Accountability Reduces Unconscious Plagiarism

BLAIRE J. WEIDLER1,2, KRISTI S. MULTHAUP2* and MARK E. FAUST3

1Washington University in St. Louis, St. Louis, MO, USA
2Davidson College, Davidson, NC, USA
3University of North Carolina at Charlotte, Charlotte, NC, USA

Summary: We investigated how holding participants accountable for their responses affected unconscious plagiarism when solving a Boggle puzzle task (finding words in a 4 × 4 letter matrix). Both experimental and control participants (N = 60) generated puzzle solutions with a computer partner, recalled their own previously generated solutions, and then produced new solutions to the puzzles. Accountability was manipulated by telling participants in the experimental group before beginning the initial-generation phase that at the end of the session, they would review their generated responses with the researcher (accountable participants). Accountable participants plagiarized less than control participants when generating words with the computer and generating new solutions on their own but not when they were attempting to recall words they initially generated. The data are discussed in terms of the leading theoretical explanation of unconscious plagiarism, the source-monitoring framework. Copyright © 2012 John Wiley & Sons, Ltd.

Cases of plagiarism are of the utmost concern in academic and non-academic contexts alike (e.g. Goodwin withdraws, 2002; Plotz, 2002), but not all plagiarism is intentional. Anecdotal cases of unconscious plagiarism have been recorded in which the perpetrator was truly unaware of his crime (e.g. Freud, 1917). Cryptomnesia, or unconscious plagiarism, occurs when remembered ideas are not recognized as memories (Taylor, 1965). In other words, cryptomnesia occurs when an individual claims an idea to be her own new creation, when in fact, she fails to remember that it was produced by someone else, or even by herself at a prior point in time (Johnson, Hashtroudi, & Lindsay, 1993; Marsh & Bower, 1993; see also Goodwin, 2002). The present study tested the hypothesis that unconscious plagiarism can be reduced by simply highlighting that participants’ responses will be reviewed with a knowledgeable other, in this case, with the researcher. To lay the foundation for this research, we introduce how cryptomnesia is investigated in the laboratory and discuss why this memory error arises. Then, we discuss the effect of motivation on unconscious plagiarism. Finally, we summarize the hypothesis in the present study.

The first laboratory investigation of cryptomnesia was performed by Brown and Murphy (1989), and their three-stage paradigm has been utilized in many subsequent investigations of the phenomenon. In the first stage, called initial generation, participants took turns within a group of four participants generating a new exemplar from common categories such as sports and clothing. In the second stage, recall own, participants recalled the four exemplars they personally generated for each category. In the third stage, generate new, participants generated four new examples that had not been stated by the group previously. Brown and Murphy defined unconscious plagiarism as when a participant recalled another person’s category exemplar when she was instructed to remember the exemplars she stated or when a participant wrote another person’s exemplar when she was instructed to write an entirely new response. They also noted that a person can plagiarize herself when attempting to generate a new word, yet she repeats one she generated earlier. Brown and Murphy found that participants unconsciously plagiarized at levels significantly greater than chance in all three stages.

Marsh and Bower (1993) adapted the Brown and Murphy paradigm, so participants generated word solutions to Boggle puzzles with a computer (located words in a 4 × 4 letter matrix). They found that this task elicited unconscious plagiarism of the computer partner and of the participants themselves. Subsequent studies indicate that cryptomnesia also occurs when stimuli to be plagiarized are more complex than single-word responses. Cryptomnesia has been documented for solutions to community problems (e.g. how to reduce traffic accidents; Bink, Marsh, Hicks, & Howards, 1999); novel uses of everyday objects (Stark & Perfect, 2006, 2008); drawing novel space creatures (e.g. Marsh, Landau, & Hicks, 1996); and for creating new works such as poems, prose, cocktails, and marketing strategies (Defeldre, 2005). In short, cryptomnesia is a reliable empirical phenomenon (see also Brown & Halldiy, 1991; Brédart, Lampinen, & Defeldre, 2003; Landau & Marsh, 1997; Marsh & Landau, 1995; Perfect, Field, & Jones, 2009; Perfect & Stark, 2008; Preston & Wegner, 2007) with implications for a range of everyday contexts.

