

Context reinstatement in recognition: Memory and beyond

Maciej Hanczakowski¹, Katarzyna Zawadzka², and Laura Coote¹

¹Cardiff University, UK

²University of Southampton, UK

Author Notes

Maciej Hanczakowski, School of Psychology, Cardiff University, UK; Katarzyna Zawadzka, Psychology, University of Southampton, UK; Laura Coote, School of Psychology, Cardiff University, UK.

We thank Bill Macken for his helpful comments on an earlier version of the manuscript.

Correspondence concerning this article should be addressed to Maciej Hanczakowski, School of Psychology, Cardiff University, Tower Building, Cardiff, CF10 3AT, United Kingdom, email: HanczakowskiM@cardiff.ac.uk, ph: +44 (0) 29 2087 5030.

Article published in [Journal of Memory and Language](#). The final version is available at <http://dx.doi.org/10.1016/j.jml.2014.01.001>

© 2014. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Abstract

Context effects in recognition tests are twofold. First, presenting familiar contexts at a test leads to an attribution of context familiarity to a recognition probe, which has been dubbed ‘context-dependent recognition’. Second, reinstating the exact study context for a particular target in a recognition test cues recollection of an item-context association, resulting in ‘context-dependent discrimination’. Here we investigated how these two context effects are expressed in metacognitive monitoring (confidence judgments) and metacognitive control (‘don’t know’ responding) of retrieval. We used faces as studied items, landscape photographs as study and test contexts and both free- and forced-report 2AFC recognition tests. In terms of context-dependent recognition, the results document that presenting familiar contexts at test leads to higher confidence and lower rates of ‘don’t know’ responses compared to novel contexts, while having no effect on forced-report recognition accuracy. In terms of context-dependent discrimination, the results show that reinstated contexts further boost confidence and reduce ‘don’t know’ responding compared to familiar contexts, while affecting forced-report recognition accuracy only when contribution of recollection to recognition performance is high. Together, our results demonstrate that metacognitive measures are sensitive to context effects, sometimes even more so than recognition measures.

Keywords: Context, Metacognition, Recognition, Confidence

Both encoding information in memory and its later retrieval occur in context. Context can be understood as any type of information that accompanies encoding and retrieval but is not itself a target of either encoding or retrieval. In experimental studies on memory, a variety of conditions have been treated as context, ranging from mood (e.g., Eich, 1985; Eich & Metcalfe, 1989) to position of a word on a computer screen (e.g., Macken, 2002; Murnane & Phelps, 1993). Context has been investigated mostly to determine if reinstating study context at the moment of testing facilitates memory retrieval. A substantial number of studies document such beneficial effects of context reinstatement in recall (e.g., Godden & Baddeley, 1975; Smith, Glenberg, & Bjork, 1978), but a somewhat more complex picture emerges in recognition, with some studies showing benefits of reinstating context for recognition performance (e.g., Geiselman & Bjork, 1980; Smith, 1985) and some showing no benefits (Hockley, 2008; see also Smith & Vela, 2001, for a review). In the present study we take a novel but complementary approach to investigating context effects in memory as we focus on how context present at retrieval affects metacognitive processes.

The metacognitive approach to memory stresses that the process of remembering does not end when retrieval of information from memory is completed. When asked a memory question, people need not only gather information from memory. They need also to assess the quality of the products of memory retrieval and to decide whether this quality is sufficient to warrant reporting of the retrieved memory. According to the framework developed by Koriat and Goldsmith (1996), the processes of responding to a memory question start with generating a candidate answer based on information stored in memory. After the memory retrieval process is complete, the metacognitive monitoring process takes over and a person assesses his/her confidence that the product of memory retrieval (the candidate response) is correct. Finally, after assigning confidence to the candidate response, in the process of exerting metacognitive control, a decision is taken whether assigned confidence is high

enough to volunteer the candidate response, or, alternatively, whether a ‘don’t know’ (DK) response to the memory question should be given. Critically, according to the Koriat and Goldsmith framework, the metacognitive processes are mixed with memory retrieval to shape memory performance in free-report tests (i.e., in tests allowing response withholding). This is so because in such tests omissions may reflect either failures to access appropriate memory information, or the state in which required information is accessible but confidence assigned to this information is too low to warrant its disclosure.

