Memory & Cognition
2007, 35 (8), 1854-1868

"I remember/know/guess that I knew it all along!": Subjective experience versus objective measures of the knew-it-all-along effect

Michelle M. Arnold
University of St Andrews, St Andrews, Scotland

and

D. Stephen Lindsay
University of Victoria, Victoria, British Columbia, Canada

The knew-it-all-along (KIA) effect occurs when individuals report that they previously knew something that they learned only recently. People often err when reporting the level of knowledge they had prior to feedback, but there is no research exploring the subjective experience of this effect. We incorporated a remember/just know/guess judgment into a traditional (Experiment 1A: rating scale) and a modified-traditional (Experiment 1B: two-alternative forced choice) KIA procedure. Experiments 2A, 2B, and 3 were similar in format to Experiments 1A and 1B, but the trivia stimuli were replaced with word puzzles, which were expected to be better suited to inducing a feeling of having known it all along, because answers to trivia questions typically seem arbitrary, whereas word puzzles often give rise to ah-ha experiences. A KIA effect was observed in all the experiments, but an accompanying subjective *feeling* of having known it all along arose only with word puzzles.

People often have difficulty retrospectively determining the level of knowledge they had prior to acquiring new knowledge (e.g., Fischhoff, 1975; Hasher, Attig, & Alba, 1981; Wood, 1978). Fischhoff (1977) coined the phrase *knew-it-all-along* (KIA) effect to describe this phenomenon, although it also is commonly referred to as *hindsight bias*. In a typical KIA paradigm, participants respond to a set of general knowledge questions, after which they are told the correct answers to a portion of the questions; the participants are later asked to respond to the questions with the same answers that they had given prior to being exposed to the feedback. The KIA effect occurs when participants give the correct answers to significantly more feedback questions than nonfeedback questions, indicating an overestimation of their prior knowledge (Fischhoff, 1977).

Hindsight bias has garnered a large volume of research since the mid-1970s, and it has been demonstrated in a wide variety of settings, such as relationship satisfaction (Halford & Griffith, 2002), forensic psychology (Williams, 1992), gustatory judgments (Pohl, Schwarz, Sczesny, & Stahlberg, 2003), and sporting events (Bonds-Raacke, Fryer, Nicks, & Durr, 2001). However, the KIA effect would be a good deal less interesting, and have far weaker practical implications, if it turned out that whenever participants demonstrate the effect they report that they are merely guessing or inferring their prior responses (cf. the eyewitness misinformation effect; e.g., Lindsay & Johnson, 1989; Zaragoza & Koshmider, 1989). If, in contrast, participants often report illusory beliefs or memories of having "known it all along," many interesting questions follow as to the mechanisms underlying those illusions. The present work takes some initial steps toward exploring conditions under which such illusions do versus do not arise.

There are different types of hindsight bias (i.e., based on how the phenomenon is measured), but our interest focuses on the KIA effect that arises with comparisons of foresight and hindsight judgments within subjects (commonly referred to as *memory designs*; see, e.g., Dehn & Erdfelder, 1998; Fischhoff & Beyth, 1975; Hasher et al., 1981). In such studies, as in the example given above, each participant completes the same set of judgments twice, once before and once after exposure to the correct answers. The participants are asked to complete the latter judgments with exactly the same answers as those that they had given prior to receiving the feedback.

Fischhoff (1975, 1977; Fischhoff & Beyth, 1975) proposed that the KIA effect results from an automatic assimilation of the correct feedback with preexisting knowledge (i.e., a memory impairment account; cf. Loftus, 1979). A feeling of KIA occurs because the assimilation of new information with prior knowledge eradicates the original knowledge state, making it impossible for an individual to

M. M. Arnold, michelle.arnold@st-andrews.ac.uk

recapture his/her previous level of knowledge. Because this process is automatic and immediate, it is difficult for people to comprehend the impact of new information on their perception of past knowledge, even when they are warned about the phenomenon. Fischhoff (1977) pointed to the failure of debiasing instructions (i.e., informing participants of the effect and cautioning them to avoid overestimating their previous knowledge) to reduce the hindsight bias as supporting the notion that participants lack awareness that the new information influenced their hindsight judgments.

As an alternative to the assimilation account, Jacoby and Kelley (1987) proposed an attributional approach to the KIA effect, according to which exposing participants to correct answers "spoils" their subjective experience, thereby contaminating the chief basis upon which they judge answers. For example, it is likely that feedback information is more accessible at test (e.g., due to recency) than is prior knowledge and that this accessibility leads to more fluent processing of the feedback information. Furthermore, individuals may erroneously attribute the fluently generated feedback information to prior knowledge, giving rise to a KIA effect.

Although Jacoby and Kelley (1987) did not address the issue explicitly, their attributional approach to the KIA effect raises the question of subjective experience: How do participants subjectively experience the KIA effect? Some researchers have discussed the KIA effect in terms that imply that participants have a *feeling* of having known the newly acquired knowledge in foresight (e.g., Mazursky & Ofir, 1990; Sanna, Schwarz, & Small, 2002; Stahlberg & Maas, 1998), but to date no published research has measured subjective experience in a memory design KIA procedure. Werth and Strack (2003) measured the impact that a participant's level of confidence in his/her response had on the size of the KIA effect, but they used a hypothetical design (i.e., participants answered all trivia questions in the presence of the correct answers and were told to respond as if they had not been shown the correct answers), rather than the memory design that we are interested in and have implemented in our experiments. A comprehensive analysis of hypothetical versus memory designs is beyond the scope and interest of the present article, but it is important to note that some researchers also have argued that the mechanisms involved in producing hindsight bias may depend on the type of paradigm used to measure the effect (see Schwarz & Stahlberg, 2003, for a more detailed discussion of this issue). Thus, although it is possible that KIA effects in a memory design paradigm are accompanied by a belief that the new knowledge was known prior to the feedback phase, it is also possible that participants feel as though they are merely guessing or inferring their prior answers (see also Marsh, Meade, & Roediger, 2003).

Most KIA paradigms have features that likely work against illusory feelings of having known it all along. For example, participants typically respond to a large number of items (e.g., Hell, Gigerenzer, Gauggel, Mall, & Muller, 1988; Sharpe & Adair, 1993), often responding on scales with numerous alternatives (e.g., Fischhoff, 1975; Hardt & Pohl, 2003). For instance, Goethals and Reckman (1973) had their participants rate each of 30 public policy statements (e.g., regarding the use of busing to achieve ethnic balance in schools) on a 31-point *agree/disagree* scale and indicate their confidence in each of those ratings on a 17-point scale. In the second part of the experiment, conducted a week later, a confederate attempted to change each participant's attitude on a target issue by presenting persuasive reasons for the belief opposite to that initially endorsed by the participant. Finally, the participants were asked to recall their initial ratings for 8 of the 30 original statements (including the target statement).

Goethals and Reckman (1973) found that participants' reports of their initial ratings were biased in the direction of the persuasive messages they had received. This change in rating demonstrates an *objective* KIA effect: As a matter of objective fact, there was a bias in the participants' reratings. The researchers claimed that the participants altered their reports of their past attitudes in the direction of their current attitudes because "this allows them to *feel* that the position they hold now is the one they have always held" (p. 498, italics added). However, there are no data that demonstrate that the participants did feel that the reratings they provided matched their original ratings; that is, it is possible that they had little or no confidence in the reratings they were forced to provide. Indeed, given the large number of items, the large rating scales, and the long delays, it seems unlikely that the participants in this particular study experienced illusions of remembering/knowing that they had made particular responses on the initial test. It seems more likely that the participants' subjective experience was one of guessing or inferring their initial responses.

The present experiments measured both the objective and the subjective characteristics of the KIA effect. To gauge the recollective experience of the KIA effect, we used a *remember/know/guess* (R/K/G) judgment (Gardiner & Java, 1990; Gardiner, Ramponi, & Richardson-Klavehn, 2002). If participants truly have the belief or feeling that they knew the answers in foresight, they often should respond "know" (or perhaps even "remember") on items for which they show a KIA effect. If, in contrast, participants do not have an accompanying subjective feeling of KIA, they should typically report that such answers were based on guessing.

The interpretation of remember/know (R/K) judgments depends on the underlying theoretical model of memory. However, the goal of the present research was to characterize the subjective phenomenology that accompanies the KIA effect, rather than to decide between rival models of mechanisms underlying the R/K distinction. Thus, in the present context, we are agnostic as to whether R responses reflect a qualitatively different or a quantitatively stronger memory than do K responses.[1] We will report both the raw rates of K responses (which proponents of the exclusivity assumption would view as an index of familiarity [F]) and estimates of F derived with Jacoby, Kelley, and Dywan's (1989) independence R/K (IRK) equation, in which F is calculated by dividing the K judgments by $(1 - R)$.

Our central question is whether people have a subjective feeling of remembering and/or knowing that they "knew it all along" (as opposed to feeling that they are guessing or inferring their prior knowledge states) when they demonstrate an objective KIA effect. We speculate that illusory feelings of having known it all along will be rare under the conditions of the typical KIA procedure but that, under other conditions (elaborated below), such illusions may arise.

