

Forced confabulation affects memory sensitivity as well as response bias

Victor Gombos · Kathy Pezdek · Kelly Haymond

Published online: 23 July 2011
© Psychonomic Society, Inc. 2011

Abstract A signal detection analysis assessed the extent to which forced confabulation results from a change in memory sensitivity (d_a), as well as response criterion (β). After viewing a crime video, participants answered 14 answerable and 6 unanswerable questions. Those in the *voluntary guess* condition had a “don’t know” response option; those in the *forced guess* condition did not. One week later, the same questions were answered using a recognition memory test that included each participant’s initial responses. As was predicted, on both answerable and unanswerable questions, participants in the forced guess condition had significantly lower response criteria than did those who voluntarily guessed. Furthermore, on both answerable and unanswerable questions, d_a scores were also significantly lower in the forced than in the voluntary guess condition. Thus, the forced confabulation effect is a real memory effect above and beyond the effects of response bias; forcing eyewitnesses to guess or speculate can actually change their memory.

Keywords Forced confabulation · Memory suggestibility · Eyewitness memory · False memory · Memory

The majority of jurors and judges consider eyewitness identification to be the most persuasive type of evidence

This research was conducted in the Department of Psychology, Claremont Graduate University, Claremont, CA. The work was supported by a grant from the Fletcher Jones Foundation to Kathy Pezdek and served as the M.A. thesis for Kelly Haymond. Correspondence should be directed to Kathy Pezdek, who can be contacted by email at Kathy.Pezdek@cgu.edu.

V. Gombos · K. Pezdek (✉) · K. Haymond
Claremont Graduate University,
Claremont, CA, USA
e-mail: kathy.pezdek@cgu.edu

administered in criminal cases (Wells & Olson, 2003). However, the forensic research on postconviction DNA exonerations has revealed that approximately 75% of these cases involve eyewitness misidentifications (<http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php>). Although these recent DNA exoneration cases are compelling, the fallibility of eyewitness memory is not a new subject among psychologists (Munsterberg, 1908).

It is well established that exposure to postevent suggestion following an observed event can affect eyewitness memory for the event (e.g., Loftus, 1975; Loftus, Miller, & Burns, 1978; Pezdek, 1977). This is referred to as *suggestibility*, or the *misinformation effect*. In a typical suggestibility study, participants observe an event as a slide sequence, video, or film. Participants in the *misled* condition are asked questions exposing them to misleading information related to the event; participants in the control condition are not exposed to the misleading information. On a subsequent recognition memory test for the observed event, participants’ false alarm rate to the misleading information is higher in the misled than in the control condition. Similarly, an intervening photographic lineup presented after an eyewitness has viewed an event suggestively decreases eyewitness identification accuracy on a subsequent lineup whether the eyewitness actually selects the innocent suspect in the initial lineup or not (Hinz & Pezdek, 2001; Pezdek & Blandon-Gitlin, 2005). Together, these findings demonstrate that one source of eyewitness error is postevent suggestion.

Despite growing empirical knowledge concerning eyewitness memory and improvements in practices within the law enforcement community, there are still a number of dubious interview procedures that may suggestively influence eyewitness memory. For example, if an interviewer believes an eyewitness observed an event, the interviewer

may press the eyewitness for information he or she admittedly cannot remember. In such interview situations, law enforcement officials are likely to encourage eyewitnesses to “just guess” or speculate about the answer to a question. In so doing, the interviewer might induce the eyewitness to self-generate misinformation. Although research has demonstrated that misleading questions and other types of postevent suggestion may introduce false information into memory, less is known about the influence of self-generated information.