Theoretical accounts of unconscious plagiarism have included a relative-strength-of-memory-trace model (e.g. Marsh & Landau, 1995; but see Stark, Perfect, & Newstead, 2005, and Perfect & Stark, 2008, which effectively refute this view) and an expended-effort model (Preston & Wegner, 2007; but this view addresses only a subset of cryptomnesia situations). The most commonly cited account, and the account that can explain the range of findings in the literature, is that cryptomnesia arises because of a source-monitoring error (e.g. Brown & Murphy, 1989; Johnson et al., 1993; Stark & Perfect, 2008). Source monitoring is the identification of the origin of knowledge, memory, or a belief (Johnson et al., 1993), and the source-monitoring framework (SMF; Johnson et al., 1993; Lindsay, 2008; Mitchell & Johnson, 2000, 2009) outlines the cognitive processes that underlie such identifications. More specifically, cryptomnesia...
can be classified as a reality-monitoring error in which participants confuse internally and externally generated responses. When participants commit unconscious plagiarism of someone else’s responses, they are attributing the source of an externally generated idea to themselves. When participants commit unconscious plagiarism of their own earlier-produced responses when assigned to create new ideas, they are attributing the memory of an earlier response to a newly created idea; because of the overt response, the original source can be labeled external as well as internal (footnote 1 of Johnson et al., 1993).

One prediction of the SMF, and the focus of the current paper, is that social contexts should affect source-monitoring performance (Johnson et al., 1993). The specific example given by Johnson et al. was that testifying in court should inspire people to avoid source-monitoring errors more than telling a story at a party would. In terms of the SMF, the goal to avoid errors could be achieved by adjusting decision criteria. When testifying, people may use relatively strict decision criteria (e.g. asserting Kim said X only if the memory includes many perceptual details and supporting information), whereas when telling a story at a party, people may use relatively loose decision criteria (e.g. asserting Kim said X based on familiarity; see also Multhaup, 1995; Multhaup & Conner, 2002). The importance of a rememberer’s goals has been reiterated in more recent reviews of source monitoring (Lindsay, 2008; Mitchell & Johnson, 2000, 2009). The resulting prediction in terms of cryptomnesia is that individuals motivated to be accurate will exhibit less unconscious plagiarism than those who are not so motivated. Given the face validity of this suggestion and the potential for reducing the problematic error of unconscious plagiarism, we tested the hypothesis that unconscious plagiarism can be reduced by simply highlighting that participants’ responses will be reviewed with a knowledgeable other at the end of the experimental session.

Previous research indicates that motivation may indeed reduce cryptomnesia errors. For example, Stark et al. (2005) offered money to any participants who could avoid unconscious plagiarism. Across experiments, the researchers found participants motivated by a financial reward showed reduced rates of unconscious plagiarism compared with controls. In addition, Marsh, Landau, and Hicks (1997) asked participants to generate solutions to common problems in a group and then participants returned to the laboratory a week later to offer new problem solutions. The experimenters stringently warned half the participants to avoid duplicating responses offered the week before. Warned participants exhibited less cryptomnesia than controls did. Similarly, Landau, Thomas, Thelen, and Chang (2002) warned half their participants that a plagiarism expert would evaluate their drawings of novel space creatures for features from the researcher’s example creatures. These participants, motivated by the instructions to avoid plagiarizing, copied fewer features from the researcher’s example creatures than participants who were told their ideas would be evaluated for creativity.

While Stark et al. (2005), Marsh et al. (1997), and Landau et al. (2002) suggest that motivated participants commit fewer cryptomnesia errors than control participants do, these studies provided experimental participants with information about how to avoid cryptomnesia as well as manipulating motivation. Stark et al. (2005) told participants they would receive a cash prize if they did “not re-produce any ideas from the initial session (previously given by themselves or any of the other participants)” (p. 568). Similarly, Marsh et al.’s (1997) stringent warning included:

When people attempt to generate ideas, often they will produce an idea that someone else had provided at an earlier point in time and not even realize they have accidentally made this error. Please be careful to avoid making this mistake when you generate your own new ideas. In other words, you want to avoid generating an idea that was offered last week. (p. 894)

Finally, plagiarism-agenda participants in Landau et al. (2002) “heard that a plagiarism expert would examine their creatures for features from the experimenter’s creatures” (p. 189).

The types of instructions noted earlier, although effective in reducing unconscious plagiarism, confound motivation and additional instructions about the cryptomnesia errors the experimental participants should avoid. Participants in the experimental groups of these studies may have plagiarized less than controls did because they were induced to be accountable for their responses. However, because experimental participants also were forewarned about how they could err and control participants were not, it is not clear how much of a role participants’ accountability (motivation) played in the reduction of cryptomnesia. The present study clarifies this issue by not explicitly defining cryptomnesia or stating what kind of errors the experimenter would be looking for; instead, the researcher induced experimental participants’ accountability for their responses (cf., Chen, 2004; Tetlock, 1985) by simply noting that the participant’s and the computer’s responses would be reviewed later.