In the first two experiments, we used standard KIA materials (trivia questions) and either a standard KIA response format (respond to questions using a number scale; Experiment 1A) or a two-alternative forced choice (2AFC) response format (Experiment 1B). In Experiments 2A and 2B, the trivia questions were replaced with word puzzles, and in the feedback phase, participants worked through the problems' solutions (rather than merely being exposed to answers). We expected that these changes would foster illusory feelings of knowing and/or remembering having given the correct answer in Test 1. Finally, Experiment 3 replicated Experiment 2B, but with a feedback-timing manipulation intended to decide between rival accounts of the different results obtained in Experiments 1B versus 2B.

EXPERIMENTS 1A AND 1B

Method

Participants. Nineteen University of Victoria students in Experiment 1A and 33 in Experiment 1B participated in exchange for optional extra credit in an introductory psychology course. The data from 8 of the 52 participants were excluded from the analyses because they apparently failed to understand the instructions for the tasks and/or the R/K/G judgment (e.g., they failed to explain the difference between R and K judgments).

Materials. A set of 100 trivia questions was constructed from various sources (e.g., Nelson & Narens, 1980). Half of the questions were critical items that were constructed to be difficult to answer (e.g., *What do you call a baby echidna?*), whereas the other 50 questions were designed to be easier to answer and were included as filler items (e.g., *Which precious gem is red?*). There were two responses assigned to each question in Experiment 1A: the correct answer and a plausible foil (e.g., *puggle* and *chuttle,* respectively, for *What do you call a baby echidna?*). The same questions and responses/foils were used in Experiment 1B, but a second plausible foil was created for each question to make the final test more difficult (i.e., both foils were presented on the final test, along with the correct response). The two foils for each question in Experiment 1B were counterbalanced (*Test 1–Test 2 foil* factor), so that each foil occurred equally often in Test 1 across participants. Furthermore, two feedback lists were constructed (*feedback list* factor) to counterbalance, between participants, which critical items were shown with feedback (i.e., the participants received feedback for one of two randomly selected sets of 25 items). A reworded trivia question was constructed from each critical item for the feedback phase, and these reworded questions always contained the answer to the critical item (e.g., *For what animal is a baby called a puggle?*).

Procedure. The participants were tested individually on an IBM-compatible personal computer using Schneider's Micro-Experimental Laboratory Professional software package (Schneider, 1988). In each phase, the experimenter read the instructions aloud. In Experiment 1A, the participants were instructed that, in Test 1, for each trial a trivia question would appear on the screen and their task was to read the question aloud; subsequently, both the correct answer and the foil for that question were displayed on the screen. The participants were told that the correct and incorrect responses would be separated vertically by a number scale ranging from 1 to 10 and that they must choose a number to indicate both their belief regarding which answer was correct and their confidence in that belief: A response of 1 or 10 was an indication that they were *absolutely sure* that the response on that end of the scale was the correct answer, whereas a response of 5 or 6 meant that they were *only guessing* that the response on that end of the scale was the correct response to the question (see Figure 1 for an example of a complete Test 1 trial for both Experiments 1A and 1B).

The Test 1 instructions given in Experiment 1B were modified to fit the 2AFC format. In Test 1, the participants were told that, after reading each question aloud, they would be shown two possible responses and that their task was to choose the response that they believed was the correct answer to the question. The participants then completed a 20-min, unrelated filler task (i.e., identifying fragmented pictures).

The feedback phase in Experiment 1A occurred immediately after the filler task. To make the feedback less obvious, the participants were informed that they were going to complete a two-part speeded reading task (SRT). In the first part of the SRT, 40 trivia questions (25 reworded critical items and 15 new filler items) were presented. For each trial, the question was presented near the bottom of the screen, and the correct answer appeared above the question; the participants were required to read both the answer and the question aloud. The 40 trivia questions then were presented two times in the second part of the SRT. The participants were instructed that on each trial, a question would appear near the top of the computer screen and that they were to read the question to themselves; they were told that once they had identified the question, they were to push a button, and the answer to the question would be presented in the center of the screen. The participants were instructed that they should say the answer as quickly as possible into the microphone and that their reaction time would be displayed on the screen.

The participants in Experiment 1B completed a feedback phase identical to that in Experiment 1A, except that their SRT occurred 24 h after Test 1. This change was intended to make Test 2 more difficult: The 2AFC format made Test 1 relatively memorable for the participants, and pilot testing indicated that it was necessary to add a delay to the procedure to avoid a ceiling effect on Test 2.

Test 2 occurred immediately after the SRT. The participants were informed that they would be presented with 50 of the trivia questions from Test 1 (to reduce the length of the testing session, only critical items were presented in Test 2) and that their task was to choose the same number (Experiment 1A) or same response (Experiment 1B) that they had given to each question in Test 1. The participants in Experiment 1B also were instructed that they would be shown three possible responses for each question: (1) the correct response that had been presented in Test 1, (2) the incorrect response that had been presented in Test 1, and (3) an incorrect response that had not been shown in Test 1. The second foil was added to Test 2 to make the task more difficult, as well as to provide a measure of consistency. The participants rarely chose the new foil on the final test, and therefore this feature will not be discussed further. In both Experiments 1A and 1B, the experimenter stressed that the researchers were interested in whether the participants could consistently select the same numbers/responses that they had chosen in Test 1 and, therefore, that it was important for the participants to ignore the SRT and concentrate on remembering the original number/response that they had given for each question in Test 1.

After choosing their Test 1 response, the participants made a remember/just-know/guess (R/JK/G) judgment; in Experiment 1A the judgment referred to the participants' memories of which side of the scale they had been on in Test 1 (e.g., whether they had chosen a number on the *puggle* or *chuttle* side of the scale), whereas in Experiment 1B it referred to their memories of selecting a response alternative (e.g., *puggle* or *chuggle*).[2] The participants were told to say "remember" if they could recollect something about making the Test 1 response and to say "just know" if they knew that they had made that response in Test 1 but could not recall anything specific about it.[3] Finally, the participants were instructed to say "guess" if they neither recollected nor knew their Test 1 response. To ensure that the participants were correctly using the R/JK/G scale, at the end of the experiment they were required to describe the three judgment options in their own words.

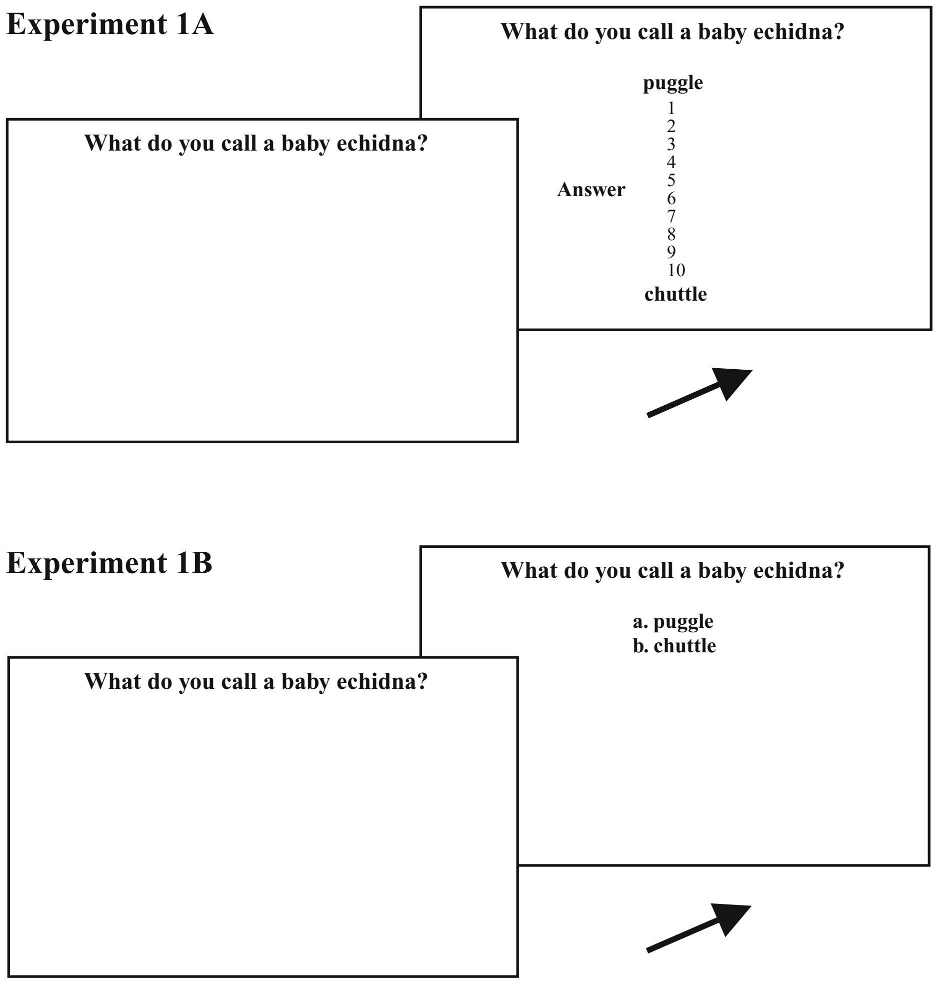


Figure 1. Example of a full trial in Test 1 of Experiments 1A and 1B.