In several studies, it has been reported that event memory can be suggestively influenced by forcing individuals to recall or speculate about answers to questions even if they are unsure of the correct answer (Ackil & Zaragoza, 1998; Hastie, Landsman, & Loftus, 1978; Roediger, Wheeler, & Rajaram, 1993; Schreiber, Wentura, & Bilsky, 2001; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001). When an individual erroneously incorporates into his or her memory for an event self-generated information that was not actually part of that event and, subsequently, recalls or recognizes this information, this effect is known as *forced confabulation* (Pezdek, 2008). However, only the recent study by Pezdek, Sperry, and Owens (2007) assessed whether information obtained by forced confabulation is as likely to persist in memory as information voluntarily self-generated by participants. That is, will individuals who were forced to guess the response to a question at time 1 subsequently be more likely to repeat their initial response than individuals who voluntarily guessed? Participants in that study were shown a video of a crime and then, using a procedure similar to that in Ackil and Zaragoza (1998), were given 16 answerable and 16 unanswerable questions immediately afterward and again 1 week later. Unanswerable questions pertained to information that was plausible but not actually presented in the video. At time 1, participants in the *voluntary guess* condition were given the response option “don’t know,” so they had to answer only questions for which they thought that they knew the answer. Participants in the *forced guess* condition were required to answer each question at time 1; they were not given the “don’t know” response option. We call the control condition in this research the *voluntary guess* condition, recognizing that some degree of response guessing is always present even when individuals are not forced to do so. One week after viewing the video, the same questions were answered with a “don’t know” response option available to all participants. Pezdek et al. (2007) reported that information generated from forced confabulations was less likely to be remembered than information voluntarily self-generated when participants were not forced to guess. Additionally, when given a chance at time 2 to respond “don’t know,” individuals were more likely to shift and respond “don’t know,” rather than give their time 1

response, in the forced than in the voluntary guess condition. Thus, although forced confabulation does occur, information that results from forced confabulation is less likely to persist in memory than information that individuals voluntarily provided because they thought that they had observed it.

Pezdek, Lam, and Sperry (2009) extended this research and compared the relative impact of *self-generated* misinformation (i.e., information fabricated by the participant, such as postevent reflection), as compared with *other-generated* misinformation (i.e., information fabricated by another person, such as an interviewer’s questions about an event) on confabulated event memory. The major results of this previous study concerned responses to unanswerable questions for which we knew that any answers provided were confabulated. In this condition, other-generated forced confabulation information was more likely to be integrated into memory and repeated at time 2 than was forced confabulation information that was self-generated by the participant. However, in both of these studies, it is unclear whether this forced confabulation effect resulted from a change in memory, as well as a change in response bias.

The present study builds upon this previous research and assesses the cognitive processes underlying the forced confabulation effect—specifically, the extent to which forced confabulation results from a change in memory sensitivity or simply response criterion (also known as response *bias*). Both memory sensitivity (d_a) and response criterion (β) are components of classical signal detection theory. On the basis of this framework, the memory on which individuals base their answers is conceptualized as consisting of some information that was observed in the video (i.e., signal) and some related schematic information not actually presented (i.e., noise). An eyewitness’s response to a question will depend on the strength of the two distributions relative to their response criterion. According to this framework, forcing individuals to guess answers to questions would have the effect of lowering their response criterion at time 1. With a higher response criterion, individuals approach decisions more carefully; with a lower response criterion, individuals are less cautious about their decisions.¹ This study tests whether forced confabulation affects d_a as well as β —that is whether it produces an actual change in event memory.

¹ This assumes a single-process, unidimensional detection theory representation of memory. Although various models of memory have been proposed (e.g., multidimensional and classification designs explained in Macmillan & Creelman, 2005), we based our predictions on the simplest unidimensional assumption. In this model, three distributions lie along a single memory-strength axis (targets with highest signal strength, self-generated nontargets with intermediate strength, and other-generated nontargets with least strength).

Method

Participants and design

Participants were 102 volunteers from five introductory psychology classes at a state university and a community college in the Los Angeles metropolitan area. This was a 2 (forced guess vs. voluntary guess condition) \times 2 (time 1 vs. time 2) mixed factorial design with the first variable manipulated between subjects.

Procedure and materials

Participants participated in two sessions lasting about 15 min each. In the first session, they viewed a 5-min video of a carjacking, the same video as that used by Pezdek et al. (2007). They were told to pay close attention to the video because, afterward, they would be asked some questions about it. The video was followed immediately by 20 open-ended questions including (1) 14 answerable questions about information in the video and (2) 6 unanswerable questions. There was a correct answer for each answerable question. Unanswerable questions prompted information not actually presented in the video. An example of an unanswerable question is “What was the logo on the perpetrator’s shirt?” when the perpetrator’s shirt did not have a logo. Questions were presented in the same order to all participants and followed the chronological order in which the relevant information appeared (or might have appeared) in the video.

Participants in the forced guess condition ($n = 52$) were instructed to answer all 20 questions, not to leave any questions unanswered, and to make their best guess of each answer even if they were unsure; they were not given an “I don’t know” response option at time 1. Participants in the voluntary guess condition ($n = 50$) were instructed to answer each question; however, they were told that if they did not know an answer, they should circle the “I don’t know” response option.

