American Psychological Association

Journal of Experimental Psychology: General

ISSN: 0096-3445

2020, Vol. 149, No. 7, 1231–1248
http://dx.doi.org/10.1037/xge0000695

Translating (Lack of) Memories Into Reports: Conversion Processes in Responding to Unanswerable Questions

Aleksandra Krogulska
University of Warwick

Zuzanna Skóra
Jagiellonian University

Alan Scoboria
University of Windsor

Maciej Hanczakowski and Katarzyna Zawadzka
SWPS University

When queried about events in the past, a person may face questions that concern details that have been witnessed—answerable questions—and details that have not been witnessed—unanswerable questions. With regard to answerable questions, the person's willingness to answer these questions increases as a function of not only information available about the queried detail itself, but also as a function of contextual information. The present research assesses whether the willingness to report specific—and thus incorrect—answers when facing unanswerable questions also increases with the amount of available contextual information. In 3 experiments, we show that when recognition questions for critical details one had not encoded are preceded by reinstated contexts, participants are less willing to respond "don't know" to these questions, thus making more commission errors. These results show how greater access to contextual information, commonly associated with better memory for answerable questions, can also lead to more incorrect responses in the case of unanswerable questions. This documents how conversion processes—metacognitive processes of monitoring retrieval from memory and controlling the quality of output—play an important role in shaping the accuracy of memory reports.

Keywords: context, conversion, memory reports, unanswerable questions

Situations involving reporting from memory often require judging whether another person's recollections of the past are accurate. Even the simplest report contains at least two indicators of memory quality. The number of correct details described by a person speaks to the quality of the report. Equally, or perhaps even more important is the number of incorrect details reported by a person (Borckardt, Sprohge, & Nash, 2003; Tenney, MacCoun, Spellman, & Hastie, 2007). Errors, intrusions, and confabulations are indicators of unreliable memory as correct memory generally serves to reduce the incidence of errors (e.g., Gallo, Weiss, & Schacter, 2004; Pierce, Gallo, & McCain, 2017). In this study, however, we examine a situation in which errors don't stem from unreliable memory for the past event. On the contrary, here we aim to demonstrate that under specific conditions, participants report erroneous responses exactly *because* they remember more rather than fewer correct details from studied materials.

Conversations about past events commonly create conditions of informational asymmetry when one person possesses some information that the other person does not know. The purpose of the conversation is then to share this information. We argue that such informational asymmetries in social sharing of memories are likely to result in asking, and thus responding to, unanswerable questions—that is, questions for which a responder cannot know the correct answer (Roebers & Fernandez, 2002; Scoboria, Mazzoni, & Kirsch, 2008; Waterman & Blades, 2011; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001). In this study we aim to demonstrate that when responding to unanswerable questions, better memory for contextual details can result in less accurate memory reports. This will serve the purpose of highlighting two key issues that so far have received relatively little conceptual or empirical attention. First, it aims to show an important contribution of conversion processes to memory reporting—the processes that are responsible for turning the products of memory retrieval into actual responses to memory questions that together constitute

This article was published Online First November 11, 2019.

Aleksandra Krogulska, Department of Psychology, University of Warwick; Zuzanna Skóra, Institute of Psychology, Jagiellonian University; Alan Scoboria, Department of Psychology, University of Windsor; Maciej Hanczakowski and Katarzyna Zawadzka, Interdisciplinary Center for Applied Cognitive Studies, SWPS University.

Alan Scoboria passed away on April 19, 2019. He will be sorely missed.

Portions of this research were presented at the 56th Annual Meeting of the Psychonomic Society, Chicago, November 19–22, 2015, the International Conference on Memory 6, Budapest, July 17–22, 2016, the 59th Annual Meeting of the Psychonomic Society, New Orleans, November 15–18, 2018, and SARMAC XIII, Brewster, June 6–9, 2019.

This work was supported by National Science Centre (Poland) Grant 2013/09/B/HS6/02642 awarded to Maciej Hanczakowski. All pictorial materials used in this study can be found at https://osf.io/37cdy/.

Correspondence concerning this article should be addressed to Katarzyna Zawadzka, Interdisciplinary Center for Applied Cognitive Studies, SWPS University, ul. Chodakowska 19/31, 03-815 Warszawa, Poland. E-mail: kzawadzka@swps.edu.pl

memory reports. Second, it highlights the importance of considering unanswerable questions—and costs related to providing specific responses to such questions—when assessing techniques designed to improve memory retrieval, as often used, for example, in the context of eyewitness testimony.

Metacognitive Processes at Retrieval

It is a widespread assumption that people's responses to memory queries directly reflect their ability to remember the past, in which case successful recall points to good memory and errors point to poor memory. However, as pointed out by Tulving (1983; see also Blank, 1998), memory responses included in final memory reports cannot be taken as direct products of retrieval from memory. Before retrieved responses are volunteered, a psychological process of *conversion* of retrieved information into a response needs to take place. Conversion depends on factors such as the quality of retrieved information as appraised by the rememberer, the understanding of the task, social conventions of memory responding, incentives for volunteering correct information, costs of making errors, and so forth. It is vital to observe that the contents of memory reports—and consequently the accuracy of reported information—depend not only on the quality of retrieval but also on how retrieved information is appraised by the rememberer and how it is ultimately communicated in the form of responses to memory queries.

Conversion processes have been studied under the umbrella of metacognitive monitoring and control of memory reporting, with the main focus on processes determining whether information will be volunteered to maximize the informativeness of reports, or whether it will be withheld to protect the reports' accuracy (e.g., Ackerman & Goldsmith, 2008; Goldsmith, Koriat, & Pansky, 2005; Goldsmith, Koriat, & Weinberg-Eliezer, 2002; Higham, 2007; Zawadzka, Krogulska, Button, Higham, & Hanczakowski, 2016). In an influential framework of strategic regulation of memory accuracy, Koriat and Goldsmith (1996b) outlined how the interplay of retrieval and conversion jointly shapes the accuracy of memory reports. In their framework, the first step in responding to a memory query involves the generation of candidate responses—a result of retrieval processes after which conversion takes over. Candidate responses are subjected to appraisal—metacognitive monitoring—by which the rememberer strives to establish whether a potential response is accurate. The products of the monitoring process feed into metacognitive control processes by which memory reporting is regulated. When confidence in the accuracy of the best candidate response is higher than the report criterion adopted for a given retrieval task, the response is volunteered (output). When confidence is lower than the report criterion, the candidate response is withheld and a socially acceptable indicator of the lack of specific response (such as "don't know") might be given instead. In this framework, both the contents and the accuracy of memory reports depend not only on the quality of retrieval which shapes candidate responses, but also on the effectiveness of monitoring processes and the adopted reporting policy–conversion processes responsible for translating candidate responses into overt responses included in a memory report.

Koriat and Goldsmith's (1996b) framework has been influential in driving the attention of memory researchers toward the issue of "don't know" responses as reflecting not solely failures of memory but also a host of conversion processes. In their work, Koriat and Goldsmith outlined two different outlooks on memory. According to what they termed a *storehouse* metaphor of memory, which they consider to be the traditional approach adopted by memory researchers, the quality of memory functioning can be assessed by evaluating *how much* of the information originally put into a memory store can be retrieved at a later time (see also Koriat & Goldsmith, 1994, 1996a, for related discussion). From this perspective, responding "don't know" to a memory query is a missed opportunity for providing a piece of originally stored—but perhaps forgotten—information, and thus can only be treated as a failure of memory. By contrast, the framework of strategic regulation of memory accuracy is based on a *correspondence* metaphor of memory, which focuses more on the *accuracy* rather than quantity of reported memories. This approach underscores an important adaptive role of "don't know" responses. These responses may reflect not only failures to remember, but also one's strategic effort to limit the incidence of errors by withholding candidate responses of potentially dubious quality—ones for which metacognitive monitoring suggests low likelihood of any candidate response being correct.

Answerable and Unanswerable Questions

The discussion of the storehouse metaphor of memory and "don't know" responses, inherent to Koriat and Goldsmith's (1996b) framework, highlights another theoretically and practically important issue. The understanding of memory as a storehouse not only downplays the adaptive role of responding "don't know" when candidate responses are perceived to be of low quality, but it also largely ignores everyday situations in which "don't know" *is* the correct response to a given question. *Unanswerable questions* are questions for which the addressee of the question cannot possess information that would allow for formulating the correct response based on direct retrieval from memory because this information was never encoded. For such questions, responses corresponding to what really happened in the past can be arrived at by various inferential processes, including reconstruction based on the overall gist of the past event or borrowing from alternatives included in the question itself. But in the strictest terms, specific information necessary for answering these questions was never memorized and thus the most accurate response reflecting one's state of memory regarding the queried information is "don't know." We argue that asking questions about details a person did not see and which thus should be met with a "don't know" response could lead to important novel insights into how memory reports are constructed, providing a window into the interplay of memory retrieval and conversion processes.

The fundamental question addressed here is what determines "don't know" responses to unanswerable questions in memory tasks. Currently, our knowledge concerning "don't know" responding builds on procedures utilizing only answerable questions—that is, questions for which the correct answer was at some point presented to the rememberer—and it points to a straightforward conclusion: the more information remembered from the time of encoding, the greater the chance that a specific response will be volunteered in lieu of "don't know." This principle is encapsulated in what Koriat (1993, 1995) has dubbed an *accessibility heuristic*, where one's feeling that an answer to a memory question is

known—a product of metacognitive monitoring related to the rate of "don't know" responses (see Hanczakowski, Pasek, Zawadzka, & Mazzoni, 2013)—is determined by the volume of information accessed at retrieval. This relation between the volume of accessed information and feeling that one knows the correct answer is independent of whether this information is correct or not (Koriat, 1993). It is also independent of whether the accessed information strictly serves to answer a particular question. For example, familiarity with the question determines whether one feels the answer is known, even though such familiarity does not necessarily serve to point to the correct answer (Schwartz & Metcalfe, 1992). Thus, stated simply, the more details related to the particular queried information can be retrieved, the higher the feeling that the answer to a given question is known and thus the higher the likelihood that a specific candidate response will be provided instead of a "don't know" response.

What, then, about unanswerable questions? To the best of our knowledge, there are currently no studies examining whether the accessibility heuristic can be applied to account for "don't know" responding to questions to which an answer cannot be known. One reason for this could be that, so far, research employing the concept of the accessibility heuristic has focused on the issue of retrieval of partial target information, for example letters of a target item (Koriat, 1993), the emotional valence of the target word (Thomas, Bulevich, & Dubois, 2011), the strategy used to encode targets (Hertzog, Fulton, Sinclair, & Dunlosky, 2014), or some other semantic or episodic aspect of the target word (Koriat, Levy-Sadot, Edry, & de Marcas, 2003; Schwartz, Pillot, & Bacon, 2014; Thomas, Bulevich, & Dubois, 2012). Clearly, all these dimensions for which partial retrieval can occur don't apply to the specific case of unanswerable questions for which, by definition, no partial retrieval of target information is possible. However, just because access to partial target information is precluded it does not mean that there is no access to memory when unanswerable questions are asked. When describing the accessibility heuristic, Koriat (1993) wrote that metacognitive monitoring may depend on access to "structural, contextual, and semantic attributes; fragments of the target; and so on" (p. 611). Here we want to focus on one of the elements highlighted by Koriat: contextual attributes. We thus ask a specific question: How does retrieval of contextual attributes affect responding to unanswerable questions?