Results and Discussion

None of the initial omnibus within-subjects ANOVAs in Experiments 1A and 1B showed significant effects of the counterbalancing factors of feedback list (Experiment 1A, all Fs ≤ 1.01, ps $\geq .33$; Experiment 1B, all Fs < 1), or Test 1–Test 2 foils (Experiment 1B, all Fs < 1), and therefore the data were collapsed across these variables.

Objective measures of the KIA effect. For comparability with Experiment 1B, responses in Experiment 1A were categorized as being on the correct or incorrect side of the rating scale. We used, as our primary measure of the KIA effect, the proportion of items answered incorrectly on Test 1 that were subsequently answered correctly on Test 2. As was expected, the participants in Experiment 1A demonstrated a KIA effect; that is, they were more likely to switch from the incorrect answer on Test 1 to the correct answer on Test 2 on feedback items ($M = .16$, $SEM = .03$) than on control items ($M = .07$, $SEM = .02$) [$F(1,15) = 9.54$, $MS_e = .01$, $p < .01$, $\eta_p^2 = .39$]. We also analyzed the proportion of items answered correctly on Test 1 for which the participants switched to giving the incorrect answer on Test 2. One might expect that exposure to correct answers during the feedback phase would reduce such switches, but in Experiment 1A the proportion of items on which the participants switched from the correct answer on Test 1 to the incorrect answer on Test 2 did not significantly differ between feedback items ($M = .15$, $SEM = .04$) and control items ($M = .10$, $SEM = .03$) [$F(1,15) = 2.15$, $MS_e = .01$, $p = .16$, $\eta_p^2 = .13$].

A KIA effect also was observed in Experiment 1B; the proportion of items on which the participants switched from the incorrect response on Test 1 to the correct response on Test 2 was higher in the feedback condition ($M = .23$, $SEM = .03$) than in the control condition ($M = .10$, $SEM = .02$) [$F(1,27) = 26.77$, $MS_e = .01$, $p < .001$, $\eta_p^2 = .50$]. Unlike in Experiment 1A, the rate of switching from correct to incorrect was significantly higher for the control items ($M = .16$, $SEM = .02$) than for the feedback items ($M = .09$, $SEM = .02$) [$F(1,27) = 5.43$, $MS_e = .01$, $p = .03$, $\eta_p^2 = .17$]. Thus, exposure to a correct answer during the SRT phase increased the likelihood that the participants who had given that answer on Test 1 also did so on Test 2. Finally, more control items were switched from correct to incorrect ($M = .16$, $SEM = .02$) than from incorrect to correct ($M = .10$, $SEM = .02$) [$F(1,27) = 5.22$, $MS_e = 0.01$, $p = .03$, $\eta_p^2 = .16$]. We

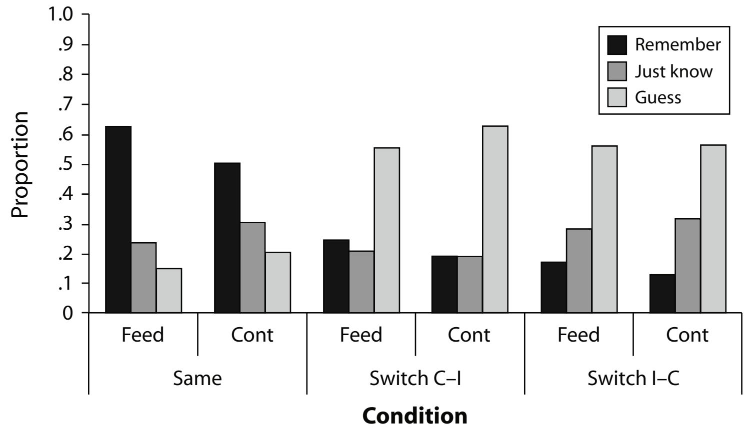


Figure 2. Remember/just-know/guess ratings for the response judgment (i.e., side of scale) in Experiment 1A for the feedback (Feed) and control (Cont) conditions (collapsed across all participants). The judgments are separated by Test 1 and Test 2 responses: (A) items given the same number on Test 1 and Test 2 (*same*); (B) items on which the participants switched from the correct response on Test 1 to the incorrect response on Test 2 (*switch C–I*); and (C) items on which they switched from the incorrect response on Test 1 to the correct response on Test 2 (*switch I–C*).

have no explanation for this effect, other than that it may be a Type I error.

Subjective measures of the KIA effect. Our central interest is in whether participants have a subjective feeling of having known it all along when they demonstrate an objective KIA effect. Figure 2 depicts the distributions of R/JK/G reports for all the participants on the feedback and control items in Experiment 1A to which the participants gave the incorrect answer on Test 1 and subsequently switched to the correct answer on Test 2 (i.e., *switch I–C* items, which correspond to our objective KIA measure). For comparison, the figure also contains the distributions of R/JK/G reports on switch C–I items and on items for which the participants responded on the same side of the scale on both tests. Figure 3 depicts those same distributions for all the participants for Experiment 1B. In both experiments, when the participants gave the same answer on Tests 1 and 2 (as they were instructed to do), they usually reported that they remembered their Test 1 response and rarely indicated that they were merely guessing. In contrast, when the participants switched from one answer to the other, they usually reported that they were guessing and rarely reported that they remembered or just knew what response they had given on Test 1. Moreover, visual inspection of Figures 2 and 3 suggests that the feedback manipulation (which led to an objective KIA effect) did not increase the likelihood that the participants would report (illusory) experiences of knowing or remembering for giving the correct answer on Test 1.

Quantitatively analyzing the R/JK/G judgment was not as straightforward as analyzing the objective data, because the subjective measure of the effect had some data cells with very few observations per participant (e.g., few R responses were given for critical items on which the participants switched from one side of the scale to the other). To alleviate this problem, we transformed the data by taking the natural log of the proportions, which resulted in more normal distributions of the data. We also added a constant of .50 to both the numerator and the denominator of the proportion equation prior to transforming the data to deal with the issue of empty R, JK, or G data cells.[4] In addition, any participants who did not have data for both the feedback and the control measures of interest were removed prior to the analyses. The inferential tests reported below—and for the subjective measures of the KIA effect in the following experiments—are based on the transformed data, but to foster clarity the accompanying means and standard error of the means are reported for the raw proportions. The means of the transformed data for the R, JK, and G ratings of the response (i.e., side of the scale) and for each experiment are shown in Appendix A (Table A1).

The transformed proportions of R/JK/G designations for the response judgment in Experiment 1A were analyzed in a 2 (item type: feedback vs. control) × 3 (judgment option: R, JK, or G) within-subjects ANOVA. The main effects of item type and judgment option are not informative (i.e., because, in terms of the raw proportions, these measures sum to 1.00), and therefore only the interaction and subsequent planned comparisons will be reported. Seven of the 19 participants were excluded from the analyses for having zero feedback and/or control items on which they switched to the correct answer on Test 2. There was no interaction between item type and R/JK/G choices for the trivia items on which participants switched from the incorrect side of the scale on Test 1 to the correct side of the scale on Test 2 ($F < 1$).

The overall proportions of items given an R, JK, or G rating for the response judgment in Experiment 1B are shown in Figure 3. The transformed proportion of R/JK/G designations for the trivia items on which the participants

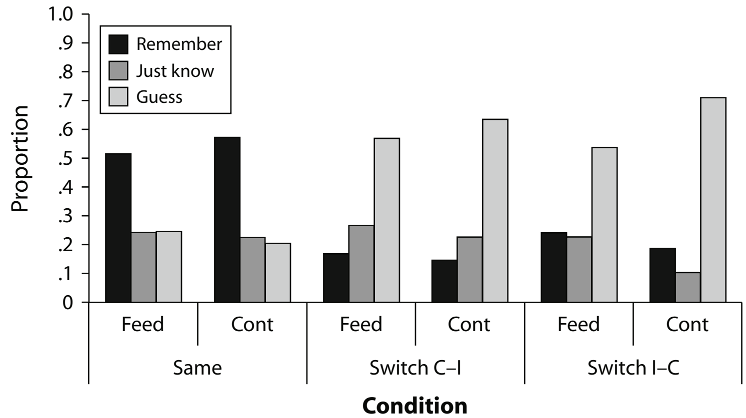


Figure 3. Remember/just-know/guess ratings for the response judgment in Experiment 1B for the feedback (Feed) and control (Cont) conditions (collapsed across all participants). The judgments are separated by Test 1 and Test 2 responses: (A) items given the same response on Test 1 and Test 2 (*same*); (B) items on which the participants switched from the correct response on Test 1 to the incorrect response on Test 2 (*switch C–I*); and (C) items on which they switched from the incorrect response on Test 1 to the correct response on Test 2 (*switch I–C*).

switched from an incorrect answer on Test 1 to a correct answer on Test 2 were analyzed in a 2 (item type: feedback vs. control) × 3 (judgment option: R, JK, or G) within-subjects ANOVA. Ten participants were excluded from the analyses for having either no feedback or no control items on which they switched to the correct answer on Test 2. There was no significant interaction between item type and judgment option ($F < 1$).