In a second experimental manipulation, Marsh et al. (1997) tested some participants individually and tape-recorded their responses and tested other participants in groups with pen-and-paper responses (relatively high and low motivation, respectively). Participants tested in the relatively high motivation context displayed lower rates of unconscious plagiarism than participants tested in the relatively low motivation context did. Perhaps most importantly, the only condition in which participants showed plagiarism rates on their generation task no higher than on their recognition task was when participants were tested individually and given the stringent warning instructions quoted earlier. This pattern is consistent with increased motivation reducing unconscious plagiarism but leaves open the question of whether motivation alone can reduce unconscious plagiarism. The present research attempts to answer exactly that question.

Whereas Marsh et al.’s (1997) interpretation of their data suggested that participants’ perception that their “ideas are being scrutinized” (p. 895) led to reduced rates of unconscious plagiarism, the current experiment directly tested the hypothesis that unconscious plagiarism can be reduced by highlighting that participants’ work will be reviewed. The specific aim of the present study is to isolate the effect of motivation on performance, and thus no information about
the nature of cryptomnesia was given to accountable participants. Furthermore, the procedure mimics everyday life situations in which plagiarism may be detected by others (e.g. professors, authors in the same field). If accountable participants (experimental group) commit fewer unconscious plagiarism errors compared with controls, the data will support the claim that motivation affects source-monitoring performance (e.g. Johnson et al., 1993).

METHOD

Participants
Sixty undergraduates (39 women) participated for psychology course credit.

Materials and apparatus
We used the four Boggle puzzles from Marsh and Bower (1993) with two minor adjustments based on a pilot study with 22 undergraduates: We added cay as a legal word in Puzzle 2, and we reordered the solutions to reflect the frequency with which pilot participants located solutions (the correlations with Marsh and Bower’s orderings for Puzzles 1 through 4 were .85, .87, .93, and .82, respectively). The computer displayed words in increasing odd serial positions (e.g. the first word the computer displayed was the one found first, on average, by the pilot participants; the second was that found third on average, etc.). The hard puzzles had 35 and 44 acceptable word solutions; the easy puzzles had 112 and 113. Puzzle order was counterbalanced across participants (the same puzzle order was used in each task for a given participant). The experiment was conducted, and data were collected using E-Prime 1.1 (Psychology Software Tools, Inc., Sharpsburg, PA, USA) software on a Dell Optiplex GX620 computer.

Design
We used a 2 × 2 mixed-factorial design. Accountability (accountable and control) was manipulated between participants; puzzle difficulty (easy and hard) was manipulated within participants. Although task (initial generation, recall old, and generate new) was technically manipulated within participants, each task yielded different error classifications (see Results section), so we analyzed the data for each task separately.

Procedure
Detailed instructions for the three tasks (initial generation, recall own, and generate new) were read aloud by the researcher and presented on the computer screen for the participants to follow along. After completing each task on the practice puzzle with the researcher present to answer any questions, a participant performed the task on the four critical puzzles alone in the testing room. Although the studied phenomenon is unconscious plagiarism, we use the word plagiarism subsequently for conciseness.

Initial-generation task
To begin, the participants learned the rules for playing Boggle, including Marsh and Bower’s (1993) criteria for acceptable words: three or more letters, not a proper noun, and the root form (e.g. hen was acceptable, but hens was not). Participants then learned that their task was to find as many words as possible together with their computer partner, so the word the participants typed during each of their turns should be new. Next, participants saw a puzzle on the computer screen and a word that the computer “found” in that puzzle. Participants hit a key verifying that they saw the word themselves, and the computer program generated two more solutions in this manner (for the first three words in the practice puzzle, participants were required to show the researcher where the word was in the puzzle). Then, the participant found and typed a word from the puzzle. If the participant typed an invalid word, the program did not allow her to proceed until she typed an acceptable word. Plagiarisms (typing a word already presented by the computer or repeating a word the participant already typed) were not counted as incorrect responses, so the participant was allowed to continue after typing a plagiarized word. If one of the words in the computer’s queue had already been found by the participant, the computer skipped this word and presented the word in the next odd serial position to avoid plagiarizing the participant. This pattern (of three computer words to one participant word) repeated four times per puzzle, resulting in 12 computer-generated and 4 participant-generated responses per puzzle.