One week later at time 2, participants were retested on memory for the video, this time with an 80-question yes/no recognition test. A unique recognition test was created for each participant on the basis of his or her time 1 responses. Each of the original 20 questions from time 1 was converted into 4 questions—one containing their answer to the time 1 question and 3 distractor questions containing the three most frequent responses to each question given by participants in Pezdek et al. (2007). For example, the question “What color were the perpetrator’s shoes?” became “Were the perpetrator’s shoes black?” “Were the perpetrator’s shoes white?” “Were the perpetrator’s shoes red?” “Were the perpetrator’s shoes brown?”—with each participant’s specific time 1 response included in one of the

questions. The four versions of each question were randomly placed throughout the recognition test to avoid comparison responses. In response to each of the 80 yes/no questions, participants also rated their confidence in the accuracy of their response on a 1 (*low*) to 5 (*high*) scale.

Results

There were two exclusion criteria, consistent with the analyses in Pezdek et al. (2007). First, 7 participants (3 from the forced guess condition and 4 from the voluntary guess condition) were excluded from analyses because, at time 1, they did not respond with one of the multiple choice answers provided but, rather, indicated that at least one unanswerable item was “not present” in the video. Because these responses did not conform to instructions, the time 1 and time 2 data from these participants were not analyzed. Second, a participant in the forced guess condition was excluded due to failure to follow instructions (i.e., the participant did not provide any answer to at least five of the six unanswerable questions at time 1). Analyses were performed on data from the remaining 94 participants (age, $M = 23.16$ years, $SD = 3.29$; 22 males and 68 females, 4 unspecified).

Analyses focused on the differences in response criterion (β) and memory sensitivity (d_a) between the forced guess and voluntary guess conditions. The time 2 yes/no recognition memory responses and 5-point confidence ratings were transformed into a 10-point ordinal scale within subjects for signal detection analyses in SYSTAT (v. 9.0) in accordance with that software’s guidelines for rating experiments. Specifically, large numbers along the scale represent positive recognition responses with high degrees of confidence (e.g., a “yes” response plus a 5 confidence rating = a final value of 10). Small numbers represent negative responses with high degrees of confidence (e.g., a “no” response plus a 5 confidence rating = a final value of 1). The middle values on the 10-point scale represented recognition responses with low ratings of confidence. These data were entered as responses in SYSTAT using default model estimations (Gaussian function set at convergence = .001 for 50 iterations) that calculated parametric signal detection measures and receiver operating characteristic (ROC) curves (including the nonparametric area under the ROC, A_g) for each condition for each participant. From the various sensitivity measures generated by SYSTAT, d_a was chosen because of its relative robustness in situations with unequal-variance distributions (Rotello, Masson, & Verde, 2008). The measure d_a is a measure of sensitivity that reflects the area under the ROC curve, calculated as the root mean square of the standard deviations (SD s) of signal and noise distributions (Banks, 1970; Macmillan & Creelman,

2005; Simpson & Fitter, 1973). This measure of sensitivity is relatively conservative because, unlike d' , the measure d_a does not assume that the underlying evidence distributions have equal variance. Analyses addressed responses to both unanswerable questions and answerable questions; however, each question type was analyzed separately because each contained unique signal and distractor response options.

Responses to answerable questions

For answerable questions, the signal distribution represents presented information, and the noise distribution represents the three other distractor items. If participants gave a correct answer to a question at time 1, then at time 2, the four recognition test items for that question included the correct answer and, as distractor items, the top three responses provided by participants in the study by Pezdek et al. (2007).² However, if participants gave an incorrect answer to a question at time 1, then at time 2, the four recognition test items included the correct answer, the participant's self-generated incorrect time 1 response, and the top two responses for each item from Pezdek et al. (2007). False alarms to other-generated and self-generated responses represent two different underlying cognitive processes. Therefore, answerable questions were analyzed separately for correct and incorrect responses at time 1 and, for incorrect responses, whether sensitivity differed for self-versus other-generated items. β and d_a were obtained for each response type.