The results of Experiment 1A clearly demonstrated a typical hindsight bias, and this pattern was found both in the number scale and in the proportion of items on which the participants switched from the incorrect side of the scale on Test 1 to the correct side of the scale on Test 2. However, the R/JK/G measure did not produce any evidence that the KIA effect was accompanied by a subjective feeling of having known the feedback information in foresight. This finding was true regardless of whether responses to the trivia questions were on a numerical rating scale or in a 2AFC format. Thus, the KIA effect was not accompanied by a subjective experience in hindsight that the correct answers had been remembered or known in foresight. Rather, when the participants showed the KIA effect, they reported that they were guessing at what their initial responses had been.

EXPERIMENTS 2A AND 2B

The lack of evidence in the first set of experiments for a subjective component to the bias suggests that the KIA effect might better be called the "I guess I knew it all along" effect. Nonetheless, we speculate that, under the right circumstances, participants sometimes have an illusory subjective feeling of knowing or remembering that they knew the correct answers in foresight.

Experiments 2A and 2B were designed to explore the KIA effect with a set of stimuli and procedures that should, on theoretical grounds, be better suited to inducing a *feeling* of KIA than were the trivia items used in the first two experiments (and in most prior within-subjects KIA demonstrations). The items used in these two experiments were word puzzles—commonly referred to as "Wordies"—which can be rearranged or translated to form common words, phrases, or clichés (see Appendix B for examples of Wordies used in this experiment). One of the important differences between the trivia items in the previous experiments and the Wordies is that whereas answers to trivia questions have an arbitrary quality (e.g., there is nothing inherently correct about *puggle* being the term for a baby echidna), solving Wordies often gives rise to "ah-ha!" experiences; that is, once the solution to a Wordie is found, one sees why it is the correct answer. Indeed, once the answer to a Wordies puzzle is discovered, it may seem obvious, as per Jacoby and Kelley's (1987) explanation regarding "spoiled" subjective experiences. Furthermore, in Phase 2 of Experiments 2A and 2B, the participants were not merely exposed to correct answers but, rather, worked through the problems' solutions, which may lead to source-monitoring problems (Johnson, Hashtroudi, & Lindsay, 1993) when they later attempt to remember their Test 1 responses. Therefore, the nature of the Wordies and the tasks involved in these experiments should lend themselves better to promoting a sense of "I knew that before the feedback!" than did the materials and procedures used in the previous two experiments (and most prior studies of the KIA effect).

Method

Participants. Twenty-four University of Victoria students participated in Experiment 2A and 23 in Experiment 2B in exchange for optional extra credit in an introductory psychology course. The data from 6 of the 47 participants were excluded from the analyses because they apparently failed to understand the instructions for the tasks and/or the R/JK/G judgment.

Materials. A pool of 92 Wordies was created from a variety of sources (e.g., Brain Teasers!, n.d.), and a plausible foil was constructed for each correct solution (e.g., *marked countertop* as a foil for the solution *check-out counter*). Two of the Wordies were used as practice trials at the beginning of Tests 1 and 2, and 10 of the Wordies were used as training items during the feedback phase. The remaining 80 items were used as test items during Tests 1 and 2 (50 critical items and 30 filler items). No counterbalancing conditions were necessary because the computer program randomly selected which critical items were presented during the feedback phase for each participant.

Procedure. The participants were tested individually on an IBM-compatible personal computer using Schneider, Eschman, and Zuccolotto's (2002) E-Prime software package. In Test 1 of Experiment 2A, the participants were told that for each trial, they would see a word puzzle appear on the screen for 3 sec; the participants were informed that after 3 sec, the puzzle would disappear and two responses separated by a vertical 10-point number scale would appear on the screen. After finishing Test 1, the participants completed a 5-min embedded figures filler task.

As in Experiment 1A, the feedback phase occurred immediately after the filler task in Experiment 2A. The participants were informed that they were going to complete a two-part Wordies judgment task. In Part 1, referred to as the *understanding-of-the-solution* judgment, they were shown 40 of the word puzzles from Test 1 (25 critical items, 15 filler items). For each trial, the puzzle was shown for 10 sec, with the correct solution directly below the puzzle. The participants' task was to decide whether they understood *how* to arrive at that correct solution for the puzzle, pressing a key when they did so (if they failed to push the key during the 10 sec, the computer went on to the next trial). As training, the participants were shown several different types of puzzles and their solutions, and the experimenter talked them through how to arrive at the solutions (e.g., "The solution is 'once upon a time' because the word 'once' has physically been placed on top of the phrase '4:56 pm,' which itself stands for the more abstract concept of 'time'").

The second part of the Wordies judgment task in Experiment 2A occurred immediately after Part 1, and its purpose was to give the participants another pass at the correct solutions to the critical items. For Part 2, referred to as the *common/rare* judgment task, the participants were shown the same 40 puzzles/solutions from Part 1 for 2 sec each, and they indicated whether the puzzle solution represented a common word or phrase (i.e., something they thought they might encounter in a typical day) or whether it represented a rare word or phrase (i.e., something they did not think they would come across in everyday conversations).

Test 2 in Experiment 2A took place immediately after the Wordies judgment. The participants were told that they would be presented with the 80 puzzles from Test 1 and that their task was to choose the exact same number that they had given to each puzzle in the first test. As in Experiment 1A, after selecting their Test 1 response, the participants in Experiment 2A were required to complete two separate R/JK/G judgments (a response judgment and a number judgment; as in Experiment 1A, only data from the response judgment will be reported here [see Note 2]). Finally, as in the previous experiments, the participants were asked to describe the definitions for the R/JK/G options in their own words.

Experiment 2B included the two major modifications from Experiment 1B; that is, the test format was changed to a 2AFC task, and the experiment was conducted over 2 days. Also, to ensure that accuracy on the 2AFC test was below ceiling, the stimulus duration in Test 1 was reduced from 3 sec in Experiment 2A to 2 sec in Experiment 2B, and the participants were given only 10 sec to respond after the response alternatives appeared on the screen (whereas there was no response deadline in Experiment 2A). Changes also were made to the feedback phase of Experiment 2B. First, the participants completed the feedback phase on the 1st day of the experiment (unlike in Experiment 1B). Second, Part 2 of the Wordies judgment feedback phase (the common/rare judgment in Experiment 2A) was a solution recall task; after completing the understanding-of-the-solution judgment (Part 1), the participants were shown the puzzles again and were asked to free-recall the solutions to each puzzle (Part 2: recall-of-the-solutions task).

Results and Discussion

Objective measures of the KIA effect. As in Experiment 1A, responses in Experiment 2A were categorized as being on the correct or incorrect side of the rating scale. As was anticipated, a KIA effect was found for switching to the correct answers, in that the participants were more likely to change to the correct answer on Test 2 for the feedback items (M = .42, SEM = .08) than for the control items (M = .19, SEM = .05) [$F(1,14) = 6.52$, $MS_e = .06$, $p = .02$, $\eta_p^2 = .32$]. There was no significant difference in the proportion of items on which the participants switched from the correct response on Test 1 to the incorrect response on Test 2 for the feedback (M = .05, SEM = .01) versus the control items (M = .07, SEM = .02) ($F < 1$).

A typical KIA effect also was observed in Experiment 2B; the proportion of items on which the participants switched from the incorrect response on Test 1 to the correct response on Test 2 was higher in the feedback condition (M = .62, SEM = .07) than in the control condition (M = .29, SEM = .04) [$F(1,19) = 32.73$, $MS_e = .03$, $p < .001$, $\eta_p^2 = .63$]. Furthermore, there was a significant effect in Experiment 2B for items on which they switched from the correct answer on Test 1 to the incorrect answer on Test 2: As in Experiment 1B, the proportion of switching from correct to incorrect was higher for the control items (M = .08, SEM = .02) than for the feedback items (M = .04, SEM = .01) [$F(1,19) = 4.70$, $MS_e = .003$, $p = .04$, $\eta_p^2 = .20$].