Context Effects on Memory and Reporting

Context effects in memory have been extensively examined for decades because context reinstatement seems to be one of the very few ways of improving memory at the time of retrieval (see Smith & Vela, 2001, for a comprehensive review). However, research on context effects on memory reporting can be actually divided into two separate streams. One line of research is focused on the issue of how access to target information can be promoted by reinstating encoding context at retrieval. By and large, studies conducted in this tradition have shown that reinstating encoding context augments retrieval of target information that was encoded in this context, although exceptions do exist. Generally, it seems that such context effects are more robust in the case of recall tests, where context can be used as an additional cue for accessing target information that was associated with the particular context at the time of encoding (Godden & Baddeley, 1975; Smith & Manzano, 2010). Much less robust patterns have been found in recognition, where benefits of context reinstatement for target recognition are found but often under somewhat constrained conditions: when perceptually rich contexts are used (Murnane, Phelps, & Malmberg, 1999), when specific encoding conditions asking for associating context to to-be-remembered information are administered (Hockley, 2008), or when target information itself is unfamiliar (Russo, Ward, Geurts, & Scheres, 1999). The important point from the present perspective, however, remains that all these studies which investigated whether context reinstatement augments memory across tests, encoding conditions, types of stimuli, and so forth, were focused inevitably on cases of answerable questions. These assessed whether context reinstatement leads to enhanced retrieval of *target information*—information solicited by particular memory queries—so unsurprisingly they did present target information to participants before the retrieval phase.

Another line of research has focused on the effect of retrieval context on metacognitive processes (Hanczakowski, Zawadzka, Collie, & Macken, 2017; Hanczakowski, Zawadzka, & Coote, 2014; Hanczakowski, Zawadzka, & Macken, 2015; Zawadzka, Simkiss, & Hanczakowski, 2018). For example, Hanczakowski et al. (2014) looked at how context reinstatement affects "don't know" responding in a recognition test and found that participants were more willing to provide specific responses to memory questions for which encoding context was reinstated. Interestingly, at least in one experiment this effect emerged even though there was no reliable effect of context reinstatement on target recognition. These results can be interpreted in terms of the Koriat's (1993) accessibility heuristic, where access to contextual memories at the time of retrieval, whenever these contexts were reinstated at test, affected metacognitive monitoring, making participants more confident that their candidate responses were correct and thus reducing the likelihood of providing a "don't know" response. To date, such an influence of context on conversion processes has also been examined exclusively with answerable questions. Importantly, however, when the metacognitive perspective is adopted, this approach seems limited and there is no apparent reason for the omission of unanswerable questions from research on context effects in remembering.

How can one reinstate context for something that was not seen, and what consequences such reinstatement may have? A simple example can be useful here. One can strive to remember what a coworker was wearing last Thursday and for this purpose one can try to augment memory by mentally recreating the office setting from that day. This operation of context reinstatement is actually independent of whether the coworker was at work on Thursday. If she was, then the question is answerable and context reinstatement may help retrieve the queried information, as numerous examples of context reinstatement benefits in memory demonstrate. But if she was absent that day, then the question becomes unanswerable—one cannot possibly know what the coworker wore if one had not seen her. One can only formulate less or more plausible guesses, for example based on how she typically dresses. What, then, would be the effects of reinstating context on the likelihood of submitting these guesses as responses to memory questions? We argue that two predictions should be considered.

Reinstating context at the time of retrieval leads to accessing more contextual information in memory that may not be spontaneously elicited by the question itself. The memory perspective

would underscore that such additional contextual information can be used as a cue for target information associated with context at encoding. But if we are dealing with an unanswerable question, then such cuing clearly cannot result in target retrieval. However, this ineffective cuing may have consequences for responding to a memory question—additional failed retrieval attempts may convince a person that the sought-after information is not available in memory at all. In other words, context reinstatement, by providing information relevant to the question, may help identify the question as unanswerable and consequently *increase* the chance that this question would be met with an appropriate "don't know" response. This would constitute a positive effect of reinstating context for unanswerable questions.

Although the scenario outlined above—more information relevant to the question translates into more appropriate responding—may seem intuitive, the accessibility heuristic allows for formulating an alternative prediction by which the effects of context reinstatement may be much less benign. One could stipulate that access to more complete memories of the queried event when context is reinstated—even if it does not result in more effective retrieval of target information (which remains unavailable)—can make the person more confident that the queried information should be known. This augmented feeling of knowing can in turn *reduce* the likelihood of providing an appropriate "don't know" response. The logic here is that memory for contexts and for targets embedded in these contexts is usually highly correlated in everyday experience, and so people learn to rely on this correlation in their memory decisions (see Hanczakowski et al., 2017, for a discussion). Thus, when a person is able to formulate a guess that would serve as a plausible candidate response to a memory question, confidence in the correctness of this guess is assessed based on the volume of memory information about the queried event that can be accessed. If a person remembers a lot about the context of the event, the candidate response under consideration seems more likely to be correct and is thus also more likely to be volunteered as a response. If a person does not remember much about the context of the queried event, the likelihood of responding "don't know" increases. The logic of the accessibility heuristic thus leads to a counterintuitive prediction by which more complete contextual memories about queried events may lead to less rather than more appropriate responding to unanswerable questions, at least when plausible candidate responses are available. This counterintuitive prediction will be assessed in the present study.

The Present Study

In this study, we set out to document a case in which context reinstatement leads to errors of commission in a memory report when unanswerable questions are asked. We conducted three experiments using a paradigm modeled on experimental work on eyewitness testimony, with the study phase consisting of a presentation of slides and text linked by a common story, and a memory test consisting of a series of specific two-alternative questions concerning details from these slides with a "don't know" option. The questions were either answerable or unanswerable based on the information presented in the slides. To manipulate context, the questions were preceded by slides—with the critical detail obscured—that either corresponded to the given question (the *reinstated* context condition), corresponded to a different question (the *repaired* context condition), or were novel and did not correspond to any of the questions (the *novel* context condition). The main focus was on the rates of "don't know" responding, particularly for unanswerable questions.

Two aspects of the current methodology warrant additional discussion. First, in the present investigation, we used recognition questions which provide participants with readily available plausible candidate responses in the form of recognition alternatives (see also Higham & Arnold, 2007; Koriat & Goldsmith, 1996b; Luna & Martín-Luengo, 2012). This methodological choice serves to largely circumvent the issues of how candidate responses for particular questions are generated. It is important to note that here we pursue the issue of how context reinstatement affects translation of candidate responses into overt reports and thus the question of how one comes up with particular candidate responses is peripheral to the issues under investigation. However, we need to acknowledge at the same time that by using recognition tests we actually limit the beneficial role context could play in memory reporting because—as already discussed—augmented retrieval due to context reinstatement is much more readily found in recall rather than recognition tests. Thus, one should bear in mind that our study serves to demonstrate the impact of context on conversion when plausible candidate responses are available, and that the dynamics of context effects may vary across different testing conditions.

Second, when considering the effects of context reinstatement it is important to focus on a critical methodological aspect: the baseline condition. In the literature, the effects of context associated with a given critical detail and later reinstated are compared against the baseline of either an entirely novel context (e.g., Franco-Watkins & Dougherty, 2006; Isarida, Isarida, & Sakai, 2012; Murnane & Phelps, 1993; Smith, Handy, Angello, & Manzano, 2014; Wong & Read, 2011) or against the baseline of a context that accompanied another detail/item at encoding (e.g., Burgess, Hockley, & Hourihan, 2017; Hockley, 2008; Koen, Aly, Wang, & Yonelinas, 2013; Macken, 2002). The first comparison has the benefit of greater ecological validity because in many everyday situations our memory for the past event is tested in contexts that have nothing to do with the original context of the queried event. The second comparison, however, has the benefits of greater experimental stringency because it allows for keeping the memory for context itself constant and thus directs the focus toward the issue of how item-to-context associations shape memory reporting. In the present investigation, we consider both comparisons. By having three levels of the context manipulation—reinstated, repaired, and novel contexts—we can assess whether reinstating context in which critical details were embedded—whether participants saw them or not—influences memory reporting compared with being tested in a familiar or in an entirely novel context.

Experiment 1

The present experiment looked at the role of context reinstatement in shaping "don't know" responding to both answerable and unanswerable questions. Apart from the context manipulations at retrieval described above, in Experiment 1 we introduced one more manipulation just before the final memory test was administered. Specifically, we delivered a specific warning about the presence of

unanswerable questions in the memory test for half of the participants, coupled with a clarification that any unanswerable question should be met with a "don't know" response. In everyday life, "don't know" response may have multiple meanings, ranging from *That information was not presented* to *I simply cannot remember an answer to this question* (Scoboria & Fisico, 2013; Scoboria et al., 2008). By including a warned group in our design, we explicitly defined that a "don't know" response may express confidence that the given target information was not studied, rendering it an unambiguously correct option for unanswerable questions. The inclusion of warning also helps us address the issue of whether participants respond differently when they expect unanswerable questions to be asked.

Method

Participants. Seventy-six participants (19 male; age range: 18–59) were students of various universities based in Kraków, Poland, who received monetary compensation for their participation. In all experiments reported here, sample sizes were determined by the number of participants willing to sign up for the study while it was being advertised, with the only constraints being imposed by the need to have equal numbers of participants per each experimental and counterbalancing condition. Participants were assigned in equal numbers to the warning-present and warning-absent groups. All participants were treated in accordance with the British Psychological Society Code of Human Research Ethics and the Polish National Science Centre guidelines for research ethics. Participation was voluntary, and participants were informed explicitly that they could withdraw their consent at any point without any adverse consequences, and that their data would be anonymized. After the study, all participants were debriefed.

Design. Three manipulations were used in the present study. First, warning/no warning concerning the presence of unanswerable questions served as a between-participants factor. Second, retrieval context was manipulated within participants, with each participant being subjected to reinstated (same context at study and at test), repaired (studied context presented with a different question at test) and novel (new context used at test) context conditions. Finally, question type was also manipulated within participants, with each participant being presented with an equal number of answerable (concerning details that were presented) and unanswerable questions (concerning details that were not presented). The assignment of questions to answerable and unanswerable conditions, and the assignment of context conditions to those questions, was counterbalanced between participants.

Materials. For study, two stories describing crimes (forging university diplomas and being a victim of mugging) were created. Each story consisted of 18 photographs, which can be found at https://osf.io/37cdy/, accompanied by brief written descriptions. For counterbalancing purposes, there were two versions of each photograph. On one version of each photo, critical information needed to answer a subsequent memory question was present, whereas this information was absent from the other version. The absent information related to items that were either obscured or entirely removed in such a way as to still enable direct questions about these items that would be unanswerable based on information provided in the photographs (see Figure 1 for an example). For the test phase, a third version of each of the 18 photographs was prepared, with the critical detail obscured by an opaque blue rectangle. These photographs served as context in the reinstated and repaired context conditions. An additional set of 12 photographs, not presented at study and unrelated to the studied stories, served as the source of novel contexts at test. A blue rectangle was also superimposed on each of these photographs. The test consisted of 36 multiple-choice questions (18 for each story) about the critical details. Three response options were provided for each question: two alternatives and a "don't know" option. One of the alternatives always correctly described the detail presented in the slide version prepared for the answerable condition, while the other alternative was always incorrect. The order in which the two alternatives were presented was randomized anew for each participant and for each question. "Don't know" was always presented as the third response option. Figure 1 presents an example of two photographs—one with critical information present, the other with critical information absent—accompanied by a written description, and the question regarding the critical detail (translated into English) with three different context photographs. In designing materials and questions, we took care to choose critical information in such a way as to make it impossible for participants to deduce correct answers based on the reinstated context alone (e.g., seeing a context photograph of a room in which the queried picture was displayed cannot help one deduce what was presented in the picture).