How does memory sensitivity for answerable questions compare in the forced guess versus voluntary guess conditions? It was hypothesized that forcing an eyewitness to guess would lower their response criterion (β). A lower response criterion would correspond to less cautious responding. This would consequently hinder participants' ability to distinguish between presented information and schematically relevant information not actually presented, resulting in a lower d_a . It was also hypothesized that participants who gave an incorrect answer at time 1 would have a lower d_a than would those who gave a correct response, because they would false alarm to their own incorrect response from time 1. Analyses of these two types of responses to answerable questions follow.

Time 1 responses Presented in Table 1 is a summary of the responses to answerable questions participants gave at time 1 in each condition. From SDT analyses based on correct time 1 responses, averaged SD ratios for each condition were calculated with SYSTAT to confirm that distributions

Table 1 Proportion of time 1 responses for answerable questions

Response Type	Proportion of Responses
Voluntary Guess Condition	
Correct	.28
Incorrect	.33
Don't know	.39
Forced Guess Condition	
Correct	.49
Incorrect	.51

Correct responses refer to accurate identifications of target stimuli at time 1, and incorrect responses refer to misidentifications at time 1. "Don't know" responses were allowed in the voluntary guess condition only

were close to normal. SD ratios are the standard deviation of the signal + noise distributions; when these are equal to 1, normal shapes are assumed (see "Normal Distribution Model for Signal Detection" explanatory notes in SYSTAT). As was predicted, SD ratios were close to 1.0: voluntary answerable questions at time 1 = 1.17, forced answerable = 1.32, voluntary unanswerable = 1.09, and forced unanswerable = 1.10.

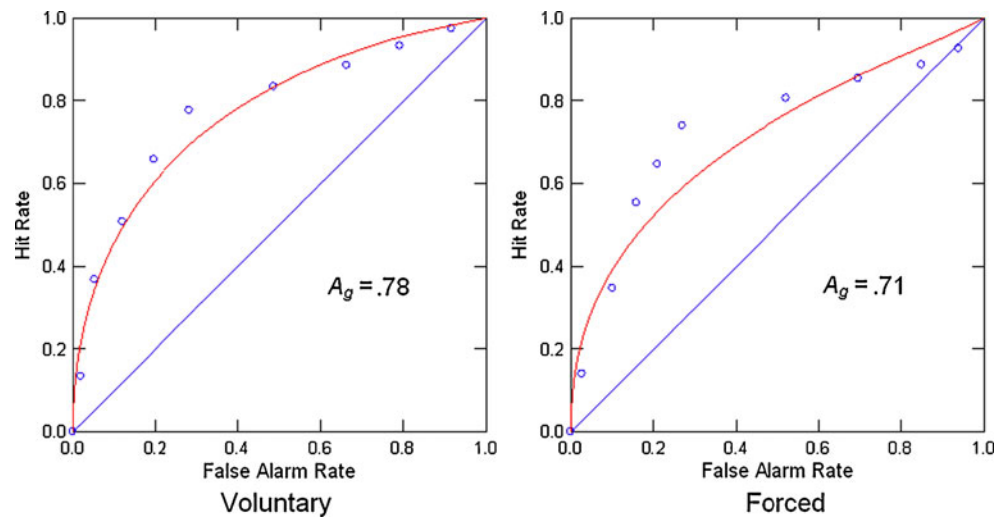
Correct responses to answerable questions The first analysis examines the time 2 responses to answerable questions for which each participant gave a correct answer at time 1. If, at time 1, a participant gave a correct answer, their time 2 recognition questions contained their answer from time 1 (from which the hit rates were derived) and three distractor options (from which the false alarm rates were derived). Results confirmed the hypothesis that participants who voluntarily guessed produced a higher mean response criterion (β) and d_a than did participants forced to guess (see Table 2). The mean response criterion (β) was significantly higher in the voluntary guess than in the forced guess condition, $t(92) = 4.93, p < .001, d = 1.01$, and the mean d_a rate was significantly higher in the voluntary guess than in the forced guess condition, $t(92) = 1.99, p < .05, d = 0.41$. The ROCs corresponding to the data presented in Table 2 are presented in Fig. 1. In Fig. 1, 9 points out of the

Table 2 Mean time 2 response criterion (β) and memory sensitivity (d_a) (with standard deviations) for responses to answerable questions with a correct response provided at time 1 for participants in the voluntary and forced guess conditions

Answerable Questions: Correct Response at Time 1	Voluntary Guess	Forced Guess
	Condition ($n = 50$)	Condition ($n = 44$)
Response criterion (β)	0.97 (0.62)	0.39 (0.52)
Memory sensitivity (d_a)	1.03 (0.47)	0.83 (0.51)

² In all analyses, the false alarm rates and, consequently, the d_a and β values were calculated using the same number of lure alternatives at time 2 for both self-generated and other-generated conditions.