Despite years of research on how retrieval context affects memory processes, the issue of whether metacognitive processes of monitoring and control are affected by changes in study-test contexts have not been systematically investigated. It is an important gap in our knowledge, since, as discussed above, metacognitive processes shape the contents of a memory report. If context present at retrieval were to affect how confident people are about the products of their memory processes, then, according to the Koriat and Goldsmith (1996) framework, it would also determine the probability of volunteering retrieved information in free-report memory tests, affecting the number of correct and incorrect details reported. Thus, for example, context reinstatement may benefit memory retrieval and at the same time make people more confident that the products of retrieval are correct, increasing the chances that retrieved details are disclosed. But it is also possible that context would affect metacognitive processes even when it has no effect on memory. As described in the next section, research on context effects in recognition identified conditions under which context present at study and later provided at test failed to affect recognition accuracy in forced-report tests. However, if context were to affect metacognitive processes under such circumstances, then it could lead to changes in free-report recognition output, demonstrating how context may exert influence upon memory performance via metacognitive, not memory processes. Such a demonstration was the main motivation behind conducting the present study.

In the present study we aim to investigate how changes in study-test contexts affect metacognitive processes. To this aim, we borrow the paradigm previously used to examine the effects of context reinstatement on recognition performance and we use this paradigm to investigate how study-test contexts determines confidence (metacognitive monitoring) and decisions whether to volunteer responses to a memory question or to respond DK (metacognitive control). In what follows, we first present an overview of the literature on context reinstatement in recognition and then we outline our predictions of how context may affect metacognitive processes.

Context effects in recognition

The most comprehensive work on context effects in recognition was conducted by Murnane and Phelps (1993, 1994, 1995) within the theoretical framework ICE (Item, Context, and Ensemble information) theory (cf. Murnane, Phelps, & Malmberg, 1999). According to this theory, when items are presented with context at study, three types of information can be encoded: a) item information, b) context information, and c) information specific to an ensemble created by an item and its context, to which we will refer here as an item-context association¹. When, in a subsequent old/new recognition test, a studied context is re-presented together with a novel or old probe, it matches the stored context information, resulting in a feeling of familiarity. Familiarity of the context is attributed to the test probe, increasing the probability of an 'old' response. Importantly, this occurs whenever studied context is used in a test, independently of whether the probe it accompanies corresponds to an item studied in this particular context. The same effect on 'old' responses occurs both for

¹ In the present work we do not differentiate between a global-matching approach to recognition memory, as advocated by Murnane *et al.* (1999), and a dual-process approach promoted by other researchers (e.g., Macken, 2002), as we believe that these two are quite similar in their descriptions of the context effects. In other words, from the perspective of our study, we do not see much difference between the concept of matching ensemble information to contents of a memory store and recollection of item-context associations. For convenience, we use the term of recollection of item-context associations in the present paper, rather than the global matching terminology of ICE.

targets studied in different contexts and for foils which were not studied at all. In consequence, studied contexts presented at test increase both hits to studied items, whether they were paired with this particular context or not, and false alarms to foils. This type of effect Murnane *et al.* dubbed *context-dependent recognition*.

A different effect may sometimes occur when at test context is re-presented with the same particular item with which it was paired at study. In such a case, inclusion of both the item and its originally paired context in a compound cue may result in recollection of the item-context association. Recollection of this association also induces more 'old' responses. However, such a recollection necessarily occurs only for studied items and thus recollection of item-context associations specifically increases hits to old items but not false alarms to foils. This type of effect Murnane *et al.* (1999) dubbed *context-dependent discrimination*, as a specific increase in hits for targets means that participants are better at discriminating between targets and foils.