Introduction of manipulation before initial-generation critical puzzles
After the participants completed initial generation on the practice puzzle, the accountability condition was induced for a randomly assigned half of the participants. Accountable participants saw text on the screen stating, “Now I have some important information for you…” The researcher then turned on the monitor of a second computer. A mock Microsoft Excel worksheet displayed “computer” and “participant” words from the practice puzzle (not from one of the four critical puzzles). The researcher pointed to the appropriate spots on the screen while reading:

After you are done, I will pull up your answers on the computer like this (this is an example with someone else’s data). We’ll be able to see the words the computer found, and the words you found, and we will go over your responses for the four puzzles you’re about to complete once you are finished.

After this point, the instructions for both groups were again the same. The researcher reminded all participants, “Please remember you are trying to find as many words as possible with your computer partner, so make sure you find a NEW word each time it is your turn.” The researcher then left, and the participants completed the initial generation task for all four critical puzzles.
Recall-own task
In the recall-own task, participants again saw the puzzles one at a time. While viewing each puzzle, participants recalled all four of the words they found and typed in during the initial-generation task. After entering each word, participants rated their confidence that it was one of the four words they typed in the puzzle earlier; the options were 4 (very confident), 3 (confident), 2 (somewhat confident), or 1 (not confident). Participants were required to type four acceptable words for each puzzle; plagiarisms were not counted as incorrect responses, so the participant was allowed to continue after typing a plagiarized word.

We note that Tenpenny, Keriazakos, Lew, and Phelan (1998) critiqued early cryptomnesia studies, pointing out that because participants were required to fill in all the blanks in the recall-own and generate-new tasks, they might be writing down an idea they knew was previously generated by another person out of frustration and then might not be exhibiting unconscious plagiarism. Therefore, many subsequent studies did not require participants to fill in all four blanks. Although we acknowledge the criticism of Tenpenny et al., we worried that the experimental manipulation would cause accountable participants to fill in more blanks than the control participants regardless of their degree of certainty in their response. Thus, we required participants to fill in all four blanks in the recall-own and generate-new tasks. We report analyses that assess the extent to which the critique of Tenpenny et al. applies to the current data.

Generate-new task
Next, the researcher reminded the participants about the instructions for finding words within a puzzle (see initial-generation task instructions). The generate-new task required that the participants find four new words within each puzzle and rate their confidence that each word was entirely new to the puzzle; they used the same 4-point scale used in the recall-own task. Participants were required to type four acceptable words for each puzzle; plagiarisms were not counted as incorrect responses, so the participant was allowed to continue after typing a plagiarized word.

RESULTS
The three tasks yielded different error classifications, so we analyzed the data for each task separately. When means and standard deviations are reported, they were computed from the proportions of the total words participants generated in that task that were plagiarisms (cf. Marsh & Bower, 1993). Figure 1 provides an overview of the plagiarism rates in each task (note that puzzle difficulty is collapsed in Figure 1, cf. Marsh & Bower, 1993). The alpha value was set at .05; non-significant results are reported only for comparisons central to the hypothesis.

Initial-generation task
For initial-generation task errors, participants could plagiarize either themselves (self-plagiarism) or the computer (other-plagiarism); a word was counted as self-plagiarism if the participant repeated a word she typed already, whereas it was counted as other-plagiarism if a participant repeated a word previously shown by the computer. Overall, the 60 participants made 154 other-plagiarisms and 20 self-plagiarisms out of 960 responses, or 18.1% total plagiarisms, slightly higher than the 12.2% reported by Marsh and Bower (1993) for their delayed condition (the one ours most closely replicated). Given that only 14 of 60 participants produced self-plagiarisms (no one produced more than two), we followed Marsh and Bower (1993) and did not separate error type in the analysis of initial-generation plagiarisms. A 2 (condition: accountable and control) × 2 (puzzle difficulty: easy and hard) mixed-factorial analysis of variance (ANOVA) on plagiarized responses yielded main effects of condition and puzzle difficulty, but no interaction. Most importantly, accountable participants showed lower plagiarism rates (M = 0.154, SD = 0.111) than control participants did (M = 0.208, SD = 0.090), F(1, 58) = 4.30, p = .043 (Figure 1), supporting the idea that motivation improves source monitoring (e.g., Johnson et al., 1993). The main effect of puzzle difficulty replicated Marsh and Bower: Plagiarism rates were lower in easy puzzles that had relatively high numbers of possible solutions (M = 0.150, SD = 0.112) than in hard puzzles (M = 0.213, SD = 0.150), F(1, 58) = 8.67, p = .005.