Fig. 1 ROCs corresponding to the voluntary guess and forced guess conditions for answerable questions with correct responses provided at time 1 (reported in Table 2)



10 possible ratings (1 rating is purposely left out for $n-1$ degrees of freedom) indicate hit rates and false alarm rates at each criterion. In this figure, greater area under the ROC (A_g) for voluntary guessing indicates better discrimination, as compared with forced guessing. Furthermore, the general pattern indicates that hit rates were slightly higher in the voluntary than in the forced guess condition, and false alarm rates were higher in the forced guess condition than in the voluntary guess condition.

Incorrect response to answerable questions The next analysis examines the time 2 responses to answerable questions for which each participant gave an incorrect answer, rather than the correct answer, at time 1. These questions are particularly interesting because time 2 response options contained three types of information: (1) the correct answer, (2) each participant’s incorrect answer from time 1, and (3) two distractor responses that were the two most frequent incorrect responses by participants in the study by Pezdek et al. (2007). Results of the analyses of these data are presented in Table 3. As was predicted, participants’ mean response criteria (β) were significantly higher in the voluntary guess than in the forced guess condition, $t(92) = 3.96, p < .001, d = 0.79$. In addition,

participants in the voluntary guess condition had a significantly higher d_a than did participants in the forced guess condition, $t(92) = 2.70, p < .01, d = 0.55$. The ROCs corresponding to the data presented in Table 3 are presented in Fig. 2. As can be seen in Fig. 2, while hit rates, on average, were equivalent between conditions, consistently higher false alarm rates at each criterion point were generated in the forced guess than in the voluntary guess condition.

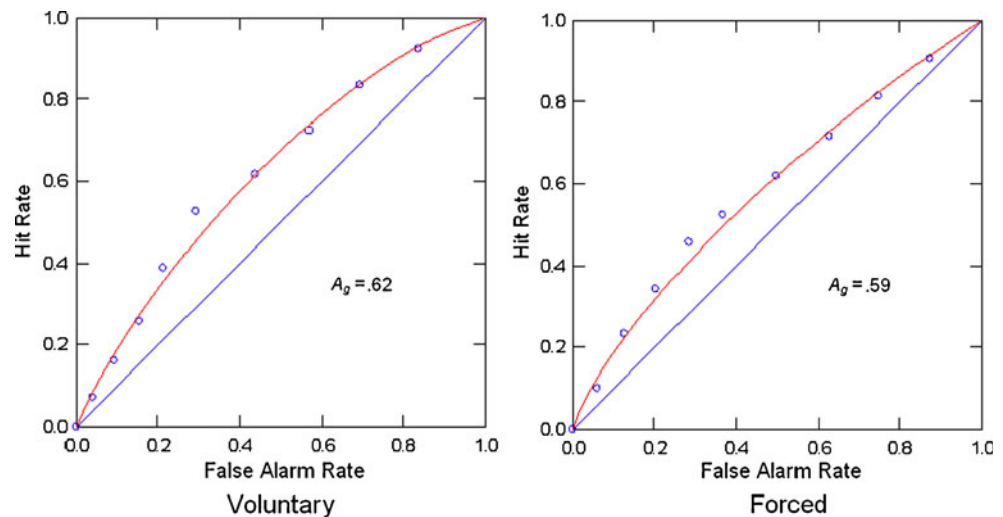
For answerable questions, as indicated above, if participants provided an incorrect response at time 1, this incorrect response became one of the distractor items on their time 2 recognition test. This produced two different types of false alarms at time 2: (1) incorrectly attributing to the video the self-generated incorrect response provided at time 1 and (2) incorrectly attributing to the video the other-generated distractors that were the two most frequent incorrect responses by participants in the study by Pezdek et al. (2007). The next analyses examine whether participants were less able to discriminate between the correct and the incorrect response at time 2 if the incorrect response (i.e., false alarm) was (1) one that they self-generated at time 1 rather than (2) one that was other-generated at time 1. Because only a subset of responses was included in this analysis, to obtain a sufficient sample size, this analysis was conducted across the voluntary and forced guessed conditions together; no differences in the results of this analysis were predicted for answerable questions across these conditions.