Subjective measures of the KIA effect. The overall proportions of items given an R, JK, or G rating for the judgment task in Experiment 2A are shown for all the participants in Figure 4. The transformed proportions of R/JK/G designations for the items on which they switched from the incorrect side of the scale on Test 1 to the correct side of the scale on Test 2 in Experiment 2A were analyzed in 2 (item type: feedback or control) × 3 (judgment option: R, JK, or G) within-subjects ANOVAs. Six participants were dropped from the analyses for having zero feedback and/or control items on which they switched to the correct answer on Test 2. The interaction between item type and judgment option failed to reach significance [$F(2,28) = 2.74$, $MS_e = .27$, $p = .08$, $\eta_p^2 = .16$], although the planned comparisons showed a significant difference for the G option. Specifically, when the participants had switched from the incorrect to the correct answer, they were less likely to choose the G option for the feedback items (M = .38, SEM = .09) than for the control items (M = .56, SEM = .10) [$t(14) = 3.16$, $p < .01$, $\eta_p^2 = .42$]. Furthermore, from Figure 4, there appears to be a difference in the R response between the feedback items (M = .34, SEM = .08) and the control items (M = .11, SEM = .06), but the data from the participants included in the follow-up analysis of the transformed data failed to reach significance [$t(14) = .70$, $p = .50$]. There was also a nonsignificant trend for higher JK response in the control condition (M = .33, SEM = .10) than in the feedback condition (M = .28, SEM = .08) [$t(14) = 1.90$, $p = .08$, $\eta_p^2 = .20$]. The data for the IRK

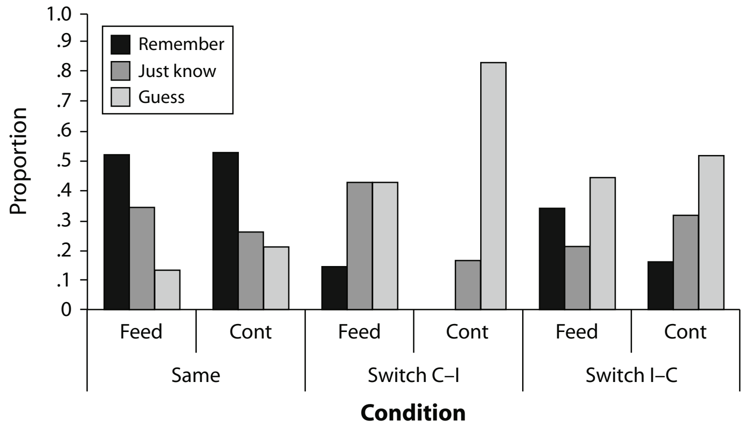


Figure 4. Remember/just-know/guess ratings for the response judgment (i.e., side of scale) in Experiment 2A for the feedback (Feed) and control (Cont) conditions (collapsed across all participants). The judgments are separated by Test 1 and Test 2 responses: (A) items given the same number on Test 1 and Test 2 (*same*); (B) items on which the participants switched from the correct response on Test 1 to the incorrect response on Test 2 (*switch C–I*); and (C) items on which they switched from the incorrect response on Test 1 to the correct response on Test 2 (*switch I–C*).

model show a reversed, but significant, trend; that is, F was higher in the feedback condition (M = .40, SEM = .11) than in the control condition (M = .34, SEM = .10) [$t(14) = 2.29$, $p < .04$, $\eta_p^2 = .27$]. Therefore, the IRK data indicate that when the participants switched from incorrect on Test 1 to correct on Test 2, they more often had an illusory experience of familiarity for feedback than for control items.

The overall proportions of items given an R, JK, or G rating for the response judgment in Experiment 2B are shown for all the participants in Figure 5. The transformed proportions of R/JK/G designations for the trivia items that switched from an incorrect answer on Test 1 to a correct answer on Test 2 were analyzed in a 2 (item type: feedback or control) $\times$ 3 (judgment option: R, JK, or G) within-subjects ANOVA. Five participants were excluded from the analyses for having either no feedback or control items on which they switched to the correct answer on Test 2. There was a significant interaction between item type and judgment option [$F(1,28) = 11.89$, $MS_e = .37$, $p = .001$, $\eta_p^2 = .46$]. Planned comparisons showed that for the items on which the participants switched from an incorrect response on Test 1 to a correct response on Test 2, they were significantly more likely to claim that they remembered having given those correct answers on Test 1 for the feedback items (M = .37, SEM = .06) than for the control items (M = .05, SEM = .04) [$t(14) = 4.95$, $p < .001$, $\eta_p^2 = .67$]. There was no difference between the feedback and the control items for the JK judgment option ($t < 1$), although there was a nonsignificant trend for higher F in the feedback condition (M = .52, SEM = .09) than in the control condition (M = .32, SEM = .07) [$t(14) = 1.99$, $p = .07$, $\eta_p^2 = .22$]. Finally, the participants were less likely to use the G option for the feedback items on which they had switched to the correct answer (M = .33, SEM = .07) than for the control items (M = .67, SEM = .08) [$t(14) = 3.85$, $p < .01$, $\eta_p^2 = .52$].

Both Experiments 2A and 2B yielded evidence of a subjective KIA effect, but of qualitatively different kinds. In Experiment 2A, when the participants shifted from the incorrect answer on Test 1 to the correct answer on Test 2, reports of guessing were rarer and estimates of F were higher for feedback items than for control items. This pattern suggests that working with a puzzle during the feedback phase led to an undifferentiated feeling during Test 2 of having produced the correct answer to that puzzle in Test 1. That same tendency also emerged (albeit not significantly) in Experiment 2B, but the novel and more pronounced effect in this experiment was that when the participants switched from incorrect to correct answers, reports of remembering were more frequent for feedback items than for control items. That is, working with a puzzle during the feedback phase often led the participants falsely to report remembering having given the correct answer in Test 1. These data suggest that the participants sometimes misattributed memories of producing answers during the free-recall portion of the feedback phase as memories of solving the corresponding problems during Test 1 (i.e., source-monitoring confusions; Johnson et al., 1993).

EXPERIMENT 3

One major methodological difference between the 2AFC trivia questions experiment (Experiment 1B) and the 2AFC Wordies puzzles experiment (Experiment 2B) was the timing of the feedback manipulation; feedback was given on Day 2 for the trivia questions paradigm, whereas it was presented on Day 1 for the Wordies version. Therefore, it is not clear whether the difference found in subjective experience between the two experiments was

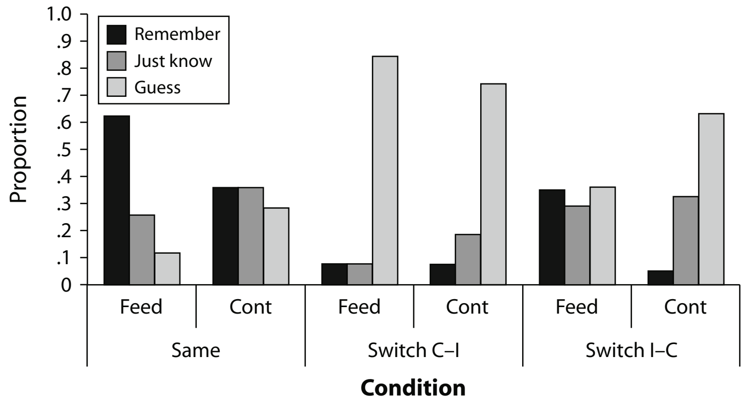


Figure 5. Remember/just-know/guess ratings for the response judgment in Experiment 2B for the feedback (Feed) and control (Cont) conditions (collapsed across all participants). The judgments are separated by Test 1 and Test 2 responses: (A) items given the same response on Test 1 and Test 2 (*same*); (B) items on which the participants switched from the correct response on Test 1 to the incorrect response on Test 2 (*switch C–I*); and (C) items on which they switched from the incorrect response on Test 1 to the correct response on Test 2 (*switch I–C*).

due to the difference in stimuli and the nature of the feedback phase (as was intended) or whether the timing of the feedback may have had an impact on the subjective measure. Experiment 3 replicated Experiment 2B with the addition of a feedback-timing manipulation.

Method

Participants. Forty University of Victoria students participated in exchange for optional extra credit in an introductory psychology course.

Materials and Procedure. The same Wordies stimuli as those from Experiments 2A and 2B were used. The same procedure as that in Experiment 2B was used, except that half of the participants completed the feedback phase on Day 1 (as in Experiment 2B) and half of the participants completed the feedback phase on Day 2 of the experiment (as in Experiment 1B).

Results and Discussion

Objective measures of the KIA effect. As was expected, a KIA effect was found; the proportion of items on which the participants switched from the incorrect response on Test 1 to the correct response on Test 2 was higher in the feedback condition ($M = .48$, $SEM = .22$) than in the control condition ($M = .27$, $SEM = .14$) [$F(1,38) = 29.17$, $MS_e = .03$, $p < .001$, $\eta_p^2 = .43$]. Furthermore, there was no effect of feedback presentation (Day 1 vs. Day 2) on the KIA effect ($F < 1$).

Unlike in Experiments 1B and 2B, there was no significant main effect for the proportion of items on which the participants switched from the correct answer on Test 1 to the incorrect answer on Test 2 between the control items ($M = .10$, $SEM = .09$) and the feedback items ($M = .11$, $SEM = .10$) ($F < 1$). There was a nonsignificant trend for the interaction [$F(1,38) = 3.17$, $MS_e = .01$, $p = .08$, $\eta_p^2 = .08$]: Switching to the incorrect answer on Test 2 was directionally higher for feedback items ($M = .12$, $SEM = .12$) than for control items ($M = .09$, $SEM = .08$) when feedback was given on Day 2, but switching to the incorrect answer on Test 2 was directionally higher for control items ($M = .12$, $SEM = .10$) than for feedback items ($M = .09$, $SEM = .08$) when feedback was given on Day 1. Although potentially interesting, neither of these patterns for switching to the incorrect answer on Test 2 was significant (ts ≤ 1.40, ps $\geq .18$). Thus, the feedback-timing manipulation had little if any effect.