Recall-own task
During the recall-own task, participants were asked to recall the four words they typed in for each puzzle. Two kinds of errors were possible, other-plagiarism and new-word intrusions. A word was counted as other-plagiarism if it was a computer-presented solution from the initial-generation task. A word was counted as a new-word intrusion if it was a correct word solution that was not presented during initial generation. If a word that had already been counted as a plagiarism in the initial-generation task appeared again in the recall-own task, it was not counted as a plagiarism the second time. In addition to acknowledging that participants were correct that they had typed in such a word during the
initial-generation task, as per the recall-own task instructions, this scoring method follows what Marsh and Bower (1993) did, facilitating comparisons across the two studies. The two types of errors were analyzed separately because a new-word intrusion is not an example of plagiarism.

The 60 participants made 270 other-plagiarisms, or 28.1%, slightly lower than the 31.8% reported by Marsh and Bower (1993) for their delayed condition. A 2 (condition: accountable and control) x 2 (puzzle difficulty: easy and hard) mixed-factorial ANOVA on other-plagiarisms yielded no significant effects. In the comparison of interest, accountable (M = 0.277, SD = 0.127) and control participants (M = 0.285, SD = 0.132) showed similar levels of other-plagiarisms, F(1, 58) = 0.06, p = .805.

Participants made 210 new intrusions, or 21.9%, slightly higher than the 18.5% reported by Marsh and Bower (1993) for their delayed condition. A 2 (condition: accountable and control) x 2 (puzzle difficulty: easy and hard) mixed-factorial ANOVA on new-word intrusions yielded no main effect of condition, F(1, 58) = .01, p = .940, (M = 0.219, SD = 0.113 and M = 0.221, SD = 0.099, for accountable and control groups, respectively). Puzzle difficulty, however, did affect the number of new-word intrusions within this task and again replicated the pattern in Marsh and Bower (1993). Participants typed fewer new words in the hard puzzles (M = 0.167, SD = 0.117) than in the easy puzzles (M = 0.273, SD = 0.170), F(1, 58) = 17.00, p < .001.

The similar patterns for the accountable and control participants in this task suggest that any group differences in plagiarism in the generate-new task are not likely to be due to condition differences in how well initial-generation responses were remembered.

Generate-new task

During the generate-new task, participants could either plagiarize themselves (self-plagiarism) or their partner (other-plagiarism). A word was counted as self-plagiarism in two cases: if a participant repeated her response from initial generation (both if the word was first self-plagiarized on the generate-new task and if the word had been counted as a self-plagiarism in initial generation already) or if a participant repeated a new-word intrusion from recall own—in all cases, the word has already been typed into the computer. A word was counted as other-plagiarism if it was a word the computer had generated earlier; this included words that had already been counted as other-plagiarisms in either the initial-generation or recall-own task. (As an analogy, if Pat plagiarizes Jess, and later Pat repeats that same idea, the second time Pat has still plagiarized Jess. In other words, the first plagiarism did not make the idea now Pat’s.)

The 60 participants made 275 other-plagiarisms, or 28.6%, akin to the 28.1% reported by Marsh and Bower (1993) for their delayed condition. Participants made 113 self-plagiarisms, or 11.8%, slightly higher than the 7.4% reported by Marsh and Bower for their delayed condition. A 2 (condition: accountable and control) x 2 (puzzle difficulty: easy and hard) x 2 (plagiarism type: other and self) mixed-factorial ANOVA yielded three main effects. Accountable participants had lower rates of plagiarism (M = 0.176, SD = 0.063) than control participants did (M = 0.228, SD = 0.084), F(1, 58) = 7.38, p = .009. Plagiarism rates were lower in the easy puzzles that had relatively high numbers of solutions (M = 0.140, SD = 0.106) than in the hard puzzles (M = 0.265, SD = 0.094), F(1, 58) = 58.90, p < .001, replicating Marsh and Bower. Participants produced significantly fewer self-plagiarisms (M = 0.118, SD = 0.088) than other-plagiarisms (M = 0.287, SD = 0.115), F(1, 58) = 94.54, p < .001. Furthermore, there was a significant Puzzle difficulty x Plagiarism type interaction, F(1, 58) = 7.76, p = .007 that was driven by a larger hard-easy difference for other-plagiarisms (0.117) than for self-plagiarisms (0.073).

In terms of the main hypothesis, the important finding is that accountable participants produced fewer plagiarism responses than control participants did, supporting the SMF account of cryptomnesia (e.g. Brown & Murphy, 1989; Johnson et al., 1993; Stark & Perfect, 2008).