The mean false alarm rate was significantly higher for self-generated distractors ($M = .39$) than for other-generated distractors ($M = .26$), $t(93) = 3.90, p < .001, d = 0.49$. Correspondingly, the mean d_a for items for which participants false alarmed to self-generated distractors ($M = .21$) was significantly lower than the mean d_a for items for which participants false alarmed to one of the other-generated distractors ($M = .33$), $t(93) = 4.26, p < .001, d = 0.52$. These

Table 3 Mean time 2 response criterion (β) and memory sensitivity (d_a) (with standard deviations) for responses to answerable questions with an incorrect response provide at time 1 for participants in the voluntary and forced guess conditions

	Answerable Questions: Incorrect Response at Time 1	
	Voluntary Guess Condition ($n = 50$)	Forced Guess Condition ($n = 44$)
Response criterion (β)	0.83 (0.34)	0.44 (0.60)
Memory sensitivity (d_a)	0.51 (0.30)	0.35 (0.28)

Fig. 2 ROCs corresponding to the voluntary guess and forced guess conditions for answerable questions with incorrect responses provided at time 1 (reported in Table 3)



results confirm that at time 2, participants were less able to discriminate between the target items presented in the video and their own self-generated time 1 responses than between the target items presented in the video and other-generated time 1 responses. Thus, when individuals speculate about an event, this has a detrimental effect on event memory, and it is specifically their self-generated responses that are more likely to hinder subsequent event memory.

Incorrect versus correct responses to answerable questions at time 1 To test the hypothesis that participants who gave an incorrect answer at time 1 would have a lower d_a at time 2 than would those who gave a correct time 1 response, a 2 (voluntary guess vs. forced guess condition) \times 2 (incorrect vs. correct response at time 1) ANOVA was performed. Participants who gave a correct response at time 1 had a significantly higher d_a ($M = 0.94$) at time 2 than did those who gave an incorrect response at time 1 ($M = 0.44$), $F(1, 90) = 72.29$, $p < .001$, $\eta^2 = .28$. Furthermore, although the mean d_a rate was significantly higher in the voluntary guess ($M = 0.77$) than in the forced guess ($M = 0.59$) condition, $F(1, 90) = 9.50$, $p < .01$, $\eta^2 = .05$, the interaction of these variables was not significant, $F(1, 90) = 0.74$, $p > .05$, $\eta^2 = .01$. These results confirm that participants who responded with an incorrect answer at time 1 had lower d_a rates than did those who gave a correct response at time 1, and we know from the ROCs presented in Fig. 2, as well as from previous results, that this is primarily attributable to the higher false alarm rate and lower d_a rate at time 2 to self-generated incorrect responses from time 1.

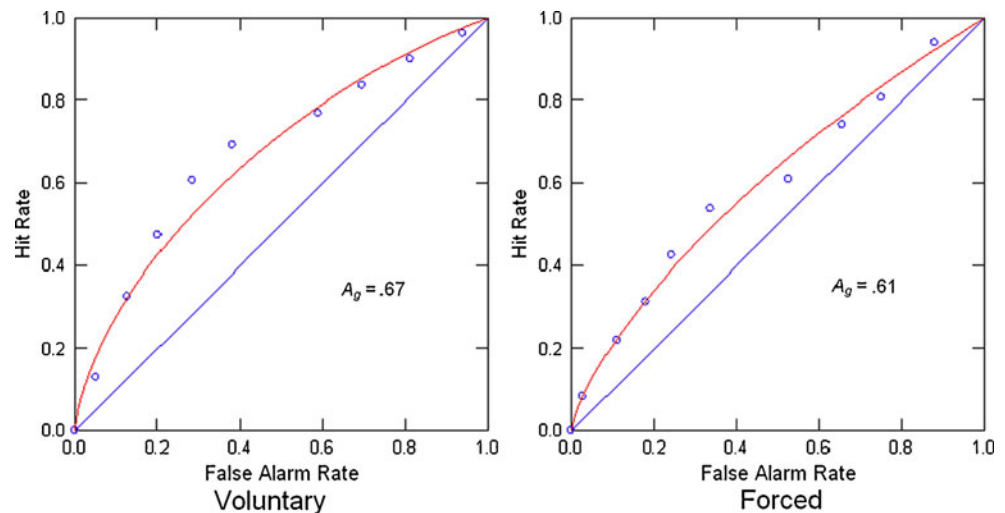
Table 4 Mean time 2 response criterion (β) and memory sensitivity (d_a) (and standard deviations) for responses to unanswerable questions for participants in the voluntary and forced guess conditions