Subjective measures of the KIA effect. The overall proportions of items given an R, JK, or G response judgment are shown in Figure 6. The transformed proportions of R/JK/G designations for the Wordies items on which the participants switched from an incorrect answer on Test 1 to the correct answer on Test 2 were analyzed in a 2 (item type: feedback or control) $\times$ 3 (judgment option: R, JK, G) $\times$ 2 (feedback timing: Day 1 or Day 2) mixed-models ANOVA, with item type and judgment option as repeated measures and feedback timing as a between-subjects factor. Four participants were excluded from the analyses for having either no feedback or no control items on which they switched to the correct answer on Test 2. As was anticipated, there was a significant interaction between item type and judgment option [$F(2,68) = 7.86$, $MS_e = .42$, $p = .001$, $\eta_p^2 = .19$]. There was no effect of feedback timing on this interaction [$F(2,68) = 1.22$, $MS_e = .42$, $p = .30$, $\eta_p^2 = .04$], and therefore this manipulation was dropped from the remaining analyses.

The planned follow-up comparisons for the interaction showed the same patterns as in Experiment 2B. That is, for the items on which the participants switched from an incorrect response on Test 1 to a correct response on Test 2, the participants were significantly more likely to claim that they remembered giving those correct answer on Test 1 for the feedback items ($M = .27$, $SEM = .05$) than for the control items ($M = .04$, $SEM = .02$) [$t(35) = 3.50$, $p = .001$, $\eta_p^2 = .26$]. Also, there was no difference between the feedback and the control items for the JK judgment option ($t < 1$), but again there was a trend toward a higher F value in the feedback condition ($M = .38$, $SEM = .06$) than in the control condition ($M = .29$,

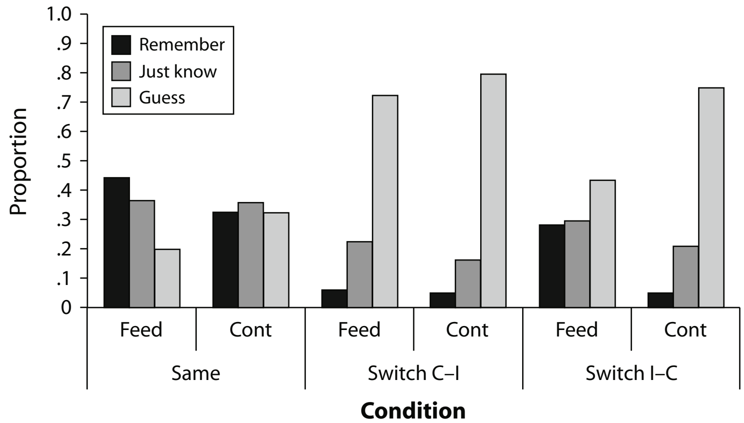


Figure 6. Remember/just-know/guess ratings for the response judgment in Experiment 3 for the feedback (Feed) and control (Cont) conditions (collapsed across all participants and feedback conditions). The judgments are separated by Test 1 and Test 2 responses: (A) items given the same response on Test 1 and Test 2 (*same*); (B) items on which the participants switched from the correct response on Test 1 to the incorrect response on Test 2 (*switch C–I*); and (C) items on which they switched from the incorrect response on Test 1 to the correct response on Test 2 (*switch I–C*).

SEM = .06) [$t(35) = 1.80, p = .08, \eta_p^2 = .09$]. Finally, the participants were less likely to choose G for the feedback items on which they had switched to the correct answer (*M* = .42, *SEM* = .05) than for the control items (*M* = .71, *SEM* = .06) [$t(35) = 3.84, p < .001, \eta_p^2 = .30$].

The results of Experiment 3 replicated the previous experiment (Experiment 2B); a subjective component to the KIA effect was found, and the participants were more likely to label feedback KIA items as remembered than control KIA items. Thus, placing the feedback phase on Day 1 (rather than on Day 2, as in the trivia KIA paradigms) was not the factor responsible for producing an accompanying subjective phenomenology to the KIA effect. Instead, it was the change in stimuli and the procedures in Experiments 2B and 3 that led to a subjective KIA effect.

GENERAL DISCUSSION

All five of the experiments presented in this article demonstrated a KIA effect in the objective measures traditionally used to quantify the phenomenon. More interesting, Experiments 1A, 1B, and 2A provided little evidence that the participants had any accompanying subjective feeling of previous knowledge for the KIA items. Rather, when the participants demonstrated a KIA effect in these experiments (for switching from the incorrect to the correct side of the scale or from the incorrect to the correct 2AFC alternative), they usually reported that they were guessing their Test 1 response. This result suggests that under the conditions most often used in prior research on the KIA phenomenon, participants rarely experienced illusory feelings of having known it all along.

Our studies with Wordies, in contrast, show that participants do sometimes report illusory experiences of knowing or remembering that they gave correct answers. When participants make such errors, they are not necessarily claiming to have "known it all along"; such participants might, instead, believe that the (erroneously) remembered/known Test 1 response was just a guess. That is, the subjective experience effect might be one of "I remember/know I guessed X on Test 1," rather than "I remember/know I knew X on Test 1." The effect would be interesting either way, but to explore this issue we conducted a subanalysis of cases in which the participants in Experiment 2A switched from the incorrect side of the number scale on Test 1 to an above-minimum-confidence answer on the correct side of the number scale on Test 2 (i.e., above a pure guess response of 5 or 6). If the KIA effect was restricted to false beliefs about correct guesses, no such effect should be obtained when minimal-confidence reports are excluded. However, even with such restrictions in place, a KIA effect was found; the proportion of items on which the participants switched from the incorrect response on Test 1 to the above-minimum-confidence correct response on Test 2 was higher in the feedback condition (*M* = .29, *SEM* = .06) than in the control condition (*M* = .11, *SEM* = .03) [$F(1,21) = 12.36, MS_e = .03, p < .01, \eta_p^2 = .38$]. The very small number of control items meeting the restriction (15 items across participants) precluded a statistical comparison of the R/JK/G data for control versus feedback items, but the subjective experiences reported for feedback items meeting the restriction showed even higher rates of false remembering/knowing (R = .43, JK = .38, and G = .19) than in the overall analysis. This pattern suggests that the participants sometimes experienced false feelings of having known it all along.

Most interesting of all, the results from Experiments 2B and 3 illustrated that individuals do sometimes report illusory recollections of answering with the feedback information in foresight. We speculate that the type of stimuli (Wordies, which once understood seem inherently correct), testing procedure (2AFC, which allows for more memorable Test 1 experiences than does a number scale), and feedback structure (training/understanding judgment followed by forced recall of solutions) all contributed to problems in memory source monitoring (Johnson et al., 1993; Mitchell & Johnson, 2000), so that memories of the training phase were misattributed to Test 1. Memories of recalling an answer in training would be quite similar to (and hence, confusable with) memories of having generated that answer during Test 1 (for related points in the context of cryptomnesia, see Bredart, Lampinen, & Defeldre, 2003; Stark, Perfect, & Newstead, 2005).

To the extent that prior treatments of the KIA effect touched on the issue of subjective experience, the implication has often been that people have a feeling of having known it all along, not that they have detailed illusory recollections of giving the feedback information in foresight. However, the more striking evidence for the existence of subjective phenomenology accompanying the KIA effect in Experiments 2B and 3 was found in the R responses. This finding is consistent with our speculation that the subjective KIA effect in these experiments arose from misattributions of memories of the training phase to Test 1; to the extent that such memories were experienced as having specific episodic content (as opposed to an undifferentiated feeling of familiarity), participants would report "remembering" Test 1.

Although the large and dramatic subjective KIA effect in these experiments was a false remembering effect, Experiments 2A, 2B, and 3 each yielded hints of a parallel effect on familiarity (as estimated with Jacoby's IRK equation): For the items on which the participants switched from being incorrect on Test 1 to being correct on Test 2, estimates of F tended to be higher on feedback items than on control items. That is, the training phase not only fostered the development of false recollections of having given particular correct answers, but also contributed to more undifferentiated feelings of having known it all along. This effect, too, can be described in terms of the source-monitoring framework. Just as the training phase can give rise to episodic recollections that can be misattributed to Test 1, it can also give rise to less source-specific memories of the sort that underlie undifferentiated feelings of familiarity, and these too can be misattributed to Test 1 (see Bodner & Lindsay, 2003; Gruppuso, Lindsay, & Kelley, 1997; Gruppuso, Lindsay, & Masson, 2007).

Implications for Theoretical Explanations of the KIA Effect

Theories pertaining to the KIA effect typically fall under two general categories: memory impairment and biased reconstruction (e.g., Schwarz & Stahlberg, 2003).