Confidence data

For two plagiarized responses in the recall-own stage and three plagiarized responses in the generate-new stage, a 0 was reported instead of a valid response (1-4) because of a computer error. These errors constituted less than 0.5% of total responses and were dropped from the analysis. Ten accountable and 10 control participants had a missing value when the puzzle difficulty (easy and hard) factor was included because they did not make every type of plagiarism error possible. By collapsing across the difficulty factor, all participants had complete data. Whether the analysis is run with or without the difficulty factor, the pattern was the same; we report the analysis without the difficulty factor to include all participants in the analysis.

A 2 (stage: recall own and generate new) x 2 (condition: accountable and control) mixed-factorial ANOVA was performed on confidence data of plagiarized responses. The only statistically significant difference was between the confidence ratings in the recall-own and generate-new stages. Participants were less confident (exhibited better judgment) in their plagiarized responses in the recall-own stage (M = 1.830, SD = 0.543) than in the generate-new stage (M = 2.311, SD = 0.706), F(1, 58) = 17.18, p < .001.

Manipulation checks

Requiring responses forced plagiarism?

Participants were required to type four responses in both the recall-own and the generate-new tasks, raising the concern that participants might knowingly plagiarize to complete the task (Tenpenny et al., 1998). If this were the case, then more plagiarism should be found in the fourth response to each puzzle than in the first response. There was, however, an equivalent number of plagiarized responses in the fourth response to all puzzles (M = 2.92, SD = 1.36) as there was in the first response (M = 2.58, SD = 1.52), F(1, 58) = 1.82, p = .183, and this pattern did not interact with accountability condition, F < 1, thus reducing this potential concern about our procedure.
Invalid responses
When participants typed a word in any task, they could not progress in the program until they entered an acceptable word solution. Collapsing across all three tasks, an independent-samples t-test revealed that, accountable participants entered fewer invalid words ($M = 5.27, SD = 6.60$) than control participants did ($M = 10.93, SD = 13.49$), $t(58) = 2.07, p = .045$. This pattern is consistent with accountable participants being more motivated than control participants were to follow task instructions.

DISCUSSION
The present study assessed whether motivation affected rates of unconscious plagiarism. Experimental participants heard that they would review their answers with the researcher after completing the Boggle task. This made the participants accountable for their performance, a situation people encounter in daily life. Accountable participants exhibited less cryptomnesia than control participants did when they were finding new solutions, that is, within both the initial-generation and generate-new tasks, but not when they were asked to remember their own words in the recall-own task. In a practical sense, this finding is encouraging in that unconscious plagiarism outside of the laboratory likely occurs more often when people are trying to generate new ideas, where the accountability manipulation reduced unconscious plagiarism, than when they are trying to recall their own contributions, where the accountability manipulation did not have a significant effect.

Although there are procedural differences across studies, we note that the present study replicates Marsh and Bower (1993, Experiment 2a) and Marsh et al. (1997, Experiment 4) in finding that an independent variable affected performance on generation tasks but not on a recall or recognition task (see also Perfect et al., 2009, for evidence that plagiarism errors on generation and recall tasks are influenced by different independent variables). We offer two explanations for the differential effect of accountability on generation and recall tasks.

One possibility is that the accountability manipulation led the experimental participants to more carefully encode, or attend to, the initial-generation experience than controls did. Indeed, this interpretation is consistent with Marsh and Bower’s (1993, Experiment 2a) study in which they manipulated how each word, whether computer or participant generated, was encoded. Participants made yes–no decisions about whether each word was four or more letters long (physical condition) or whether each word was “something good” (p. 679; semantic condition). Participants in the semantic condition showed less plagiarism than did participants in the physical condition for the initial-generation and generate-new phases but not for the recall-own phase. Marsh and Bower argued that this pattern arose because deeper encoding during the initial-generation phase strengthened the availability of both self-generated and other-generated words compared with new words. Therefore, for the semantic condition, all old words were easier to distinguish from new words (thus reducing plagiarism in the generation stages), but the manipulation did not aid distinctions between self-generated and computer-generated words. Our accountability manipulation may have led experimental participants to encode the initial-generation experience more deeply (e.g. Craik & Lockhart, 1972) than control participants did and thus to exhibit the same pattern of fewer plagiarisms in generation stages only.