Procedure. At study, participants were presented with the two crime stories one after another. Each story consisted of 18 brief statements (one or two sentences long) followed by a photograph, and presented one at a time. The statements were the same for all participants. Within each story, participants were presented with nine photographs containing critical information (on which answerable questions were based), and nine omitting this information (on which unanswerable questions were based). Presentation times were determined on the basis of a pilot experiment. As a result, each statement was presented for 4 s, and each photograph for 2 s. There was a 2-s interval between the presentation of statements and photographs. The order in which the stories were presented (forgery first or mugging first) was counterbalanced across participants.

Immediately after study, participants were given two tests, one for each of the stories. The tests were presented in the same order as the stories. The warning group received additional instructions informing them that some questions in the tests may be unanswerable, that no relevant information had originally been presented in the slides for these questions, and that they should respond "don't know" to questions they believe to be unanswerable. The no-warning group received standard test instructions that made no mention of unanswerable questions and no clarifications regarding the "don't know" option; however, the "don't know" option was clearly presented to them as a response option during the test. Each test consisted of 18 questions presented in a random order. On each test trial, a context photograph was first presented for 5 s, with a blue rectangle obscuring part of it (see Figure 1). Of the 18 context photographs presented at test, six were based on the photograph which originally contained the critical detail (reinstated condition), six were based on a photograph taken from the other story (repaired condition), and six were novel photographs, not presented before (novel condition). After the photograph was removed from the screen, a question was presented. Each question contained

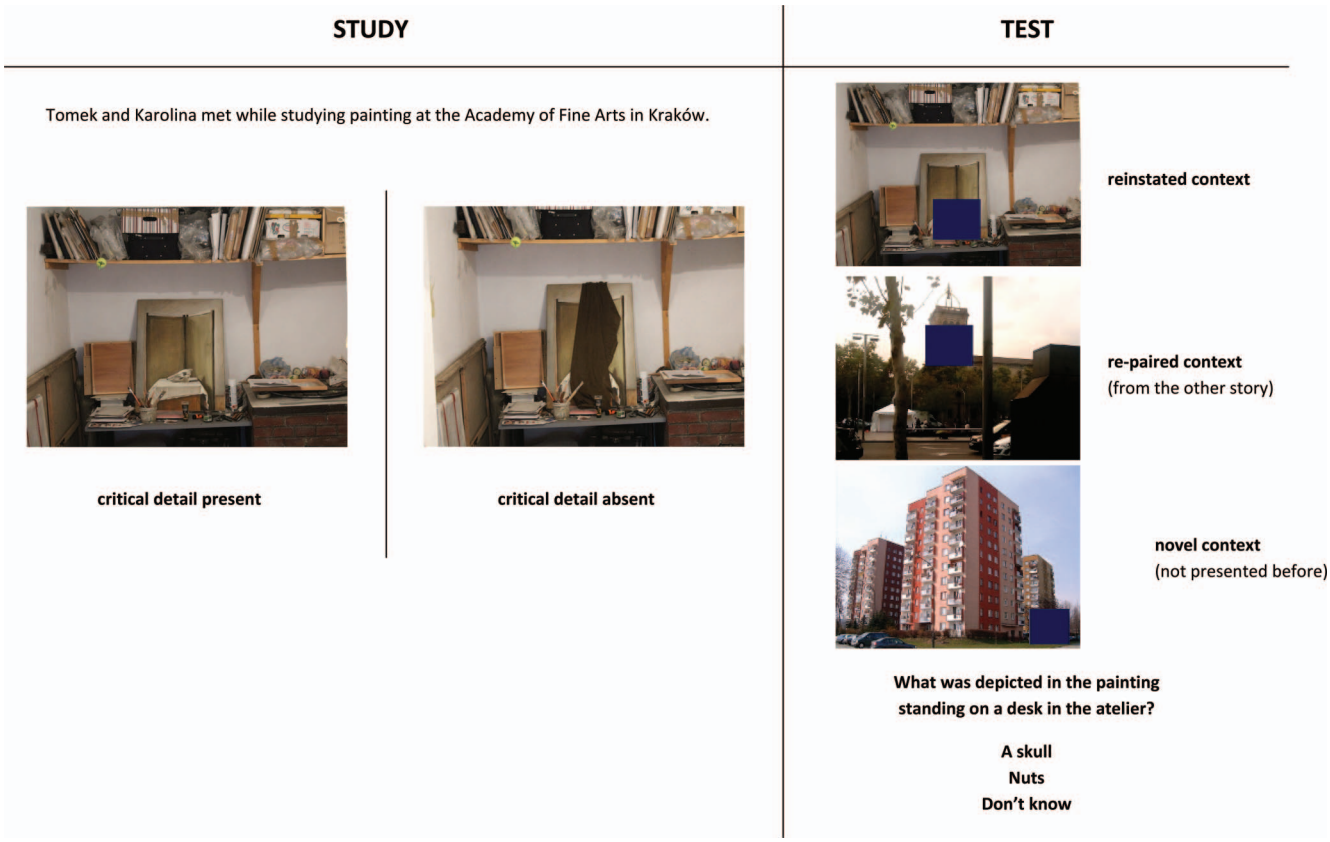

Figure 1. Examples of a study and test trial (translated into English) in Experiment 1. At study, half of the participants were presented with the critical detail (left panel), whereas for the other half it was obscured (middle panel). At test, participants were presented with a question accompanied by one of three context photographs. In the reinstated context condition, the same context picture was used as in the study phase (top right). In the repaired context condition, a context picture was taken from the other story (middle right). In the novel context condition, a new photograph was used (bottom right). Independent of condition, an opaque rectangle obscured either the critical detail or the place where this detail could have been plausibly presented. See the online article for the color version of this figure.

enough details taken from both the relevant studied picture and the accompanying sentence(s) to make sure that participants knew which specific part of the story it was referring to (e.g., if the question was about a house or building number, it clearly explained whether it was referring to the protagonists' dream house or the address of the local police station). Participants' task was to either choose one of two alternatives or respond "don't know," and an answer was required to progress to the next trial. Test questions and alternatives were the same for all participants.

Results and Discussion

The mean rates of "don't know" responses are presented in Table 1, and the mean accuracy scores for recognition decisions for answerable questions—the number of hits divided by the number of questions for which a specific (rather than "don't know") response was given[1]—are presented in Table 2. The analyses focused on "don't know" responses to answerable and unanswerable questions. The accuracy scores for answerable questions were not analyzed because in a free-report procedure with the "don't know" option these rates cannot be clearly interpreted. This happens because accuracy in free-report tests reflects how a certain manipulation affects both retrieval of target information and the process of conversion (Koriat & Goldsmith, 1996b).

The rates of "don't know" responses were analyzed as a function of type of question (answerable vs. unanswerable), warning (present vs. absent), and context (reinstated, repaired, novel), using a three-factor mixed analysis of variance (ANOVA). Significant effects of context were followed up with specific comparisons, comparing the reinstated context condition against the baseline of novel context and, in a more stringent comparison, against the baseline of repaired context. It is vital to note that the rates of "don't know" responses are equivalent to the reversed rates of volunteering responses in a memory report. This is particularly

[1] This measure is often referred to as output-bound accuracy because it looks at the accuracy of responses that were volunteered. An alternative measure is input-bound accuracy, which looks at the proportion of hits for all questions that were asked.

Table 1
Mean Rates of "Don't Know" Responses in Experiment 1 (as a Function of Warning Group, Context at Retrieval, and Type of Question) and Experiments 2 and 3 (as a Function of Context at Retrieval and Type of Question Group)

Experiment	Warning			No warning		
	Reinstated context	Repaired context	Novel context	Reinstated context	Repaired context	Novel context
Experiment 1						
Answerable questions	.28 (.03)	.40 (.04)	.42 (.03)	.30 (.04)	.25 (.04)	.32 (.04)
Unanswerable questions	.61 (.04)	.69 (.04)	.71 (.04)	.42 (.05)	.43 (.05)	.51 (.05)
Experiment 2						
Answerable questions	—	—	—	.23 (.02)	.26 (.02)	—
Unanswerable questions	—	—	—	.46 (.03)	.51 (.03)	—
Experiment 3						
Answerable questions	—	—	—	.25 (.03)	.28 (.03)	.31 (.03)
Unanswerable questions	—	—	—	.51 (.04)	.56 (.04)	.58 (.04)

Note. Standard errors of means are given in parentheses.

important for unanswerable questions for which any drop in the rate of "don't know" responses translates to an increased propensity to volunteer responses that cannot match any record in participants' memory.

The 2 (question type) × 2 (group) × 3 (context condition) mixed ANOVA on the mean rates of "don't know" responses yielded a significant main effect of question type, $F(1, 74) = 144.51$, $MSE = .04$, $p < .001$, $\eta_p^2 = .661$, indicating a greater inclination to respond "don't know" to unanswerable questions ($M = .56$, $SD = .27$) than to answerable questions ($M = .33$, $SD = .18$). A significant main effect of warning, $F(1, 74) = 10.92$, $MSE = .23$, $p = .001$, $\eta_p^2 = .129$, indicated a greater inclination to respond "don't know" in the group that received a warning about unanswerable questions ($M = .52$, $SD = .17$) than the group that received no warning ($M = .37$, $SD = .37$). There was also a significant main effect of context, $F(2, 148) = 10.48$, $MSE = .03$, $p < .001$, $\eta_p^2 = .124$. These effects were qualified by a significant interaction of context and warning, $F(2, 148) = 4.75$, $MSE = .03$, $p = .010$, $\eta_p^2 = .060$, and a significant interaction of question type and warning, $F(1, 74) = 14.37$, $MSE = .04$, $p < .001$, $\eta_p^2 = .163$. The three-way interaction was not significant, $F(2, 148) = 1.25$, $p = .288$, $\eta_p^2 = .017$.

To disentangle the significant interactions, we first examine the effects of retrieval context, which constitute the main focus of the present study. Given that context did not interact with question type, we collapsed the data across the latter variable for the overall analyses, although we present the crucial findings for unanswerable questions at the end of the relevant sections given that they constitute the main novel contribution of this study. To follow-up on the significant context by warning interaction, two 2 × 2 mixed ANOVAs on the rates of "don't know" responses were conducted, one focusing on the comparison of the reinstated context condition against the baseline of novel context condition and one against the baseline of repaired context condition. The first 2 (warning: present vs. absent) × 2 (context condition: reinstated vs. novel) mixed ANOVA yielded a significant main effect of warning, $F(1, 74) = 6.35$, $MSE = .08$, $p = .014$, $\eta_p^2 = .079$, showing more "don't know" responses when warning was present ($M = .51$, $SD = .18$) rather than absent ($M = .39$, $SD = .23$). More interestingly, the main effect of context was also significant, $F(1, 74) = 22.31$, $MSE = .01$, $p < .001$, $\eta_p^2 = .232$, but the interaction was not, $F(1, 74) = 2.99$, $p = .088$, $\eta_p^2 = .039$. The significant effect of context arose because participants were less inclined to provide "don't know" responses in the presence of reinstated contexts ($M = .40$, $SD = .23$) than in the presence of novel contexts ($M = .49$, $SD = .22$). Importantly, when the same comparison was made for unanswerable questions only, the reduction in the rate of "don't know" responding in the reinstated context condition relative to the novel context condition remained reliable, $t(75) = 3.52$, $p < .001$, $d = 0.40$.