The empirical studies on context effects largely followed the directions outlined within the ICE theory. Hockley, Bancroft, & Bryant (2012) reviewed the results of a number of conditions employed in various studies on context effects in recognition and found that false alarm rates to foils tested in studied contexts are invariably larger than false alarm rates to foils tested in novel contexts, which supports the idea of context-dependent recognition (see also Dodson & Shimamura, 2000; Hockley, 2008). However, the issue of context-dependent discrimination remains more controversial. To make a strong case for context-dependent discrimination, a comparison is needed between a condition in which the exact context accompanying encoding of a given item is reinstated at test for this item (a reinstated context condition) with a condition in which a studied context is used at test in a novel target-context configuration (a re-paired context condition). Such a comparison holds the familiarity of a context equal, eliminating the influence of context-dependent recognition, and thus

allows for revealing context-dependent discrimination. A study by Macken (2002) employed such a comparison in a remember/know paradigm and found evidence for context-dependent discrimination in recognition judgments accompanied by 'remember' responses (see also Tibon, Vakil, Goldstein, & Levy, 2012, for related results). On the other hand, a comprehensive investigation of context effects by Hockley (2008) showed that context-dependent discrimination requires a specific orienting task at encoding. Only when participants were explicitly instructed to associate items with their contexts at study, did the hit rate in the reinstated context condition (and also the rate of 'remember' responses) exceed the hit rate in the re-paired context condition, indicating context-dependent discrimination. Thus, the study by Hockley suggests that the amount of attention devoted to encoding item-context associations is crucial for the presence of context-dependent discrimination. When attention is not directed toward such associations, they are not very likely to be retrieved even when the exact context is reinstated at test, impeding the chances of documenting context-dependent discrimination.

Several studies used faces, the stimuli adopted for the present investigation, as materials in research on context reinstatement (Kerr & Winograd, 1982; Watkins, Ho, & Tulving, 1976; Winograd & Rivers-Bulkeley, 1977). These studies found indications of both context-dependent recognition (increase in false alarms with familiar contexts) and context-dependent discrimination in memory for faces (as assessed either with corrected recognition scores or forced-choice recognition accuracy). Importantly, all these studies used associative instructions at study, designed to promote creation of associations between target faces and their contexts. Two more recent studies that examined context effects in memory for faces included a re-paired context condition, which allows for dissociating the effects of context familiarity and reinstating the specific item-context pairing. Gruppuso, Lindsay, and Masson (2007) used associative study instructions at study and found a higher hit rate in the re-paired

context condition compared to the novel context condition and a higher false alarm rate to foils in studied contexts compared to foils in novel contexts, supporting the concept of context-dependent recognition. They also found an increased hit rate and the rate of ‘remember’ responses in the reinstated context condition as compared to the re-paired context condition, consistent with the pattern observed by Macken (2002) and also by Hockley (2008, Experiment 5B), and supporting context-dependent discrimination. Reder *et al.* (2013) also used associative study instructions but found no evidence for context-dependent discrimination for unknown faces (although such evidence emerged for famous faces). Reder *et al.* hypothesized that it could be particularly difficult for participants to create associations between unknown faces and context images that would later support recollection. These results underscore the point made by Hockley (2008) that documenting context-dependent discrimination is unlikely when creation of item-context associations at study is impeded.

The present study

In the present study we asked how context-dependent recognition and discrimination affect metacognitive processes of monitoring and control of retrieval. We start outlining our predictions with context-dependent discrimination, which is a facilitation of retrieval of item-context associations for reinstated contexts. Numerous studies have shown that metacognitive processes at retrieval are strongly affected by recollective processes (e.g., Bulevich & Thomas, 2012; Koriat, 1993). It is a straightforward prediction that whenever recollection is triggered by context reinstatement, participants should become more confident in their candidate responses, which should be reflected in both confidence judgments concerning forced-report responses (confidence should be higher when recollection is augmented) and also their decisions to volunteer candidate responses in a free-report test (more responses should be volunteered when recollection is augmented). Thus, in the case of context-

dependent discrimination, we essentially predict parallel effects on memory and metacognitive measures.