	Unanswerable Questions	
	Voluntary Guess Condition ($n = 50$)	Forced Guess Condition ($n = 44$)
Response criterion (β)	1.16 (0.74)	0.37 (0.40)
Memory sensitivity (d_a)	0.56 (0.28)	0.40 (0.24)

Responses to unanswerable questions

For unanswerable questions, there was no true signal in the video because the information necessary to answer each question was not actually presented. Therefore, each participant's time 1 response to each question was coded as a signal (from which the hit rates were calculated). The three distractor items (from which the false alarm rates were calculated) were the top three participant-provided responses for each question from Pezdek et al. (2007). The analyses of responses to unanswerable questions reveal the extent to which each participant's forced answers provided at time 1 persisted in their memory at time 2. It was predicted that the availability of the "don't know" response option at time 1 in the voluntary guess condition would produce a higher mean response criterion, corresponding to more cautious responding. With a higher mean response criterion, answers generated at time 1 were likely to be based on higher strength memories than were those generated in the forced guess condition. The analyses of responses to unanswerable questions confirm this prediction (see Table 4). The response criterion (β) levels were significantly higher in the voluntary guess than in the forced guess condition, $t(92) = 6.27$, $p < .001$, $d = 1.29$. Furthermore, the mean d_a rate was significantly higher in the voluntary guess than in the forced guess condition, $t(92) = 2.98$, $p < .01$, $d = 0.61$. Given the way these d_a values were computed, this finding means that participants in the voluntary guess condition were more likely to repeat their time 1 response at time 2 than were those in the forced guess condition. These findings are

Fig. 3 ROCs corresponding to the voluntary guess and forced guess conditions for unanswerable questions (reported in Table 4)



consistent with results of Pezdek et al. (2009) that individuals discriminate better between self- and other-generated items in the voluntary than in forced guess condition. The ROCs corresponding to the data presented in Table 4 are presented in Fig. 3. As can be seen in Fig. 3, voluntary guessing is characterized by higher hit rates and false alarm rates than in the forced guessing condition. The higher level of sensitivity (based on d_a or A_g) in voluntary guessing, as compared with forced guessing, is attributable to the larger difference in hit rates than in false alarm rates.

Discussion

This study assessed the cognitive processes underlying the forced confabulation effect. The results confirmed earlier findings of Pezdek et al. (2007) and Pezdek et al. (2009) that information generated from voluntary confabulation is more likely to be remembered than forced self-generated confabulated information. In all three studies, information resulting from forced confabulation at time 1 was less likely to be recalled at time 2 than information that individuals voluntarily provided. This study extended the findings to determine that the forced confabulation effect results both from a change in response bias (β) and, more significant, from changes in memory sensitivity (d_a).³

Although the results indicated significantly higher response criteria (β) and memory sensitivity (d_a) in the voluntary guess than in the forced guess conditions for

both answerable and unanswerable questions, the results for answerable questions should be interpreted differently than those for unanswerable questions. While results for unanswerable questions provide information on the impact of confabulation on information that an eyewitness never actually saw, results for answerable questions pertain to memory for information that was actually presented.

How does forced confabulation affect memory sensitivity? For answerable questions, for which the information necessary to answer the question was presented in the target event, individuals may be reluctant to provide an answer at time 1 because the signal strength of the memory is relatively weak or their response criterion is too high. Forcing the individual to guess lowers his or her response criterion, reflecting less cautious responding, and results in a higher hit rate to presented information. However, information that was forcibly guessed is less likely to be remembered on subsequent memory tests because it resulted from memories of average lower strength than those produced in the voluntary guess condition. In contrast, individuals who voluntarily provide answers at time 1 have a higher response criterion, reflecting more cautious responding, and are more likely to recognize their volunteered answers on a subsequent memory test because of the higher strength of this information in memory. Furthermore, memory sensitivity depends on whether the answer provided initially was correct or incorrect. If an individual initially provides a correct answer to an answerable question, he or she is likely to recognize that answer on a subsequent test. However, if an individual provides an incorrect response to an answerable question at time 1, this reduces his or her ability to distinguish subsequently between information actually presented and nonpresented information, especially if the incorrect time 1 response was self- rather than other-generated.