The second 2 (warning: present vs. absent) × 2 (context condition: reinstated vs. repaired) mixed ANOVA yielded a significant main effect of warning, $F(1, 74) = 9.61$, $MSE = .08$, $p = .003$, $\eta_p^2 = .115$, again showing more "don't know" responses when warning was provided ($M = .50$, $SD = .19$) rather than absent ($M = .35$, $SD = .22$). More importantly, the main effect of context was also significant, $F(1, 74) = 4.32$, $MSE = .01$, $p = .041$, $\eta_p^2 = .055$, and so was the context by warning interaction, $F(1, 74) = 9.52$, $MSE = .01$, $p = .003$, $\eta_p^2 = .114$. Follow-up comparisons focused on context reinstatement separately for the warning and no-warning groups. These revealed that the context reinstatement effect on "don't know" responses was significant when warning was provided, $t(37) = 3.40$, $p = .001$, $d = 0.57$, with fewer "don't know" responses in the reinstated context condition ($M = .45$, $SD = .21$) than in the repaired context condition ($M = .55$, $SD = .20$). By contrast, the context reinstatement effect on "don't know" responses was absent when there was no warning, $t < 1$, $d = 0.12$, with a comparable rate of "don't know" responses

Table 2
Mean Hit Rates to Answerable Questions (After Exclusion of "Don't Know" Responses) in Experiment 1 (as a Function of Warning Group and Context) and Experiments 2 and 3 (as a Function of Context)

Experiment	Reinstated	Repaired	Novel
Experiment 1			
Warning	.64 (.04)	.51 (.04)	.50 (.04)
No warning	.59 (.04)	.60 (.03)	.53 (.04)
Experiment 2	.84 (.02)	.86 (.01)	—
Experiment 3	.64 (.02)	.59 (.02)	.59 (.02)

Note. Standard errors of means are given in parentheses.

in the reinstated ($M = .36$, $SD = .25$) and repaired ($M = .34$, $SD = .23$) context conditions. Thus, for the more stringent comparison of the effect of reinstating context against the baseline of equally familiar contexts that nevertheless did not match the questions, the reduction in the rate of "don't know" responses occurred for participants warned about the presence of unanswerable questions in the test. To confirm that this effect emerged for unanswerable questions specifically, we conducted a separate comparison for the warned group, confirming that participants in this group were indeed less likely to respond "don't know" to unanswerable questions in the presence of reinstated rather than merely familiar contexts, $t(37) = 2.29$, $p = .025$, $d = 0.38$.[2]

Moving to the issue of how the warning manipulation affected "don't know" responding independently of variations in retrieval context (a significant warning by type of question interaction in the overall ANOVA), separate comparisons of the effect of warning for answerable and unanswerable questions, collapsing across context conditions, were conducted. These revealed that warning reliably affected "don't know" responding to unanswerable questions, $t(74) = 3.96$, $p < .001$, $d = 0.92$, with more "don't know" responses when warning was provided ($M = .67$, $SD = .21$) rather than absent ($M = .45$, $SD = .27$). The same comparison for answerable questions was not significant, $t(74) = 1.84$, $p = .07$, $d = 0.46$, although a numerical trend still suggested slightly more "don't know" responses when the warning was provided ($M = .37$, $SD = .16$) rather than when it was absent ($M = .29$, $SD = .19$). Altogether, these results indicate that warning participants about unanswerable questions reliably increases the rate of correct "don't know" responses to these questions, while having a smaller (and not statistically significant in the present investigation) effect on "don't know" responses to answerable questions.

To summarize, the rate of "don't know" responses was affected by both the warning and the context manipulations. With regard to warning, it generally raised the rate of "don't know" responses, but did so to a markedly greater extent for unanswerable questions. If warning served only to make participants overall more willing to use the "don't know" option, we would expect equal changes in the "don't know" rates across answerable and unanswerable questions. A more robust effect observed for unanswerable questions suggests that warned participants were able to positively identify at least some of the questions as unanswerable. We return to this issue in the General Discussion.

With regard to context, context reinstatement generally lowered the rate of "don't know" responses, both compared with being tested in the presence of a novel context and in the presence of a repaired context, although in the latter case the effect unexpectedly emerged only for participants who were specifically warned about the presence of unanswerable questions in the memory test. Crucially, the same pattern of context effects was observed for answerable and unanswerable questions alike. Thus, reinstating the context of the original event served to increase the rate of commission errors for unanswerable questions, consistent with our prediction that relying on the accessibility heuristic can distort memory reports when unanswerable questions are considered.

We note, however, that the finding of an interaction of context reinstatement and warning when a more stringent definition of context reinstatement was adopted (a comparison against the baseline of a repaired context) somewhat undermines the consistency of our results. On the one hand, it is reassuring that the context reinstatement effect for unanswerable questions emerged when participants were explicitly informed that they should use the "don't know" option for unanswerable questions. This confirms that our results are not merely due to some misunderstanding as to what the phrase "don't know" means. On the other hand, this interactive pattern is surprising given that if this particular effect emerges even when there is absolutely no ambiguity as to how the "don't know" option should be utilized, then we see no reason for it to be absent when the "don't know" option is not clearly specified.

One way to explain this unexpected interaction is to argue that warned participants focused more on contextual cues at test, utilizing all available information to identify which questions were answerable and which were not. Obviously, if context is barely attended to, it cannot serve as an effective cue for contextual associations, and with no additional access to memory information no effect on conversion and reporting would be expected. Given this ambiguity concerning the results of Experiment 1, an additional experiment was conducted to ensure that the impact of context reinstatement on responding to unanswerable questions is robust. To circumvent the problem of insufficient attention devoted to contextual information, in Experiment 2 we modified the testing instructions to emphasize that contexts provided by the experimenter could be helpful in responding to memory queries.

Experiment 2

The main aim of the present experiment was to examine once again how context reinstatement affects "don't know" responding to answerable and unanswerable questions. No effect of context reinstatement when compared against the baseline of repaired contexts was obtained in Experiment 1 when participants were not warned about the presence of unanswerable questions in the test. Here we replicated this no-warning condition, but to induce more attention to contextual information, we instructed participants that contextual photographs they were presented with at test could help them retrieve relevant information and thus they should pay careful attention to these photographs. Because the focus of the present experiment was on the more stringent definition of context reinstatement—as the comparison of reinstated and novel context conditions provided robust findings in Experiment 1—the novel context condition was not included.

Apart from clarifying the surprising pattern concerning context reinstatement from Experiment 1, this experiment provided an opportunity to examine one additional issue. The assumption of Koriat and Goldsmith's (1996b) framework is that "don't know" responding is closely related to confidence in candidate responses. This is also an assumption adopted for the present investigation, as

[2] An anonymous reviewer wondered whether context effects on "don't know" responding change in the course of a memory test. To assess this possibility, we reran the main 2 (question) × 2 (group) × 3 (context) mixed ANOVA by including an additional factor of test—whether the examined questions were asked in the first or the second test (with different tests relating to different stories participants read). For Experiment 1, this ANOVA yielded no reliable interactions including factors of context and test, smallest $p = .19$ for the interaction of test and context. Thus, the present data provide no evidence that context effects on "don't know" responding change in the course of the memory task. We performed the same analyses for Experiments 2 and 3 with the same outcomes, which we don't report here for the sake of brevity.

we argue that access to additional memory information triggered by contexts present at retrieval increases one's confidence in candidate responses, translating into reduced rates of "don't know" responses. To provide support for this assumption and thus to confirm the locus of the context effects at the stage of metacognitive monitoring of candidate responses, in Experiment 2 confidence judgments for forced-report recognition decisions were collected. We expected that these would also be affected by the context reinstatement manipulation.

Method

Participants. Ninety-six participants (24 male; age range: 19–55) were students of various universities based in Kraków, Poland, who received monetary compensation for their participation. As with Experiment 1, we tested all participants who responded to our advertisement. We chiefly aimed here to replicate the effect of context reinstatement on "don't know" responding when contexts are attended to, which in Experiment 1 corresponded to $d = 0.57$ when collapsed across answerable and unanswerable questions and $d = 0.38$ when only unanswerable questions were considered. To obtain power of .90 and with two-tailed tests, 35 participants would be needed for the effect collapsed across questions and 75 participants for the effect for unanswerable questions only.

Materials and design. The study and test materials were the same as in Experiment 1, with the exception of the 12 photographs used for the novel condition which were not utilized in this study. The same test materials were used in free- and forced-report tests. In this experiment, two within-participant manipulations were used. This time, the context manipulation had only two levels (reinstated, repaired). This translated into a larger number of test questions assigned to each of the context conditions (18 across both tests, as compared with 12 in Experiment 1). The question type manipulation was the same as in Experiment 1, with half of the questions being answerable and the other half unanswerable. The assignment of questions to the context and question type conditions was the same on free- and forced-report tests.

Procedure. The study and free-report test phases of the present experiment were modeled on those from the no-warning group in Experiment 1, with one exception. The instructions displayed before the start of the first free-report test underscored that each question would be preceded by a photograph, and that participants should "pay close attention to this photograph, as for some questions it might help recall the right answer to the following memory question." After completing both free-report tests, participants were asked to answer the same questions again, but this time without an option to say "don't know." They were told to guess if they did not know the answer.[3] Again, instructions regarding the context photographs were provided, underscoring the need to attend to these photographs. At test, for each question participants had to choose one of two response options. After confirming their choice, they were presented with a prompt for a confidence rating, and a scale from 0 (*Guess—I don't know whether my answer is correct*) to 4 (*I am absolutely certain that my answer is correct*). They could advance to the next question only after a rating had been provided.

Results and Discussion

The mean rates of "don't know" responses are presented in Table 1, whereas the mean accuracy scores for recognition decisions for answerable questions are presented in Table 2.[4] The analyses focused again primarily on the rates of "don't know" responses, but this time the design of the experiment also allowed for examining forced-report recognition performance and retrospective confidence judgments.

A 2 (question type: answerable vs. unanswerable) × 2 (context condition: reinstated vs. repaired) within-participant ANOVA on the rate of "don't know" responses yielded a significant main effect of question, $F(1, 95) = 127.10$, $MSE = .04$, $p < .001$, $\eta_p^2 = .572$, which arose because participants gave more "don't know" responses to unanswerable ($M = .48$, $SD = .29$) than to answerable questions ($M = .25$, $SD = .17$). A main effect of context condition was also significant, $F(1, 95) = 7.49$, $MSE = .02$, $p = .007$, $\eta_p^2 = .073$, as participants gave fewer "don't know" responses in the reinstated context condition ($M = .35$, $SD = .23$) than in the repaired context condition ($M = .38$, $SD = .22$). The interaction was not significant, $F < 1$. To ensure that the pattern of reduced "don't know" responding in the presence of reinstated contexts was reliable specifically for unanswerable questions, we conducted an additional comparison, which indeed showed a significant difference between reinstated and repaired context conditions, $t(95) = 2.46$, $p = .016$, $d = 0.26$.

An analysis of forced-report recognition performance was conducted for answerable questions because only for these questions one of the alternatives presented in the forced-report recognition test was correct. A planned comparison showed no reliable difference between the reinstated context ($M = .77$, $SD = .16$) and repaired context ($M = .74$, $SD = .14$) conditions, $t(95) = 1.14$, $p = .255$, $d = 0.16$. Thus, there was only a numerical difference favoring recognition performance in the reinstated context condition, indicating that reinstating context had at most a very small effect on retrieval of target information when such information could have been stored and then retrieved in response to answerable questions.