Memory impairment. Fischhoff (1977) is credited with spawning a great deal of interest and follow-up research on testing theoretical explanations of the KIA effect (Christensen-Szalanski & Willham, 1991; Dehn & Erdfelder, 1998), and his automatic assimilation hypothesis is the most well-known memory impairment theory of the KIA effect. Fischhoff (1975, 1977; Fischhoff & Beyth, 1975) hypothesized that the KIA effect results from an automatic assimilation of the correct feedback with preexisting knowledge. It is not clear how the assimilation account would explain differences in the subjective phenomenology of the hindsight bias. That is, an objective hindsight bias was found in all five experiments reported in this article, but strong evidence for an accompanying belief that the feedback information was known in foresight was demonstrated only in the final two experiments (i.e., 2AFC with word puzzles).

If automatic assimilation is the underlying mechanism of the objective effect, one would expect that participants would report a subjective feeling of having known it all along whenever they displayed hindsight bias. One of the main shortcomings of the automatic assimilation theory, though, is that it fails to take context into account: Our results indicate that whether an individual experiences a *feeling* of having known some piece of information in foresight is dependent, at least to some degree, on whether the situation in which the judgment is made supports illusory recollection/familiarity.

Biased reconstruction. A more recent theoretical account of hindsight bias is the biased reconstruction approach (e.g., Sanna, Schwarz, & Small, 2002; Schwarz & Stahlberg, 2003). This approach emphasizes the contribution of aspects of subjective experience to the creation of hindsight bias, although its proponents have not measured the subjective experience of hindsight bias itself.

Sanna, Schwarz, and Stocker (2002) argued that two different kinds of information become available when an individual tries to recall details from memory: *accessible content* (the details that come to mind) and *accessibility experiences* (how easy/difficult it was to bring those details to mind). Furthermore, they stated that accessibility experiences are an important type of information because they "qualify" the inferences that an individual draws from the accessible content (for related ideas, see Jacoby et al., 1989; Whittlesea, 2004). To test this assertion, Sanna, Schwarz, and Stocker manipulated the number of participants' thoughts pertaining to different outcomes of KIA items because "the most frequent recommended remedy for debiasing the hindsight effects is to search for reasons why the event might have turned out otherwise" (p. 497). The researchers hypothesized that this strategy would not work under all conditions because accessibility experiences influence the conclusions made from bringing these alternatives to mind. They created a hypothetical KIA paradigm (using event stimuli) in which participants had to generate either many thoughts regarding how the event could have turned out differently from the feedback they received (*10-thought* condition) or a few thoughts regarding alternatives (*2-thought* condition). The participants subsequently were required to judge the probability that the event could have had a different outcome. Results from the 10-thought versus 2-thought conditions demonstrated what Sanna, Schwarz, and Stocker (2002) referred to as a "backfire effect"; that is, the participants in the 10-thought condition judged the probability of a different outcome's occurring for the event as less likely than did the participants in the 2-thought condition.

Similar to Sanna, Schwarz, and colleagues' approach to the KIA effect (Sanna & Schwarz, 2003; Sanna, Schwarz, & Small, 2002; Sanna, Schwarz, & Stocker, 2002), Schwarz and Stahlberg (2003) proposed a biased reconstruction view of hindsight bias. More specifically, they argued that participants in a KIA paradigm use the feedback they are given as a basis for recreating their pre-outcome response to an item, and one of their main criticisms of the automatic assimilation approach to hindsight bias is that it does not factor in the use of metacognitive processes. In their biased reconstruction view of the KIA effect, "hindsight distortions will only occur when people forget their original prediction and—while reconstructing it—have reason to believe that their initial estimate must have been close to what is now the known outcome" (Schwarz & Stahlberg, 2003, p. 398). To investigate the use of metacognitive processes, Schwarz and Stahlberg manipulated participants' beliefs regarding how close they had been in the foresight judgment to the feedback information that they subsequently received. The researchers predicted that a higher hindsight bias would be found for participants who had been led to believe that their foresight judgments were relatively good (i.e., close to the feedback information), whereas a smaller effect should be found when participants believed that their original judgments were quite poor in comparison with the correct answers, and the data from their studies have largely conformed to this prediction.

Despite their emphasis on subjective experience as a causal agent, proponents of the biased reconstruction approach to the KIA effect have not assessed participants' subjective experience of the effect itself. Also, at this point, the approach seems rather simplistic (e.g., recollection of the original judgment either does or does not occur). Furthermore, no mention is made as to what specific factors may influence recollection of original judgments, and whether these same factors may influence the reconstruction of prefeedback responses. Nonetheless, the present findings are more compatible with the general spirit of the biased reconstruction approach than with that of the memory impairment approach.

Summary and Conclusions

The results of the present experiments demonstrated that the KIA effect sometimes has an accompanying component of subjective phenomenology, but the data also showed that it cannot be assumed that the measure typically used to characterize the effect (i.e., a numerical move toward, or switch to, the feedback information on the final test) specifies anything about the subjective phenomenology of the bias. Separating the objective and subjective components of hindsight bias is important for a variety of reasons. For example, the results from the present five experiments clearly indicate that treating the phenomenology that accompanies the KIA effect (whether it be "remembering," "knowing," or "guessing") as a mere by-product of the objective measure is unfounded and misleading: Whether participants claim to have remembered or to have known being correct in foresight will rely on whether the context that they are in lends itself to such feelings and beliefs. Investigating the situations under which individuals do versus do not have an accompanying feeling of having known in foresight things that they really learned subsequently will tell us more about both the nature of the effect itself and its underlying mechanisms.

Of particular interest in this regard is the question of the subjective experiences accompanying real-world KIA effects. Katz (1989), for example, found that baseball fans (but not nonfans) misremembered their predictions of team rankings as having been closer to the actual rankings than they had been. Ross (1989) summarized numerous studies in which participants' current knowledge/beliefs regarding real-world matters (e.g., their current feelings toward a romantic partner) distorted their reports of their prior knowledge/beliefs regarding those matters. Researchers exploring such KIA-like effects in real-world settings have not, to the best of our knowledge, assessed participants' subjective experience.

Similar issues pertain to other sorts of memory biases and errors. We mentioned at the outset the analogy between hindsight bias and the eyewitness misinformation effect: Just because participants report a suggested detail does not necessarily mean that they have a subjective experience of remembering witnessing that detail (e.g., Lindsay & Johnson, 1989). For example, several researchers have found that participants are much more likely to correctly attribute misinformation to its source (e.g., as occurring in a narrative and not the witnessed event) when a short delay is inserted prior to testing their memory for the event (Frost, Ingraham, & Wilson, 2002; Lindsay & Johnson, 1989; Zaragoza & Koshmider, 1989). Furthermore, just as we have argued in this article against the automatic assimilation interpretation of the KIA effect, Lindsay and Johnson claimed that the reduction in the misinformation effect that is found for source-monitoring tests (i.e., as compared with the standard *yes/no* testing format) goes against a memory impairment account of the effect (e.g., Loftus, 1979). Although there are some major differences between hindsight bias and the misinformation effect (e.g., the feedback in our KIA paradigm was never false information), it is possible that similar reductions could be found with a source-monitoring test; that is, requiring participants to complete a source-monitoring test in a KIA paradigm may reduce the objective and/or subjective components of the effect (i.e., reduce the size of the hindsight bias and/or eliminate subjective phenomenology of KIA items).

Beyond misinformation paradigms, there has also been considerable interest in the subjective experience that accompanies false memory reports in the procedure developed by Deese (1959; e.g., Neuschatz, Payne, Lampinen, & Toglia, 2001). Likewise, schema- and script-driven reconstructive memory errors (as explored in classic research by Bartlett, 1932, and Bransford & Johnson, 1972) may or may not be experienced as memories per se (see Lampinen, Faries, Neuschatz, & Toglia, 2000). Thus, we see our research on the phenomenology of the KIA effect as part of a more general zeitgeist emphasizing the importance of the subjective experience of memory phenomena (e.g., Rubin, Schrauf, & Greenberg, 2003; Tulving, 2005).

AUTHOR NOTE

Portions of this research were reported at the 45th Annual Meeting of the Psychonomic Society, Minneapolis, Minnesota, November 2004, and at the 4th International Conference on Memory (ICOM), Sydney, Australia, July 2006. This research was supported by the Natural Science and Engineering Research Council (NSERC) of Canada through a postgraduate scholarship to M.M.A. and by NSERC Grant OGPO121316-95 to D.S.L. Correspondence concerning this article should be addressed to M. M. Arnold, School of Psychology, University of St Andrews, St Andrews, Fife KY16 9JP, Scotland (e-mail: michelle.arnold@st-andrews.ac.uk).

REFERENCES

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Bodner, G. E., & Lindsay, D. S. (2003). Remembering and knowing in context. *Journal of Memory & Language*, **28**, 563-580.

Bonds-Raacke, J. M., Fryer, L. S., Nicks, S. D., & Durr, R. T. (2001). Hindsight bias demonstrated in the prediction of a sporting event. *Journal of Social Psychology*, **141**, 349-352.

Brain Teasers! (n.d.). Retrieved September 8, 2003, from www.billsgames.com/brainteasers/.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*, **11**, 717-726.