A potentially more interesting scenario arises for context-dependent recognition, where we predicted a dissociation of memory and metacognitive measures. The effect of context-dependent recognition occurs when context familiarity is attributed to a recognition probe, increasing both hits and false alarms in old/new recognition test but not affecting recognition discrimination. The question that we asked was whether context familiarity would also increase confidence in a candidate response and increase the probability that this candidate response would be volunteered. Our previous investigation of the effects of familiarity on metacognitive processes (Hanczakowski, Pasek, Zawadzka, & Mazzoni, 2013) suggests that familiarity that is irrelevant for recognition discrimination may nevertheless affect confidence and DK responding. In this study, we showed that when paired associates are studied and the recognition task requires indicating a target which was paired with a given cue, cue familiarity is attributed to a candidate response in a recognition test. This inflates confidence that this response is correct and makes it more likely that this response would be volunteered. Although the cue in the paired-associates task differs from context inasmuch as the cue is vital for the task at hand (choosing target associated with this cue), whereas context is not, we predicted that the effects of familiarity will be the same in these two scenarios. Thus, we predicted that context familiarity would increase confidence that a candidate response is correct and reduce DK responding, while having no effect on forced-report recognition accuracy.

To test our predictions, we adopted the paradigm used to investigate context reinstatement in recognition. In our particular version of the paradigm participants studied compounds of face and landscape photographs and their memory for faces was later tested with a two-alternative forced-choice (2AFC) recognition test with a context (landscape)

photograph placed between test stimuli.² Three context conditions were investigated: reinstated context, re-paired context and novel context. A comparison of re-paired and novel context conditions speaks to the effects of context familiarity (context-dependent recognition) whereas a comparison of reinstated and re-paired context conditions speaks to the effects of reinstating the association between the test item and its context, while keeping context familiarity constant (context-dependent discrimination).

Critically, we supplemented our 2AFC recognition test with a stage in which metacognitive monitoring in the shape of confidence was assessed, as well as with a stage in which exerting metacognitive control was allowed and participants could withhold a response. Thus, in each trial of the test, participants were presented with test stimuli (two faces and a context photograph) and they were first asked to indicate the studied face only when they were sure their response was correct and to respond DK otherwise (a free-report step). After that, they were asked to indicate which they thought was the studied face, even if it required guessing (a forced-report step). Finally, they were asked to rate their confidence in the forced-report response (see Hanczakowski et al., 2013, for this type of testing). The means of confidence judgments and the proportions of DK responses across conditions were our measures of metacognitive processes whereas the hit rate on forced-report recognition was our measure of memory retrieval processes (cf. Higham, 2002; Koriat & Goldsmith, 1996).

²The 2AFC test was used here rather than the standard old/new recognition test in order to circumvent the problems of distinguishing between context effects on bias and discrimination. In the present study we were interested in examining context effects on metacognitive processes both when context affects participants' ability to discriminate between targets and foils and also when it does not. Although numerous methods for deriving measures of discrimination on the old/new test have been proposed (some based on signal detection theory and some on multinomial processing tree models), there is currently no consensus on which of these measures appropriately distinguish between discrimination and bias to call any test probe 'old'. There is, however, a consensus that 2AFC tests are essentially bias-free (except for the left-right bias which is further discussed in footnotes 4, 5 and 6) and that any differences between conditions on such a test reflect differences in participants' ability to discriminate between targets and foils.

To reiterate our general predictions as applied to the procedure used for the present study, we predicted that familiarity of the context (novel vs. re-paired context conditions comparison) will not affect forced-report 2AFC recognition accuracy but will determine level of confidence and the rate of DK responses.³ By contrast, we predicted that context reinstatement (reinstated vs. re-paired context conditions) will lead to parallel effects on memory and metacognitive measures. Thus, whenever context reinstatement is potent to augment retrieval of item-context associations and lead to context-dependent discrimination reflected in increased forced-report 2AFC recognition accuracy, this should also be reflected in increased confidence and reduced rate of DK responses.

Experiment 1

Experiment 1 used the general procedure outlined earlier coupled with passive-viewing instructions at study. The passive-viewing instructions can be described as non-associative, as participants are not asked to attend to item-context associations. Under such instructions, previous studies have often failed to document context-dependent discrimination (see Hockley, 2008). Thus, in Experiment 1, under non-associative instructions we expected minimal effects of context reinstatement on retrieval of item-context associations and, accordingly, we predicted no differences between reinstated and re-paired context conditions in either memory or metacognitive measures. By contrast, we predicted that context familiarity would be attributed to candidate responses, leading to differences in metacognitive measures between re-paired and novel context conditions, while having no effect on the memory measure of hits in forced-report 2AFC recognition test.