Mean confidence judgments for forced-report recognition decisions were analyzed with a 2 (question type) × 2 (context condition) within-participant ANOVA. This yielded a significant main effect of question type, $F(1, 95) = 416.39$, $MSE = 0.34$, $p < .001$, $\eta_p^2 = .814$, which arose because participants were more confident in their responses to answerable ($M = 2.38$, $SD = 0.59$) than unanswerable questions ($M = 1.17$, $SD = 0.75$). The main effect

[3] In the present experiment, we used a free-then forced-report order of tests, just as in our previous work on "don't know" responding (Hanczakowski et al., 2013, 2014). The order of tests has often been assumed to play a negligible role in shaping the memory and metamemory patterns (Koriat & Goldsmith, 1996b), although some recent work suggests that free- then forced-report order may actually promote more efficient metacognitive control and even better memory retrieval (Hollins & Weber, 2017).

[4] As Table 2 shows, hit rates are markedly higher in this experiment than in Experiments 1 and 3. This is likely because, owing to a change of the experimental software, for Experiment 2 we had to reduce the size of context pictures from 800 × 600 to 640 × 480 pixels. This likely made it easier for participants to attend to the whole picture within the 2-second time window at study. Nevertheless, this had no impact on the rates on "don't know" responding, as shown in Table 1.

of context was also significant, $F(1, 95) = 9.70$, $MSE = 0.24$, $p = .002$, $\eta_p^2 = .093$, as participants were more confident in the correctness of their recognition responses accompanied by reinstated ($M = 1.86$, $SD = 0.65$) rather than repaired contexts ($M = 1.70$, $SD = 0.65$). The interaction was also significant, $F(1, 95) = 4.12$, $MSE = 0.15$, $p = .04$, $\eta_p^2 = .042$. Although the difference in confidence between context conditions was reliable for answerable questions, $M = 2.50$, $SD = 0.65$ for the reinstated condition and $M = 2.27$, $SD = 0.71$ for the repaired condition, $t(95) = 3.27$, $p = .001$, $d = 0.33$, it was not for unanswerable questions, $M = 1.21$, $SD = 0.80$ for the reinstated condition and $M = 1.14$, $SD = 0.79$ for the repaired condition, $t(95) = 1.36$, $p = .18$, $d = 0.13$, with only a numerical difference in the predicted direction. These results are largely similar to the results obtained for "don't know" responses, bar the significant interaction obtained for confidence but not volunteering rates. We don't put much weight to this interaction, as it is likely to reflect the fact that confidence judgments for unanswerable questions were essentially at floor and thus relatively less sensitive as a measure of cognitive states. We thus suggest that the main message derived from the present analyses is that context reinstatement generally affects retrospective confidence in forced-report recognition decisions.[5]

To summarize the results of the present experiment, context reinstatement—compared with being tested in merely familiar contexts—again has been shown to affect "don't know" responding to answerable and unanswerable questions alike, while at the same time making participants more confident in recognition responses they provided when responding "don't know" was precluded. These results indicate that the effects of context reinstatement on conversion are robust, suggesting that the unexpected interaction obtained in Experiment 1—a failure to find the effects of context reinstatement on "don't know" responses for the no-warning group—most likely stemmed from insufficient use of context information at retrieval, remedied here by specific experimental instructions.

Interestingly, the accuracy of responses provided in the forced-report recognition test for answerable questions was similar for the reinstated and repaired context conditions. This could be taken to suggest that context reinstatement had a negligible effect on retrieval of target information potentially stored at encoding. In other words, although context reinstatement clearly facilitated retrieval of a number of details related to memory queries—as evidenced by the "don't know" and confidence data—it might have been less effective in facilitating retrieval of the specific target information that would allow for correctly answering the particular questions included in the memory tests. We would like to sound a word of caution here, however. Although these results can be taken to indicate that any effect observed for responding to answerable questions is driven by changes in conversion processes, spurred by retrieval of contextual details rather than target information per se, one should be careful before drawing firm conclusions from null results in performance measures. After all, performance measures are never process-pure measures of cognitive processes and thus changes in the probability of target retrieval could be masked by other processes, for example if recognition tests can also be completed based on familiarity (see Hanczakowski et al., 2014, Experiment 1).

Returning to the main focus of the present study—the issue of unanswerable questions—the results presented so far demonstrate how an empirical pursuit of unanswerable questions can reveal novel patterns related to memory reporting, such as negative effects context reinstatement may have for the accuracy of information provided in memory reports. An outstanding issue concerns the conditions under which context information could be discounted as a cue shaping "don't know" responding for unanswerable questions. In Experiment 1, the warning group was informed about the presence of unanswerable questions in the memory tests and instructed to use "don't know" responses to explicitly express their conviction that a given question is unanswerable. This warning, however, did not alter the influence of context information on conversion processes, as participants still used additional retrieval of details other than the target information—precipitated by contextual cues—when deciding whether a "don't know" response should be given. In Experiment 3, we tested another condition under which people may be discouraged from using retrieval of nontarget information for the purpose of conversion. Specifically, we changed the manipulation of the type of question from a within- to a between-participants one, and thus included a group that was asked exclusively unanswerable questions. By facing a series of 36 questions for which no target information can be retrieved, participants could potentially discount memory information elicited by the context when deciding whether the correct answers to memory queries are known. In this case, we would predict a different pattern of results regarding "don't know" responses between the answerable questions group—for which we would expect to replicate the basic results of Experiments 1 and 2—and the unanswerable-questions group.

Experiment 3

The aim of the present experiment was to assess whether context information affects "don't know" responding to unanswerable questions when retrieval of target information for each and every question in the memory test is not possible. To this aim, we replicated the design of Experiment 1, but instead of manipulating the type of question variable within participants, we tested two groups of participants, one receiving only answerable and the other receiving only unanswerable questions (see Cox & Dobbins, 2011, for a similar idea). We expected to replicate the pattern of Experiment 1 in the answerable-questions group—fewer "don't know" responses when contexts are reinstated. If participants learn to

[5] The simultaneous use of a free-report and a forced-report test coupled with confidence ratings allows also for verifying whether "don't know" responses are indeed related to the products of metacognitive monitoring in the form of confidence in candidate responses, as the conceptual framework of Koriat and Goldsmith (1996b) stipulates. To assess this, we computed mean confidence in forced-report answers, separately for questions for which a specific or a "don't know" response was given in the free-report test. From these results it is clear that confidence was higher whenever a specific response was given rather than a "don't know" response, both for answerable questions, $M = 2.97$, $SD = 0.47$ versus $M = 0.75$, $SD = 0.71$, $t(84) = 26.09$, $p < .001$, $d = 2.83$, and for unanswerable questions, $M = 1.71$, $SD = 0.78$ versus $M = 0.65$, $SD = 0.61$, $t(84) = 12.74$, $p < .001$, $d = 1.38$ (degrees of freedom differ between the analyses reported here and those in the main text because 12 participants failed to provide at least one type of response across all trials). This confirms that for both answerable and unanswerable questions, decisions whether to volunteer a specific response to a memory query were related to confidence in candidate responses. We thank Morris Goldsmith for suggesting this analysis.

discount retrieval of related details that don't serve to answer specific memory queries as cues to shaping conversion processes, then a different pattern should emerge in the unanswerable-questions group, with no effects of context on "don't know" responding. If, however, participants are unable to learn in this specific testing condition that retrieval from memory does not help in answering questions posed in the memory test, then the same patterns should emerge as in the answerable-questions group. In this experiment, we again assessed the effects of context reinstatement against the baseline of novel contexts and against the more stringent baseline of contexts familiar from the study phase.

Method

Participants. One hundred two participants (21 male; age range: 18–31) were students of various universities based in Kraków, Poland, who received monetary compensation for their participation. They were randomly assigned in equal numbers to the answerable- and unanswerable-questions groups. Once again, we tested all participants who responded to our advertisement within the testing period.

Materials, design, and procedure. All materials were the same as in Experiment 1. The slides were arranged in such a way that one group of participants was provided only with slides that rendered subsequent questions answerable whereas the other group of participants was provided only with slides that rendered subsequent questions unanswerable. The novel context condition was again included, resulting in three levels of the context manipulation (reinstated, repaired, novel). Only free-report tests were conducted, and the testing procedure was the same as in Experiment 1 for the group that received no warning and no clarification as to the meaning of the "don't know" option.

Results and Discussion

The mean rates of "don't know" responses are presented in Table 1 whereas the mean accuracy scores (hit rates) for recognition decisions for answerable questions (after eliminating "don't know" responses) are presented in Table 2. The analyses focused again on "don't know" responses to answerable and unanswerable questions.

The 2 (group: answerable vs. unanswerable questions) × 3 (context: reinstated, repaired, novel) mixed ANOVA yielded a main effect of group, $F(1, 100) = 37.03$, $MSE = .16$, $p < .001$, $\eta_p^2 = .268$, which reflected higher rates of "don't know" responding for unanswerable ($M = .55$, $SD = .27$) than for answerable questions ($M = .27$, $SD = .18$). The main effect of context was also significant, $F(2, 200) = 7.96$, $MSE = .01$, $p < .001$, $\eta_p^2 = .073$, but the interaction was not, $F < 1$. Data were thus collapsed across the groups and the effect of context reinstatement was assessed against the baseline of novel and repaired contexts separately. For the reinstated-to-novel comparison, there were fewer "don't know" responses when context was reinstated, $M = .38$, $SD = .28$, rather than novel, $M = .44$, $SD = .28$, $t(100) = 3.83$, $p < .001$, $d = 0.38$. When the baseline of repaired contexts was adopted, the same effect emerged, with more fewer "don't know" responses for the reinstated rather than repaired context condition, $M = .42$, $SD = .29$, $t(100) = 2.76$, $p = .007$, $d = 0.27$. Both of these effects remained reliable when the analyses were restricted to unanswerable questions only, $t(50) = 2.46$, $p = .017$, $d = 0.34$, and $t(50) = 2.28$, $p = .027$, $d = 0.32$, respectively.

Overall, the results of the present experiment largely replicated the results of the previous experiments, despite the change to the experimental design. Once again, reinstating the encoding context at the time of retrieval reduced "don't know" responding compared with a situation in which contexts were either novel (as in Experiment 1) or repaired (as in Experiments 1 and 2), and this effect emerged for both answerable and unanswerable questions. The fact that the patterns of how context information affected "don't know" responses in the present experiment remained largely unchanged despite using a between-participants manipulation of question type testifies to the robustness of the effect context-induced retrieval has on conversion processes.

Combined Analysis

Given the novelty of our main finding that context reinstatement—a manipulation commonly thought to improve memory at retrieval—can actually disrupt conversion processes, resulting in a decrease in correct "don't know" responding when unanswerable questions are asked, we performed additional analyses on combined data. To quantify the strength of evidence in support of our conclusions, we conducted Bayesian repeated-measures ANOVAs on unanswerable questions data using the JASP software (JASP Team, 2019). The first Bayesian ANOVA included the factors of context (reinstated, novel) and group (Experiment 1 – warning, Experiment 1 – no warning, Experiment 3 – unanswerable questions).[6] As Table 3 shows, the model that garnered the most support from the data included two main effects but no interaction, and all other models fared substantially worse (the model with the second strongest degree of support, which included only the main effect of context, was 10 times less likely than the model with both main effects included). An analysis of effects across matched models revealed that there was extreme (Lee & Wagenmakers, 2013) evidence for including the effect of context in the model ($BF_{inclusion} = 545.96$) and moderate evidence for including the effect of group ($BF_{inclusion} = 9.62$), and at the same time there was strong support against the interaction ($BF_{inclusion} = 0.09$).