³ Differences exist between the findings of this study and those in Pezdek et al. (2009). These can be attributed to several methodological differences between these studies—primarily, the fact that whereas in the previous study, other-generated items included information that participants actually read in the time 1 test questions themselves, in the present study, other-generated items were items participants never actually experienced.

On the other hand, responses to unanswerable questions are generated from schematically related information, rather than from information actually presented. This study demonstrates that when confabulated answers to unanswerable questions are generated, they can become integrated into participants' memory and then incorrectly recognized on a subsequent memory test. Forcing an eyewitness to guess an answer to an unanswerable question lowers his or her response criterion, which reflects less cautious responding and increases the rate of false confabulations. However, the confabulated information is more likely to be repeated on subsequent memory tests in the voluntary than in the forced guess condition, either because, at time 1, this information was more strongly associated with information in the presented event or because the source-monitoring error that produced the response at time 1 is likely to be repeated in the recognition memory test at time 2.

It is clear that manipulating the response criterion of eyewitnesses by forcing them to guess significantly impacts the veracity of the information they provide. Pressuring eyewitnesses to answer questions when they indicate that they do not know the answer can be viewed as a double-edged sword. If a question is answerable—the eyewitnesses really did see some information about the target event, but the signal strength of the memory is relatively weak or the individuals' response criterion is too high—then forcing the eyewitnesses to guess might reveal more correct information from event memory. In many cases, this is probably the assumption of investigating officers interviewing reluctant witnesses. However, if the question is unanswerable—the eyewitnesses really never did see any information pertaining to the target event, but the investigating officer presses the eyewitnesses to respond because they believe that they are holding back information—then forcing the eyewitnesses to speculate can elicit erroneous information that nonetheless becomes integrated into the eyewitnesses' memory, subsequently reducing their ability to distinguish correct information from falsely speculated about information. The results of this study suggest that forcing eyewitnesses to guess or speculate does more than just influence how they respond to questions; it can result in real changes in memory.

References

- Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: age differences in susceptibility to false memories. *Developmental Psychology, 34*, 1358–1372.
- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin, 74*, 81–99.
- Hastie, R., Landsman, R., & Loftus, E. F. (1978). Eyewitness testimony: the dangers of guessing. *Jurimetrics Journal, 19*, 1–8.
- Hinz, T., & Pezdek, K. (2001). The effect of exposure to multiple lineups on face identification accuracy. *Law and Human Behavior, 25*, 185–198.
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560–572.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah: Erlbaum.
- Munsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. Garden City: Doubleday.
- Pezdek, K. (1977). Cross-modality semantic integration of sentence and picture memory. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 515–524.
- Pezdek, K. (2008). Forced confabulation. In B. L. Cutler (Ed.), *Encyclopedia of psychology and law* (pp. 324–325). Thousand Oaks: Sage.
- Pezdek, K., & Blandon-Gitlin, I. (2005). When is an intervening lineup most likely to affect eyewitness identification accuracy? *Legal and Criminological Psychology, 10*, 247–263.
- Pezdek, K., Lam, S. T., & Sperry, K. (2009). Forced confabulation more strongly influences event memory if suggestions are other-generated than self-generated. *Legal and Criminological Psychology, 14*, 241–252.
- Pezdek, K., Sperry, K., & Owens, S. M. (2007). Interviewing witnesses: the effect of forced confabulation on event memory. *Law & Human Behavior, 31*, 463–478.
- Roediger, H. L., III, Wheeler, M. A., & Rajaram, S. (1993). Remembering, knowing, and reconstructing the past. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 97–134). San Diego: Academic.
- Rotello, C. M., Masson, M. E. J., & Verde, M. F. (2008). Type I error rates and power analyses for single-point sensitivity measures. *Perception & Psychophysics, 70*, 389–401.
- Schreiber, N., Wentura, D., & Bilsky, W. (2001). What else could he have done? Creating false answers in child witnesses by inviting speculation. *Journal of Applied Psychology, 86*, 525–532.
- Simpson, A. J., & Fitter, M. J. (1973). What is the best index of detectability? *Psychological Bulletin, 80*, 481–488.
- Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54*, 277–295.
- Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: forced confabulation and confirmatory feedback increase false memories. *Psychological Science, 12*, 473–477.