On the other hand, if the accountability effect is due to encoding, moving the accountability manipulation to a later part of the procedure, after initial generation, should reduce its impact. To our knowledge, the study that comes closest to having done this is that of Marsh et al. (1997, Experiment 4). Their motivation manipulation, the warning quoted in the introduction combined with dividing participants into individual or group settings, occurred after initial generation was completed. After that, Marsh et al.’s (1997) participants completed two tasks. The first, called final generation, was analogous to the generate-new task in the current study. In the second, participants identified whether an idea was previously generated by them, previously generated by another, or not generated earlier (new). This recognition task, although procedurally different from the recall-own task in the current study, required participants to make self-other source-monitoring judgments such as the ones made during our recall-own task. Marsh et al. (1997) found that participants tested individually exhibited less plagiarism in the final-generation task than did participants tested in a group (cf. the accountable and control generate-new data, respectively, in Figure 1). By contrast, when participants had to make self-other judgments in the recognition task, the individual and group conditions’ scores were very similar (cf. the recall-own data in Figure 1). The Marsh et al. (1997) data indicate that a motivation manipulation combined with instructions on how to avoid plagiarism errors may lead to reduced cryptomnesia on a generation task even when the manipulation is not introduced until after initial generation is completed, thus undermining the encoding account. To fully test the encoding account of the present data, however, one would need to disentangle motivation and additional instructions while systematically varying where the motivation manipulation is introduced in the procedure.

There is, however, another reason to question the interpretation of the present data as due largely to deeper encoding of the initial-generation phase by accountable participants than by control participants. Macrae, Bodenhausen, and Calvini (1999) found that distraction during encoding, in the form of a radio program playing in the background, increased plagiarism errors in the recall-own phase but not in the initial-generation or generate-new phases. This pattern is inconsistent with the encoding interpretation of the present data because the group with relatively impoverished encoding in the present study (controls) did not show higher rates of cryptomnesia in the recall-own phase than our accountable participants did, nor did they show similar rates of cryptomnesia in the generate-new phase as accountable participants did.

A second account of the pattern in Figure 1 is rooted in the SMF’s proposal that people can vary the stringency of decision criteria when determining source (e.g. Johnson
et al., 1993). In the initial-generation and generate-new tasks, participants could avoid plagiarism by relying on familiarity or recent activation (relatively lenient decision criteria); as long as a word was entirely new to the experimental session, they would not be plagiarizing. During the recall-own task, however, participants could not rely on such information and instead had to distinguish between responses generated by themselves or the computer, that is, participants would need to rely on more source-specific information to avoid plagiarism (relatively stringent decision criteria). Note that the use of familiarity in the initial-generation and generate-new tasks does not make them simply old-new recognition tasks. For example, determining that can is old in this context is a source decision because it involves identifying that this familiar word came from a puzzle as opposed to being read recently. Thus, a second account of the present data is that motivation affected performance when relatively lenient decision criteria could be used (initial generation and generate new) but not when more stringent decision criteria were already required by the nature of the task (recall own).

This interpretation begs the question of why, if participants are using relatively stringent decision criteria in the recall-own phase, would plagiarism error rates appear relatively high in this phase (Figure 1). First, consider the recall-own other-plagiarism errors compared with the initial-generation errors (RO-O and IG-O&Self in Figure 1, respectively; RO-New errors are not plagiarism). Retention interval and proactive interference may contribute to higher plagiarism rates in the recall-own phase compared with the initial-generation phase. Initial-generation errors occurred during the first viewing of each puzzle. Plagiarism rates might be expected to be lower in this more immediate condition than the in the recall-own condition, which occurred after all critical puzzles had been viewed once and the practice puzzle had been viewed twice. Next, consider the recall-own plagiarisms compared with the generate-new plagiarisms. Importantly, the plagiarism rates in the recall-own phase are lower than the rates in the generate-new phase (the GN-O and GN-Self columns in Figure 1 need to be added together to obtain a picture of total plagiarisms in the generate-new phase). That being said, the plagiarism rates are relatively similar for recall-own other and generate-new other-plagiarisms (RO-O and GN-O in Figure 1, respectively). Notably, participants were significantly less confident in the plagiarized responses they produced in the recall-own phase than they were in the generate-new phase (this pattern held when the confidence analysis was restricted to only other-plagiarisms). This relatively reduced confidence in the recall-own plagiarisms is consistent with participants evaluating their memories particularly carefully in this phase, that is, that they used relatively stringent decision criteria in the recall-own phase. In other words, the use of stringent decision criteria does not mean that errors will necessarily be avoided, although they may be in some cases (cf. Lindsay & Johnson, 1989; Multhaup, De Leonardis, & Johnson, 1999).

The more general point is that motivation may affect performance when relatively lenient decision criteria could be used (generation tasks) but not when more stringent decision criteria are already required by the nature of the task (recall task). This could be due to motivation inspiring more careful examination of details stored in memory, and when that is already demanded by the task, such as the recall-own phase, there is no further effect that motivation can have on performance. If future studies continue to find that motivation has an effect on generation but not on recall-own tasks, this may suggest that there is a boundary condition on where motivation instructions will be effective.