To quantify support for the more stringent of the two context reinstatement effects, we conducted another Bayesian repeated-measures ANOVA with the factors of context (reinstated, repaired) and group (Experiment 1 – warning, Experiment 1 – no warning, Experiment 2, Experiment 3 – unanswerable questions). As Table 4 shows, the results closely mirrored those from the reinstated-to-novel analysis. Again, the model with two main effects but no interaction received the strongest support from the data, and the second best model including only the main effect of context was the second most supported, with evidence weaker by a factor of 10. An analysis of effects across matched models revealed that there was very strong evidence for including the effect of context in the model ($BF_{inclusion} = 38.94$) and strong evidence for including the

[6] Note that Experiment 2 does not feature in this analysis, because there was no novel context condition in this experiment.

Table 3
Bayesian Model Comparisons for Combined Analysis of Unanswerable Question Data Including Context (Reinstated, Novel) and Group (Experiment 1 – Warning, Experiment 1 – No Warning, Experiment 3 – Unanswerable Questions) as Factors

Model	p(M)	p(M|Data)	BF_M	BF_{10}	Error%
Context + Group	.20	.843	21.450	1.000	
Context	.20	.078	0.337	0.092	4.486
Context + Group + Context × Group	.20	.056	0.236	0.066	6.075
Group	.20	.022	0.088	0.026	5.411
Null	.20	.002	0.008	0.003	4.324

Note. p(M) shows uniform prior probabilities for all competing models. p(M|Data) shows posterior model probabilities. BF_M shows the change from prior to posterior odds. BF_{10} shows how much support compared to the best model (top row) each additional model provides, with values above 1 showing increasing evidence for the model, and below 1 showing increasing evidence against the model.

effect of group ($BF_{inclusion}$ = 10.82), as well as strong support against the interaction ($BF_{inclusion}$ = 0.07).[7]

Together, the results from the combined analyses confirm the main novel conclusion from this study. When unanswerable questions are asked, context reinstatement results in an increase in commission errors stemming from a decreased willingness to correctly answer "don't know." This conclusion holds true regardless of whether the baseline condition against which the effects of context reinstatement are measured holds context familiarity constant, as in some recognition memory studies on context reinstatement (e.g., Hockley, 2008; Macken, 2002), or uses novel, unfamiliar contexts, as in applied investigations of context effects (e.g., Saenz & Smith, 2018; Wong & Read, 2011).

General Discussion

As students of memory, we often face the question, be it from colleagues at conferences or recruitment panels, regarding what it is that we do in our research. When we mention memory as our area of interest, it seems immediately obvious that we present participants in our experiments with some materials and later assess participants' memory for these materials. Indeed, it is very likely that students in our laboratories share this outlook at the experiments they take part in. This kind of research has been described by Koriat and Goldsmith (1994) as adopting the storehouse metaphor of memory: We investigate memory by asking people to encode information and later assessing how much of the originally encoded information can be retrieved and included in a memory report. From this perspective, the idea of a memory experiment in which the focus is on asking questions about details that were never presented to participants may seem peculiar. We argue, however, that the use of unanswerable questions provides unique opportunities for revealing the dynamics of memory reporting. In the present study, unanswerable questions were used to reveal how commission errors in a memory report can reflect more rather than less memory information at the responder's disposal.

When discussing unanswerable questions, it is important to distinguish between the perspective of an experimenter and the perspective of a participant. Unanswerable questions can be seen in at least two different ways. From the perspective of the experimenter, these questions may provide a methodological tool for investigating how conversion processes contribute to the accuracy of memory reports. The case of unanswerable questions directly highlights that the accuracy of participants' responses to memory questions does not depend solely on how effective memory retrieval is; it also depends on how people interpret the products of retrieval. The perspective of the experimenter was adopted throughout the present study, given that we argued that there is a strong distinction between answerable and unanswerable questions—a distinction we, as experimenters, were privy to because of the way we designed our materials and questions.

From the perspective of the participant, however, unanswerable questions may often differ much less from the answerable ones than the experimenters would like to assume. Obviously, for some of the unanswerable questions no plausible candidate response is likely to be generated and in those cases conversion processes will likely not play much of a role in memory reporting. For other unanswerable questions, however, at least some plausible candidate responses might be available and then the dynamics of responding for these questions will be very much the same as for answerable questions. This is clearly evidenced by the current results: for answerable and unanswerable questions alike, candidate responses were subjected to metacognitive monitoring, the accessibility of contextual information was used as a cue for this monitoring process, and control processes operated on the products of monitoring. As a result of this commonality of metacognitive processes engaged in responding to answerable and unanswerable questions, the distinction between these two types of memory queries—delineated by the experimenter—becomes blurred. This has important consequences for the accuracy of the reported responses: retrieval of *correct* contextual information can then lead to reporting of *incorrect* responses when unanswerable questions are asked.

[7] In both cases, the main effect of group stemmed from the inclusion of the warned group from Experiment 1 in the analysis. This was the only group in which participants were explicitly informed about the presence of unanswerable questions, and instructed to answer "don't know" whenever such a question was detected. These instructions resulted in an increase in the number of "don't know" responses above the level found in the remaining groups. When this group was excluded from both analyses, the model with only the main effect of context received the strongest support, and Bayes Factors for inclusion in the model actually revealed anecdotal evidence against including the factor of group.

Table 4
Bayesian Model Comparisons for Combined Analysis of Unanswerable Question Data Including Context (Reinstated, Repaired) and Group (Experiment 1 – Warning, Experiment 1 – No Warning, Experiment 2, Experiment 3 – Unanswerable Questions) as Factors

Model	p(M)	p(M|Data)	BF_M	BF_{10}	Error%
Context + Group	.20	.834	20.073	1.000	
Context	.20	.087	0.380	0.104	5.319
Context + Group + Context × Group	.20	.078	0.337	0.093	8.155
Group	.20	.002	0.006	0.002	7.695
Null	.20	1.55e-4	6.21e-4	1.86e-4	5.094

Note. p(M) shows uniform prior probabilities for all competing models. p(M|Data) shows posterior model probabilities. BF_M shows the change from prior to posterior odds. BF_{10} shows how much support compared to the best model (top row) each additional model provides, with values above 1 showing increasing evidence for the model, and below 1 showing increasing evidence against the model.

The Experimenter Perspective

The main feature of the experimenter perspective on unanswerable questions is that these questions differ sharply from answerable questions. Whereas target information for answerable questions can be stored and subsequently retrieved, such direct retrieval is precluded for unanswerable questions. This creates a unique situation in which any type of a specific response—other than an opt-out option such as "I don't know," or an explicit rejection of the question (stating that the question cannot be answered, or that the information needed was not available in the original stimuli; see Scoboria & Fisico, 2013)——can be treated as an error. Obviously, such a specific response may sometimes correspond to what happened in the past, if, for example, a person is lucky when formulating guesses or can infer the answer on some basis. But even in this case the experimenter knows that such a correct response cannot result from retrieval of target information that was solicited by a particular memory question.

Throughout this paper, we have argued that using unanswerable questions in experimental designs highlights the contribution of conversion processes to memory performance. Our theorizing remains largely based on the conceptual framework developed by Koriat and Goldsmith (1996b), who criticized the prevalent approach to memory which adopts what they dubbed the storehouse perspective on memory and which looks at how much of the originally encoded information can be reproduced at a later memory test. According to Koriat and Goldsmith, memory in everyday situations is often assessed based on the correspondence criterion: how well the reported information matches what actually happened in the past. The correspondence perspective shifts researchers' attention to two related aspects of memory reports: reporting of information not corresponding to the events from the past—often examined under the umbrella term of *false memory*—and the use of "don't know" responses as the means of editing out such erroneous reports. The present study follows this perspective by focusing on erroneous reports and "don't know" responding. The case for the use of unanswerable questions in memory research serves to underscore the importance of the correspondence perspective in assessing memory. While posing unanswerable questions makes little sense from the memory-as-a-storehouse perspective, it is immediately apparent how unanswerable questions relate to the correspondence perspective: Specific responses to unanswerable questions break correspondence between memory reports and originally encoded information.

The framework of Koriat and Goldsmith (1996b) provides a description of a particular type of conversion processes—translating candidate responses into overt responses included in memory reports—that was investigated in the present study. However, while to date these processes have been described in relation to answerable questions, the present study is the first one to demonstrate that similar dynamics in memory reporting can be observed for unanswerable questions. Thus, following Koriat and Goldsmith, we argue that when faced with a memory question, participants first try to formulate candidate responses to these questions and if this generation process is successful, candidate responses are subjected to metacognitive monitoring. The process of monitoring is based on the overall accessibility of memory information (see Koriat, 1993), with an important caveat that information feeding into the metacognitive process of appraisal does not need to strictly answer a particular memory question. As evidenced in our study, contextual information that remains related to the queried detail also determines how confident people are about their candidate responses and, as a result, how likely they are to volunteer these responses in a memory report.

By demonstrating the effects of contextual information on conversion, apart from retrieval of target information, we were able to reveal how context reinstatement can lead memory reporting astray. Interestingly, this is not the first demonstration of the perils of context reinstatement. Recently, Doss, Picart, and Gallo (2018) showed how reinstating original contexts for lures that were conceptually highly similar to, yet perceptually different from studied items elevates false alarms to these lures in recognition tests. This illusion of oldness likely results from the fact that context reinstatement aids retrieval of contextual associations that although correct—that is, corresponding to information encoded during study—leads to false reports via faulty conversion processes. For example, reinstating the encoding context for a picture of a table tennis racket can serve to augment retrieval of various semantic associations established at study—how one played table tennis at school, how one watched *Forrest Gump* for the first time, and so forth. All these true memories testify to the fact that a tennis racket was presented during study. However, the memory question is not really whether there was a tennis racket presented during study but whether it was the particular tennis racket that is included in a recognition test (which is not the case for highly similar lures). The results reported by Doss et al. suggest that despite this specific nature of a memory query, participants are willing to endorse

similar lures as studied owing to improved access to semantic information that does not serve to answer the memory question at hand. We argue that a common theme for our study and the study by Doss et al. is that there is a wide variety of information that can be retrieved due to context reinstatement—semantic, perceptual, emotional, and so forth—and all these types of information will determine responses to memory queries. This happens even when the retrieved information is not strictly relevant, that is, it does not provide access to target information solicited by the particular question. Thus, there is a host of conversion processes that mediate between retrieval and memory responses that are capable on their own of producing errors in memory reports. In other words, not all memory errors stem from false recollections of details that were not in fact encoded.

The Participant Perspective

When designing a memory experiment, the experimenter knows what participants are presented with for study. By being privy to the materials and design used in a given experiment, the experimenter thus knows that some questions are unanswerable because participants were not presented with information directly relevant to these questions. This knowledge, however, is not available to participants. Participants may either expect all questions to be answerable based on the materials that were studied, or suspect that unanswerable questions are asked. These two situations may create different demands for memory reporting.

Starting with the scenario in which participants don't suspect that unanswerable questions are asked, we believe it to be a common situation for participants in laboratory experiments on memory. As researchers, we usually inform our participants that they will be presented with certain materials to learn and later tested on their memory which may create expectations that indeed only questions about studied materials will be asked. This creates a situation similar to a student facing an exam: the student can legitimately expect that only questions pertaining to the materials that were presented in lectures or included in the reading materials will be asked. Importantly, this may constitute very specific and restricted conditions for memory reporting. In various situations, people may approach a memory task with very different assumptions. An eyewitness in an interview may fully expect to be asked unanswerable questions concerning information that was never encoded in the first place.

