 1990a, 1990b, 1991.

[6] We thank Ed Hurt and several anonymous reviewers for raising some of these possibilities.

Table 1

Predicted and actual proportion of answer changes from right to wrong, wrong to right, and wrong to wrong, Study 1.

	Predicted	Actual
Wrong to right	33%	51%
Right to wrong	42%	25%
Wrong to wrong	24%	23%

Table 2

Predicted and actual proportion of individuals who changed one or more answer that were helped by answer-switching, hurt by answer-switching, or neither, Study 1.

	Predicted	Actual
Helped by switching	33%	54%
Hurt by switching	38%	19%
Neither	29%	27%

Table 3

Anticipated reaction to switching when one should have stuck versus sticking when one should have switched, Study 2.

	switching when I should have stuck	sticking when I should have switched	neither
Which would you regret most?	74%	0%	26%
Which would make you feel foolish for missing?	61%	13%	26%
Which would make you think you ‘should have known better?	48%	17%	35%

Figure Captions

Figure 1. Actual and remembered number of problems answered correctly versus incorrectly, Study 3.

Figure 2. Path analysis between switching, frustration, memory, and the first instinct fallacy, Study 4.

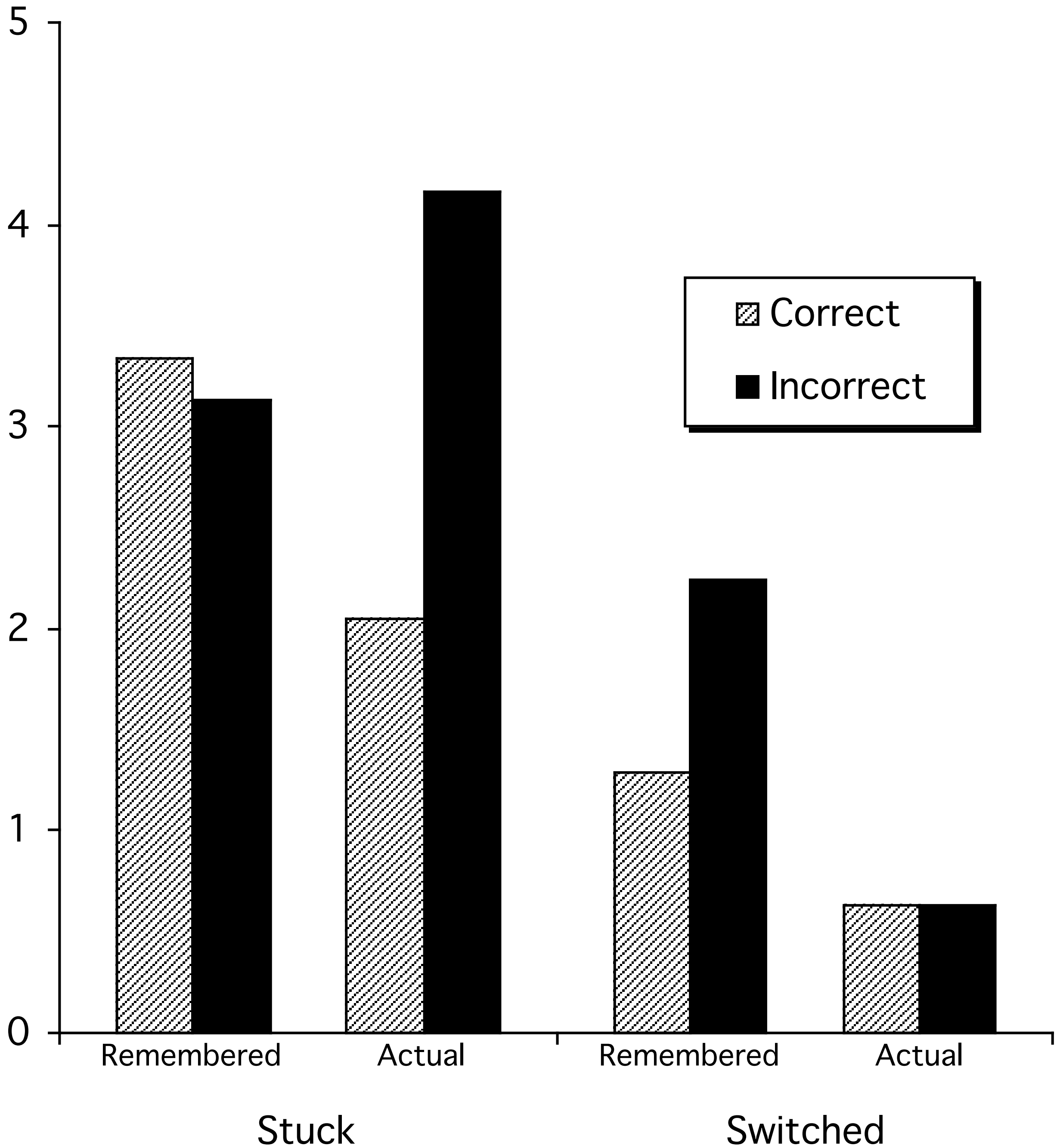

Figure 1. Actual and remembered number of problems answered correctly versus incorrectly, Study 3.

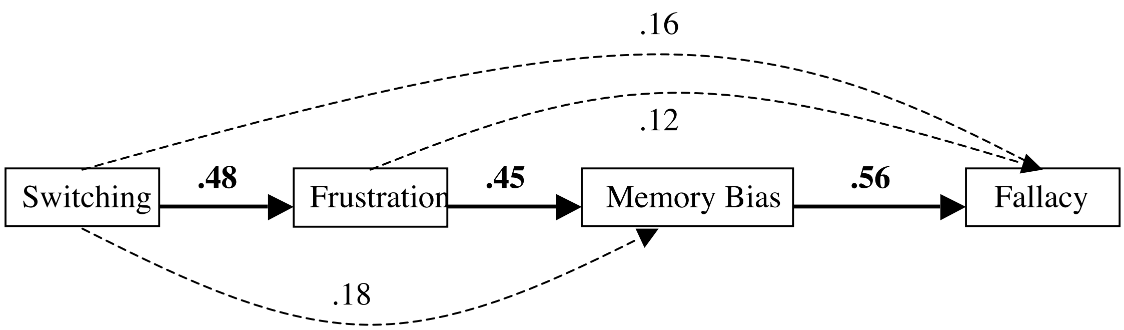

Figure 2. Path analysis between switching, frustration, memory, and the first instinct fallacy, Study 4.