The present study also replicated Marsh and Bower’s (1993) findings of (i) higher plagiarism rates on hard compared with easy puzzles on the generation tasks, (ii) no effect of puzzle difficulty on recall-own task plagiarisms, and (iii) higher recall-own new-intrusion rates for easy compared with hard puzzles. We agree with Marsh and Bower’s explanation that participants may have interpreted ease-of-word identification as a signal that they found the word in the puzzle in the initial-generation phase. This view accounts for the higher recall-own new-intrusion rate for easy puzzles, which had so many solutions available, than for hard puzzles which did not. The lack of a difficulty effect on the recall-own plagiarism errors may be explained by the initial-generation phase instructions that participants find computer-generated words in the puzzle before pressing a key to continue. Assuming the participants followed instructions, they would have found an equal number of words in the hard and easy puzzles before continuing, thus leaving old words from hard and easy puzzles similarly primed for ease of identification during the recall-own phase. Finally, we return to the higher plagiarism rates for hard versus easy puzzles for the generation tasks. When trying to find a new word in hard puzzles with fewer solutions than easy puzzles, participants may “relax their [generate-new] criterion . . . so that relatively more old . . . items would be accepted as novel and thus . . . plagiarized” (Marsh & Bower, 1993, p. 687).

Marsh and Bower’s (1993) account of difficulty effects is consistent with our interpretation of accountability effects arising when a task inspires relatively lenient decision criteria (generation tasks) but not when a task demands relatively strict decision criteria (recall-own task). Both accounts suggest that participants use general information, such as ease of identification or familiarity, on this puzzle task. This information is often sufficient to separate completely new solutions from prior solutions (all that is needed on the generation tasks) but cannot do so on the recall-own task, which requires more information to avoid plagiarisms. As outlined earlier, the use of ease of identification or familiarity also accounts for the effects of puzzle difficulty on plagiarism and recall-own new-intrusion errors. Further research may help us refine this explanation. For example, recording response times to solutions or procedures that use response deadlines (cf. Johnson, Kounios, & Reeder, 1994) may find that self-plagiarism and other-plagiarism arise on different timelines. Such data may yield clues about how people weight information such as ease of identification in puzzle solution or idea-generation tasks.

We have argued that the pattern of data in Figure 1 suggests that motivation induced by accountability for participants’ responses reduced plagiarism when relatively lenient decision criteria could be used to avoid plagiarism.
(initial-generation and generate-new tasks) but not when relatively stringent decision criteria were needed (recall-own task). Of course, the varying accounts of the data in Figure 1 are not mutually exclusive; there could be some effect of encoding and some effect of decision criteria. On the basis of the analysis above, we favor a decision-criteria interpretation of the present data. Further research is needed to clarify how motivation affects unconscious plagiarism. The SMF (e.g. Johnson et al., 1993; Lindsay, 2008; Mitchell & Johnson, 2009) may be particularly helpful in shaping future studies given that it accounts for memory errors in a range of paradigms.

The present research supports the assertion that source-monitoring errors, specifically reality-monitoring errors, underlie unconscious plagiarism (e.g. Brown & Murphy, 1989; Johnson et al.,1993; Stark & Perfect, 2008). Moreover, the present research demonstrates that source monitoring is affected by motivational factors (Johnson et al., 1993; Lindsay, 2008; Mitchell & Johnson, 2000, 2009). The key pattern of lower plagiarism rates for accountable compared with control participants did not interact with task difficulty level and is seen in plagiarism from multiple kinds of sources (e.g. see GN-O and GN-Self in Figure 1), pointing to the robustness of the finding. Our modification of Marsh and Bower’s (1993) procedure offers one method that (i) simulates everyday experiences in which a person’s work will be reviewed by a knowledgeable other, (ii) highlights accountability before the work begins, and (iii) avoids confounding accountability (motivation) and additional instructions on how to improve source-monitoring performance. Variations of this procedure can be used to further explore the mechanisms underlying the effect accountability has on rates of unconscious plagiarism.

**ACKNOWLEDGEMENTS**

Portions of this research were funded by an Abernethy Grant to BJW and a Faculty Study and Research Grant to KSM (both Davidson College). The data were presented at the 2010 Psychonomics Society meeting. We thank Hannah Bobrink for her help in collecting and scoring data, and the reviewers for helpful comments.

**REFERENCES**