When participants don't expect unanswerable questions to be asked, those questions may often become subjectively very similar to answerable questions for which target information is not remembered. In both cases, candidate responses need to be generated by means other than direct retrieval of target information. When such responses are in fact generated, confidence in their accuracy is assessed based on various factors such as their plausibility or—as demonstrated in the present study—retrieval of related or contextual details, before a decision is made whether these candidate responses should be volunteered or withheld. Consistent with the argument according to which answerable and unanswerable questions are subjected to the same set of conversion processes governing their reporting is the observation that in all three experiments the manipulations of retrieval context had indistinguishable effects on reporting for these two types of questions. The only exception to this pattern—an apparently smaller effect of context reinstatement on retrospective confidence in Experiment 2 for unanswerable questions—is most plausibly explained as a statistical artifact reflecting scaling effects. We argue that this effect stems from generally low levels of confidence for responses to unanswerable questions. This low confidence indicates that whenever a specific response is provided to an unanswerable question, it likely reflects an inference, not a subjectively compelling false memory (for evidence that context reinstatement can sometimes create false memories see Dodson, 2007).

Whenever inferences are provided in response to memory queries, one should also consider social aspects of memory reporting. Specifically, we argue that the norms of communication and the implicit requirement to be informative dictate that not all questions should be met with a "don't know" response (Ackerman & Goldsmith, 2008; Martín-Luengo, Shtyrov, Luna, & Myachykov, 2018). This is particularly relevant when considering the results of Experiment 3, in which we provided one group of participants with unanswerable questions only. In this case, participants were not able to retrieve any target information throughout the entire test. Still, they did provide specific responses and the rates of that specific responding were again affected by retrieval context. If false recollections are rare under conditions created in our study, as confidence data from Experiment 2 would suggest, then the task of responding to a set of questions for which one remembers very little turns into inferring for which of these questions one's candidate responses are most likely to be correct, even if they don't reflect retrieval of target information from memory. In other words, participants likely feel they need to use whatever information they have at their disposal to solve the problem of identifying questions for which the chances of reducing the accuracy of a memory report—by providing specific answers—are minimized. As a result, their responses reflect inferences based, among other factors, on recognized or recalled contextual details.

So far we have discussed a scenario in which participants assume that all questions are answerable. But our results suggest that conversion processes may change with changes in participants' assumptions. In Experiment 1, we tested participants for whom no attempt was made to revise their initial assumptions as to the nature of a memory task, as well as participants who were explicitly informed that some of the questions may be unanswerable. This warning manipulation had an effect on how likely participants were to volunteer responses in the final memory test. Specifically, warned participants volunteered fewer responses, using the "don't know" option more often instead. This effect, however, was specific as it occurred for unanswerable questions, with only a trend in the same direction observed for answerable questions.

There seem to be two explanations for this interactive pattern. First, one can assume that the warning makes participants overall less willing to report candidate responses, thus effectively changing the report criterion, as described in the model by Koriat and Goldsmith (1996b) in relation to incentive manipulations. The most straightforward formulation of this mechanism would predict similar effects of warning for answerable and unanswerable questions—akin to response criterion shifts affecting equally hits and false alarms in recognition tests. However, a more complex account can also be built, according to which candidate responses for unanswerable questions are on average closer in terms of confidence to the criterial level of confidence warranting inclusion in a

memory report and thus they are more affected by changes in this criterial level. Although such an account is possible, we note that it does face a challenge of explaining why a similar interaction did not emerge for the manipulation of context. Specifically, if confidence in candidate responses for unanswerable questions is closer to the criterial level of confidence than confidence in candidate responses for answerable questions, then changing these baseline levels of confidence via context reinstatement should also affect the former type of questions more than the latter. This, however, was not observed across our three experiments.

Second, if the warning manipulation did not serve simply to alter the report criterion, then the specific effect it had on unanswerable questions suggests the change in the basis of responding to these questions. It is possible that participants used the warning to search information that would allow them to identify particular questions as unanswerable. A study on the meaning of "don't know" responses to answerable and unanswerable questions conducted by Scoboria et al. (2008) showed that participants using "don't know" responses may try to express two different cognitive states. In the *not knowing* state, the "don't know" option reflects insufficient information concerning the queried detail. This meaning of the "don't know" option is addressed in the current models of response volunteering, such as the framework developed by Koriat and Goldsmith (1996b). In the *knowing not* state, however, the "don't know" option reflects participants' certainty that relevant information was not available to be encoded and thus a given question is unanswerable. A similar use of "don't know" responses for general knowledge questions was recently reported by Coane and Umanath (2019). It is this latter meaning of the "don't know" option that was promoted by the warning provided in Experiment 1 of the present study. Indeed, Scoboria and Fisico (2013) have demonstrated that specific warning about unanswerable questions, coupled with instructions designed to encourage the use of a "don't know" option, can serve to increase the rate by which "don't know" responses reflecting 'knowing not' states are given. If our manipulation of warning had the same effect, then this suggests that participants in laboratory settings may generally not realize that unanswerable questions are present unless explicitly informed by the experimenter or interviewer—a reasonable assumption given that most experiments indeed require participants only to remember the information provided in the study materials.

Finally, we note that a specific warning may not always be necessary for positively identifying questions as unanswerable. One can certainly envision questions that are sure to alert participants as to their nature. This can be achieved, for example, by providing particularly implausible alternatives that participants would not regard as potential candidate responses. Much like highly conspicuous misinformation is likely to serve as an implicit warning as to unreliability of the account in which misinformation is embedded (Loftus, 1979), highly conspicuous unanswerable questions may serve as a warning about the possibility of unanswerable questions being asked. Such a warning could also help responders to arrive at the knowing not state, where they use the "don't know" option to express their conviction that target information was not encoded. Thus, although we have argued earlier that in relation to conversion processes unanswerable questions often behave like answerable ones—as evidenced by the results of contextual manipulations included in the present study—this should not be taken as a general law for which there are no exceptions.

Future Directions

The present investigation constitutes the first step toward elucidating the role of conversion processes in responding to unanswerable questions and thus it used laboratory procedures providing a high degree of experimental control over manipulated variables, in particular over the effects of context reinstatement. For the sake of this control, we compared the reinstatement effects against the baseline of repaired context, thus equating the two compared conditions in terms of memory for the context itself. However, we did also include a less stringent comparison of reinstated contexts to novel contexts. We argue that such a comparison—generally producing larger effects on "don't know" responding in Experiments 1 and 3 (see Table 1)—provides the basis on which further investigations into context effects on unanswerable questions should proceed. This is because this comparison is of much more use from an applied perspective. Remaining of particular interest here are various issues related to eyewitness testimony. The context reinstatement technique already serves as an important tool for augmenting eyewitnesses' memory in forensic contexts, being the part of the current golden standard of interviewing protocols, the Cognitive Interview (Fisher & Geiselman, 1992). Context reinstatement in this applied instantiation means that eyewitnesses are usually tested in novel contexts such as a police station and they are directed toward mental context reinstatement, trying to imagine the particulars of physical context, mood, thoughts, and so forth that accompanied the event they testify to. What our research suggests is that even though there are clear benefits of context reinstatement when answerable questions are asked, there may also be potential costs associated with this technique when an interviewer asks an unanswerable question. Of course, our laboratory investigations differ on many counts from context reinstatement as used in applied settings, particularly in the fact that we provided contexts for our participants rather than asked participants to mentally reinstate them, but our findings at least point to the need for a systematic pursuit of the issue of unanswerable questions in applied settings. Indeed, our findings suggest that context reinstatement at least in certain cases may prove problematic for responding to unanswerable questions exactly because it is such an effective tool for augmenting retrieval of correct information, which also points to the need for examining other techniques designed to augment memory in relation to unanswerable questions.

One other important difference between our laboratory-based procedures and situations in which unanswerable questions can be asked in nonlaboratory settings is that we pursued "don't know" responding in recognition tests, whereas outside the laboratory unanswerable questions are more likely to take the form of recall questions. Our design was dictated by the focus on conversion processes, but we noted earlier that the dynamics of responding to unanswerable questions can look different when the whole process of responding is more dependent on generating plausible candidate responses, a process to a large extent circumvented with the use of recognition alternatives. Specifically, it is possible that context reinstatement may have different effects depending on whether plausible candidate responses suggest themselves when a memory

query is posed. When such candidate responses are not generated, as may sometimes happen for unanswerable questions in a recall format, retrieval of contextual associations may actually alert people that they deal with questions for which target information was never encoded, facilitating "don't know" responding reflecting knowing not states. Although at present speculative, the role of availability of plausible candidate responses on responding to unanswerable questions is a clear avenue for further research.

From a broader perspective, it seems vital to pursue a more general question of how appropriate "don't know" responding to unanswerable questions can be promoted. The creation of an appropriate task representation, one that would reflect the possibility of being asked unanswerable questions, is clearly one potential avenue, as the current consideration of the warning manipulation suggests. We suggest that another avenue could be found in a burgeoning literature on monitoring incorrect responses in recognition tests (see Goldsmith, 2016). These monitoring processes involve the mechanisms of early selection—the restriction of memory search to features diagnostic of study occurrence (Jacoby, Shimizu, Daniels, & Rhodes, 2005; Zawadzka, Hanczakowski, & Wilding, 2017), and the mechanisms of late correction—such as the distinctiveness heuristic, by which a person is able to judge an item as new because it possesses distinctive features that would be remembered had the item actually been studied (Dodson & Schacter, 2002; Hanczakowski & Mazzoni, 2011). It would be of interest to see whether the same mechanisms are responsible for monitoring one's state of knowledge to arrive at a judgment concerning the answerability of a question. For example, can one decide that a question is not answerable—and hence a "don't know" response should be given—based on the potential distinctiveness of the queried detail? The unanswerable question may concern such a conspicuous feature of an environment that a person can be quite certain that this feature was not properly perceived during the event because otherwise it would be remembered. We believe that the ways of promoting appropriate "don't know" responding to unanswerable questions by examining response formats, warning interventions, and a variety of monitoring strategies should become a priority for research into unanswerable questions, both in laboratory and applied settings.

Context of Research

The current research was conceived as a project linking interests of two different research teams. Alan Scoboria, who sadly passed away in April 2019, was initially interested in the consequences of posing unanswerable questions in the applied context of forensic interviews. Alan was one of the first researchers who delved into specifics of what people try to convey when they say "don't know" in response to memory questions, pointing out that at least sometimes these "don't know" responses are exactly what the interviewer should expect from an interviewee. This reorientation of the understanding of "don't know" responses—away from treating them as errors of omissions and toward their understanding as correct responses to unanswerable questions—served as an impetus for the remaining authors of the paper, who have for some time worked on describing the processes involved in metacognitive regulation of memory reporting, particularly as a function of context manipulations. This led directly to a question of whether the same dynamics that shape metacognitive processes when answerable questions are concerned operate—and serve directly to change the accuracy of reported responses by modulating "don't know" responding—in the case of unanswerable questions.