Bredart, S., Lampinen, J. M., & Defeldre, A. (2003). Phenomenal characteristics of cryptomnesia. *Memory*, **11**, 1-11.

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

Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General*, **126**, 393-413.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58,** 17-22.

Dehn, D. M., & Erdfelder, E. (1998). What kind of bias is hindsight bias? *Psychological Research*, **61**, 135-146.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 288-299.

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

Fischhoff, B., & Beyth, R. (1975). "I knew it would happen": Remembered probabilities of once-future things. *Organizational Behavior & Human Decision Processes*, **13**, 1-16.

Frost, P., Ingraham, M., & Wilson, B. (2002). Why misinformation is more likely to be recognized over time: A source monitoring account. *Memory*, **10**, 179-185.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition*, **18**, 23-30.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, **10**, 83-98.

Goethals, G. R., & Reckman, R. F. (1973). The perception of consistency in attitudes. *Journal of Experimental Social Psychology*, **9**, 491-501.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 259-278.

Gruppuso, V., Lindsay, D. S., & Masson, M. E. J. (2007). I'd know that face anywhere! *Psychonomic Bulletin & Review*, **14**, 1085-1089.

Halford, K. W., & Griffith, U. (2002). "How has the week been for you two?" Relationship satisfaction and hindsight memory biases in couples' reports of relationship events. *Cognitive Therapy & Research*, **26**, 759-773.

Hardt, O., & Pohl, R. F. (2003). Hindsight bias as a function of anchor distance and anchor plausibility. *Memory*, **11**, 379-394.

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

Hell, W., Gigerenzer, G., Gauggel, S., Mall, M., & Muller, M. (1988). Hindsight bias: An interaction of automatic and motivational factors? *Memory & Cognition*, **16**, 533-538.

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

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391-422). Hillsdale, NJ: Erlbaum.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.

Katz, A. N. (1989). Autobiographical memory as a reconstructive process: An extension of Ross's hypothesis. *Canadian Journal of Psychology*, **43**, 512-517.

Lampinen, J. M., Faries, J. M., Neuschatz, J. S., & Toglia, M. P. (2000). Recollections of things schematic: The influence of scripts on recollective experience. *Applied Cognitive Psychology*, **14**, 543-554.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, **17**, 349-358.

Loftus, E. F. (1979). The malleability of human memory. *American Scientist*, **67**, 312-320.

Marsh, E. J., Meade, M. L., & Roediger, H. L., III (2003). Learning facts from fiction. *Journal of Memory & Language*, **49**, 519-536.

Mazursky, D., & Ofir, C. (1990). "I could never have expected it to happen": The reversal of the hindsight bias. *Organizational Behavior & Human Decision Processes*, **46**, 20-33.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179-195). Oxford: Oxford University Press.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning & Verbal Behavior*, **19**, 338-368.

Neuschatz, J. S., Payne, D. G., Lampinen, J. M., & Toglia, M. P. (2001). Assessing the effectiveness of warnings and the phenomenological characteristics of false memories. *Memory*, **9**, 53-71.

Pohl, R. F., Schwarz, S., Sczesny, S., & Stahlberg, D. (2003). Hindsight bias in gustatory judgments. *Experimental Psychology*, **50**, 107-115.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, **96**, 341-357.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum–difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review*, **111**, 588-616.

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition*, **31**, 887-901.

Sanna, L. J., & Schwarz, N. (2003). Debiasing the hindsight bias: The role of accessibility experiences and (mis)attributions. *Journal of Experimental Social Psychology*, **39**, 287-295.

Sanna, L. J., Schwarz, N., & Small, E. M. (2002). Accessibility experiences and the hindsight bias: I knew it all along versus it could never have happened. *Memory & Cognition*, **30,** 1288-1296.

Sanna, L. J., Schwarz, N., & Stocker, S. L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 497-502.

Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instruments, & Computers*, **20**, 206-217.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user guide*. Pittsburgh: Psychology Software Tools.

Schwarz, S., & Stahlberg, D. (2003). Strength of hindsight bias as a consequence of meta-cognitions. *Memory*, **11**, 395-410.

Sharpe, D., & Adair, J. G. (1993). Reversibility of the hindsight bias: Manipulation of experimental demands. *Organizational Behavior & Human Decision Processes*, **56**, 233-245.

Stahlberg, D., & Maass, A. (1998). Hindsight bias: Impaired memory or biased reconstruction? In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 8, pp. 105-132). Chichester, U.K.: Wiley.

Stark, L., Perfect, T. J., & Newstead, S. E. (2005). When elaboration leads to appropriation: Unconscious plagiarism in a creative task. *Memory*, **13**, 561-573.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *Origins of self-reflective consciousness* (pp. 3-56). New York: Oxford University Press.

Werth, L., & Strack, F. (2003). An inferential approach to the knew-it-all-along phenomenon. *Memory*, **11**, 411-419.

Whittlesea, B. W. A. (2004). The perception of integrality: Remembering through the validation of expectation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 891-908.

Williams, A. D. (1992). Bias and debiasing techniques in forensic psychology. *American Journal of Forensic Psychology*, **10**, 19-26.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, **11**, 616-641.

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

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, **130**, 361-379.

Zaragoza, M. S., & Koshmider, J. W. (1989). Misled subjects may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 246-255.

NOTES

1. There are many articles that focus specifically on issues surrounding R/K judgments, such as Gardiner et al. (2002); Rotello, Macmillan, and Reeder (2004); Wixted and Stretch (2004); and Yonelinas (2001). Also, see Gruppuso et al. (1997) and Bodner and Lindsay (2003) for an explanation of the R/K distinction that combines aspects of both the qualitative and the quantitative approaches.

2. In Experiments 1A and 2A, the participants also made a second R/JK/G judgment pertaining to their memories of choosing a particular number during Test 1. The participants very rarely reported remembering or knowing precise numbers, so this measure yielded little of interest and will not be discussed further in this article. Details of these data are available from the first author.

3. We changed the traditional *know* judgment label to *just know* because we believed that the participants would better understand the task with this alteration (e.g., "Even though I don't remember any specific details, I *just know* that I gave that response in the first test!"; see Conway, Gardiner, Perfect, Anderson, & Cohen, 1997, for a more detailed discussion of this distinction). Therefore, the terms *know* (K) and *just know* (JK) will be used interchangeably throughout the remainder of this article.

4. We thank Michael A. Hunter for suggesting this transformation approach to analyzing the subjective data.

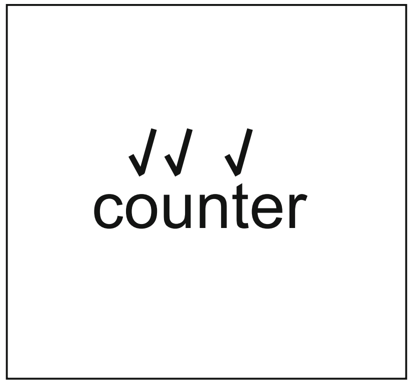
APPENDIX A

Table A1
Mean Proportions of Response Judgment Remember/Just-Know/Guess (R/JK/G) Designations (Experiments 1–5) for the Natural Log Transformed Data (Items That Participants Switched From the Incorrect Response on Test 1 to the Correct Response on Test 2) for Feedback and Control Items

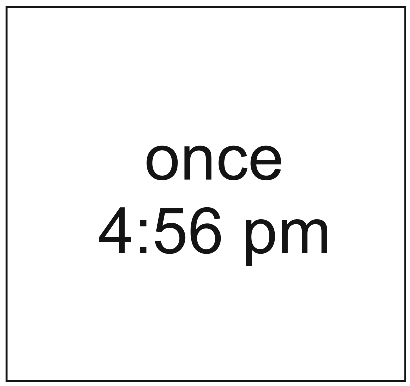
	Feedback			Control		
	R	JK	G	R	JK	G
Experiment 1	−1.27	−1.06	−0.70	−1.06	−0.88	−0.54
Experiment 2	−1.20	−1.30	−0.63	−1.13	−1.28	−0.35
Experiment 3	−1.08	−1.25	−0.97	−1.21	−0.93	−0.50
Experiment 4	−0.93	−1.08	−1.09	−1.76	−1.02	−0.40
Experiment 5	−1.17	−1.10	−0.81	−1.57	−1.15	−0.40

(*Continued on next page*)

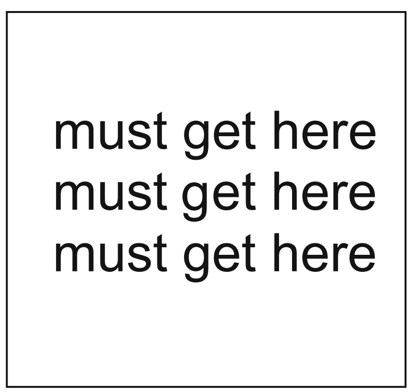
APPENDIX B



solution: check-out counter



solution: once upon a time



solution: the three musketeers

(Manuscript received October 18, 2006;
revision accepted for publication March 8, 2007.)