References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting—With and without satisfying knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1224–1245. http://dx.doi.org/10.1037/a0012938

Blank, H. (1998). Memory states and memory tasks: An integrative framework for eyewitness memory and suggestibility. *Memory, 6,* 481–529. http://dx.doi.org/10.1080/741943086

Borckardt, J. J., Sprohge, E., & Nash, M. (2003). Effects of the inclusion and refutation of peripheral details on eyewitness credibility. *Journal of Applied Social Psychology, 33,* 2187–2197. http://dx.doi.org/10.1111/j.1559-1816.2003.tb01880.x

Burgess, N., Hockley, W. E., & Hourihan, K. L. (2017). The effects of context in item-based directed forgetting: Evidence for "one-shot" context storage. *Memory & Cognition, 45,* 745–754. http://dx.doi.org/10.3758/s13421-017-0692-5

Coane, J. H., & Umanath, S. (2019). I don't remember vs. I don't know: Phenomenological states associated with retrieval failures. *Journal of Memory and Language, 107,* 152–168. http://dx.doi.org/10.1016/j.jml.2019.05.002

Cox, J. C., & Dobbins, I. G. (2011). The striking similarities between standard, distractor-free, and target-free recognition. *Memory & Cognition, 39,* 925–940. http://dx.doi.org/10.3758/s13421-011-0090-3

Dodson, C. S. (2007). Retrieval-based illusory recollections: Why study-test contextual changes impair source memory. *Memory & Cognition, 35,* 1211–1221. http://dx.doi.org/10.3758/BF03193595

Dodson, C. S., & Schacter, D. L. (2002). When false recognition meets metacognition: The distinctiveness heuristic. *Journal of Memory and Language, 46,* 782–803. http://dx.doi.org/10.1006/jmla.2001.2822

Doss, M. K., Picart, J. K., & Gallo, D. A. (2018). The dark side of context: Context reinstatement can distort memory. *Psychological Science, 29,* 914–925. http://dx.doi.org/10.1177/0956797617749534

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.

Franco-Watkins, A. M., & Dougherty, M. R. (2006). Word concreteness and encoding effects on context-dependent discrimination. *Memory & Cognition, 34,* 973–985. http://dx.doi.org/10.3758/BF03193245

Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recognition tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory and Language, 51,* 473–493. http://dx.doi.org/10.1016/j.jml.2004.06.002

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66,* 325–331. http://dx.doi.org/10.1111/j.2044-8295.1975.tb01468.x

Goldsmith, M. (2016). Metacognitive quality-control processes in memory retrieval and reporting. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 357–385). New York, NY: Oxford University Press.

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language, 52,* 505–525. http://dx.doi.org/10.1016/j.jml.2005.01.010

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size in memory reporting. *Journal of Experimental Psychology: General, 131,* 73–95. http://dx.doi.org/10.1037/0096-3445.131.1.73

Hanczakowski, M., & Mazzoni, G. (2011). Both differences in encoding processes and monitoring at retrieval reduce false alarms when distinc-

tive information is studied. *Memory, 19,* 280–289. http://dx.doi.org/10.1080/09658211.2011.558514

Hanczakowski, M., Pasek, T., Zawadzka, K., & Mazzoni, G. (2013). Cue familiarity and 'don't know' responding in episodic memory tasks. *Journal of Memory and Language, 69,* 368–383. http://dx.doi.org/10.1016/j.jml.2013.04.005

Hanczakowski, M., Zawadzka, K., Collie, H., & Macken, B. (2017). Metamemory in a familiar place: The effects of environmental context on feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 59–71. http://dx.doi.org/10.1037/xlm0000292

Hanczakowski, M., Zawadzka, K., & Coote, L. (2014). Context reinstatement in recognition: Memory and beyond. *Journal of Memory and Language, 72,* 85–97. http://dx.doi.org/10.1016/j.jml.2014.01.001

Hanczakowski, M., Zawadzka, K., & Macken, B. (2015). Continued effects of context reinstatement in recognition. *Memory & Cognition, 43,* 788–797. http://dx.doi.org/10.3758/s13421-014-0502-2

Hertzog, C., Fulton, E. K., Sinclair, S. M., & Dunlosky, J. (2014). Recalled aspects of original encoding strategies influence episodic feelings of knowing. *Memory & Cognition, 42,* 126–140. http://dx.doi.org/10.3758/s13421-013-0348-z

Higham, P. A. (2007). No special K! A signal detection framework for the strategic regulation of memory accuracy. *Journal of Experimental Psychology: General, 136,* 1–22. http://dx.doi.org/10.1037/0096-3445.136.1.1

Higham, P. A., & Arnold, M. M. (2007). How many questions should I answer? Using bias profiles to estimate optimal bias and maximum score on formula-scored tests. *European Journal of Cognitive Psychology, 19,* 718–742. http://dx.doi.org/10.1080/09541440701326121

Hockley, W. E. (2008). The effects of environmental context on recognition memory and claims of remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1412–1429. http://dx.doi.org/10.1037/a0013016

Hollins, T. J., & Weber, N. (2017). Evidence of a metacognitive benefit to memory? *Memory, 25,* 317–325. http://dx.doi.org/10.1080/09658211.2016.1171363

Isarida, T., Isarida, T. K., & Sakai, T. (2012). Effects of study time and meaningfulness on environmental context-dependent recognition. *Memory & Cognition, 40,* 1225–1235. http://dx.doi.org/10.3758/s13421-012-0234-0

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review, 12,* 852–857. http://dx.doi.org/10.3758/BF03196776

JASP Team. (2019). JASP (Version 0.9) [Computer software]. Retrieved from https://jasp-stats.org/

Koen, J. D., Aly, M., Wang, W. C., & Yonelinas, A. P. (2013). Examining the causes of memory strength variability: Recollection, attention failure, or encoding variability? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1726–1741. http://dx.doi.org/10.1037/a0033671

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639. http://dx.doi.org/10.1037/0033-295X.100.4.609

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124,* 311–333. http://dx.doi.org/10.1037/0096-3445.124.3.311

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123,* 297–315. http://dx.doi.org/10.1037/0096-3445.123.3.297

Koriat, A., & Goldsmith, M. (1996a). Memory metaphors and the real life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences, 19,* 167–188. http://dx.doi.org/10.1017/S0140525X00042114

Koriat, A., & Goldsmith, M. (1996b). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517. http://dx.doi.org/10.1037/0033-295X.103.3.490

Koriat, A., Levy-Sadot, R., Edry, E., & de Marcas, S. (2003). What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1095–1105. http://dx.doi.org/10.1037/0278-7393.29.6.1095

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge, UK: Cambridge University Press.

Loftus, E. F. (1979). Reactions to blatantly contradictory misinformation. *Memory & Cognition, 7,* 368–374. http://dx.doi.org/10.3758/BF03196941

Luna, K., & Martín-Luengo, B. (2012). Improving the accuracy of eyewitnesses in the presence of misinformation with the plurality option. *Applied Cognitive Psychology, 26,* 687–693. http://dx.doi.org/10.1002/acp.2845

Macken, W. J. (2002). Environmental context and recognition: The role of recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 153–161. http://dx.doi.org/10.1037/0278-7393.28.1.153

Martín-Luengo, B., Shtyrov, Y., Luna, K., & Myachykov, A. (2018). Different answers to different audiences: Effects of social context on the accuracy-informativeness trade-off. *Memory, 26,* 993–1007. http://dx.doi.org/10.1080/09658211.2017.1420196

Murnane, K., & Phelps, M. P. (1993). A global activation approach to the effect of changes in environmental context on recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 882–894. http://dx.doi.org/10.1037/0278-7393.19.4.882

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128,* 403–415. http://dx.doi.org/10.1037/0096-3445.128.4.403

Pierce, B. H., Gallo, D. A., & McCain, J. L. (2017). Reduced interference from memory testing: A postretrieval monitoring account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 1063–1072. http://dx.doi.org/10.1037/xlm0000377

Roebers, C. M., & Fernandez, O. (2002). The effects of accuracy motivation on children's and adults' event recall, suggestibility, and their answers to unanswerable questions. *Journal of Cognition and Development, 3,* 415–443. http://dx.doi.org/10.1080/15248372.2002.9669676

Russo, R., Ward, G., Geurts, H., & Scheres, A. (1999). When unfamiliarity matters: Changing environmental context between study and test affects recognition memory for unfamiliar stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 488–499. http://dx.doi.org/10.1037/0278-7393.25.2.488

Saenz, G. D., & Smith, S. M. (2018). Testing judgments of learning in new contexts to reduce confidence. *Journal of Applied Research in Memory & Cognition, 7,* 540–551. http://dx.doi.org/10.1016/j.jarmac.2018.07.003

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1074–1083. http://dx.doi.org/10.1037/0278-7393.18.5.1074

Schwartz, B. L., Pillot, M., & Bacon, E. (2014). Contextual information influences the feeling of knowing in episodic memory. *Consciousness and Cognition, 29,* 96–104. http://dx.doi.org/10.1016/j.concog.2014.08.018

Scoboria, A., & Fisico, S. (2013). Encouraging and clarifying "don't know" responses enhances interview quality. *Journal of Experimental Psychology: Applied, 19,* 72–82. http://dx.doi.org/10.1037/a0032067

Scoboria, A., Mazzoni, G., & Kirsch, I. (2008). "Don't know" responding to answerable and unanswerable questions during misleading and hypnotic interviews. *Journal of Experimental Psychology: Applied, 14,* 255–265. http://dx.doi.org/10.1037/1076-898X.14.3.255

Smith, S. M., Handy, J. D., Angello, G., & Manzano, I. (2014). Effects of similarity on environmental context cueing. *Memory, 22,* 493–508. http://dx.doi.org/10.1080/09658211.2013.800553

Smith, S. M., & Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods, 42,* 292–301. http://dx.doi.org/10.3758/BRM.42.1.292

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8,* 203–220. http://dx.doi.org/10.3758/BF03196157

Tenney, E. R., MacCoun, R. J., Spellman, B. A., & Hastie, R. (2007). Calibration trumps confidence as a basis for witness credibility. *Psychological Science, 18,* 46–50. http://dx.doi.org/10.1111/j.1467-9280.2007.01847.x

Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2011). Context affects feeling-of-knowing accuracy in younger and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 96–108. http://dx.doi.org/10.1037/a0021612

Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2012). An analysis of the determinants of the feeling of knowing. *Consciousness and Cognition, 21,* 1681–1694. http://dx.doi.org/10.1016/j.concog.2012.09.005

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Clarendon Press/Oxford University Press.

Waterman, A. H., & Blades, M. (2011). Helping children correctly say "I don't know" to unanswerable questions. *Journal of Experimental Psychology: Applied, 17,* 396–405. http://dx.doi.org/10.1037/a0026150

Wong, C. K., & Read, J. D. (2011). Positive and negative effects of physical context reinstatement on eyewitness recall and identification. *Applied Cognitive Psychology, 25,* 2–11. http://dx.doi.org/10.1002/acp.1605

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. http://dx.doi.org/10.1111/1467-9280.00388

Zawadzka, K., Hanczakowski, M., & Wilding, E. L. (2017). Late consequences of early selection: When memory monitoring backfires. *Journal of Memory and Language, 92,* 114–127. http://dx.doi.org/10.1016/j.jml.2016.06.003

Zawadzka, K., Krogulska, A., Button, R., Higham, P. A., & Hanczakowski, M. (2016). Memory, metamemory, and social cues: Between conformity and resistance. *Journal of Experimental Psychology: General, 145,* 181–199. http://dx.doi.org/10.1037/xge0000118

Zawadzka, K., Simkiss, N., & Hanczakowski, M. (2018). Remind me of the context: Memory and metacognition at restudy. *Journal of Memory and Language, 101,* 1–17. http://dx.doi.org/10.1016/j.jml.2018.03.001

Received June 27, 2017
Revision received August 8, 2019
Accepted August 30, 2019 ■