³Note that our prediction of context familiarity increasing confidence differs from a common formulation of how confidence is shaped in 2AFC recognition, which assumes that confidence is a function of a magnitude of a difference in memory evidence for two test stimuli (cf. Clark, 1997). Context familiarity should not affect this difference.

Method

Participants. Fifty undergraduates from Cardiff University participated for course credit.

Materials and Design. A set of 96 black-and-white photographs of faces (a mixed set with male and female faces in equal proportions, all of them Caucasian) from the Psychological Image Collection at Stirling and another of 64 black-and-white photographs of landscapes from various Internet sources were assembled. The set of faces was divided into two subsets (A and B) of 48 faces (with approximately equal proportions of male and female faces in both subsets). One of these subsets served as study material and the other as a source of foils for a recognition test, which was counterbalanced across participants. Each face from the subset A was randomly yoked with a face from the subset B and with an individual context photograph of a landscape. The yoking of faces was done without regard to the sex of the yoked faces, which meant that pairs consisting of two male, two female, and one male and one female faces were created.

At study, faces were presented with their yoked context photographs. At test, both yoked faces were presented with a context photograph. If a studied face was assigned to the reinstated context condition, the context photograph presented at test was a photograph yoked with this pair of faces. If a studied face was assigned to the re-paired context condition, the context photograph presented at test was a photograph yoked with another pair of faces. If a studied face was assigned to the novel context condition, the context photograph presented at test was one of the 16 landscape photographs not yoked with any pair of faces. An equal number of studied faces were assigned to each of the three conditions (16) and this assignment was counterbalanced across participants.

The design had one independent variable manipulated within-participants: context condition (reinstated, re-paired, novel). Dependent variables of interest were: the rate of DK responses in free-report recognition, accuracy in forced-report recognition and mean confidence judgments in forced-report decisions.

Procedure. In the study phase participants were presented with 48 images of a face together with a landscape photo. The images shared a border and the face was always presented on the right. Each face-landscape pair was presented for 5 s and participants were instructed to look at both photographs. They were also forewarned about an unspecified memory test. The test directly followed the study phase. Each trial at test included three steps. First, two faces were presented on a screen with a landscape photo between them. Half of the studied faces were presented on the left and half were presented on the right. Participants were instructed to indicate which of the presented faces, the one on the left or the one on the right, was presented at study. They were informed that a context photo may help them make this decision. They were further instructed that accuracy was crucial in this step and thus an additional option to respond DK was made available and they should use it whenever they were not sure of their response. Second, immediately after the free-report decision, the same set of photographs was presented and participants were asked to indicate the face presented at study even if it required guessing. Third, they were asked to provide a confidence judgment concerning the correctness of their response in the preceding, forced-report step. The judgment was made by typing a value between 50 and 100 and participants were instructed that 50% corresponds to a chance level. The photographs were not presented on the screen during the confidence judgment stage of the test trial. All stages of the test were self-paced.

Results and discussion

The descriptive statistics are presented in Table 1. We start with the presentation of forced-report recognition accuracy, which is assumed to reflect memory processes unaffected by metacognitive effects (cf. Higham, 2002; Koriat & Goldsmith, 1996). A one-way ANOVA on accuracy of forced-report recognition failed to reveal a significant effect of the context condition, $F(2, 98) = 1.580$, $MSE = .022$, $p = .211$, $\eta_p^2 = .031$. No difference between novel and re-paired context conditions was expected as familiarity of the test context should not affect participants' ability to discriminate between targets and foils. Of main interest, however, is that accuracy in the reinstated context condition did not differ reliably from accuracy in the re-paired context condition. Any such difference would indicate context-dependent discrimination, which is better memory performance in the presence of the exact context that was paired with a given target at study. Such enhanced performance would suggest that reinstated contexts aided recollection of item-context associations. However, based on the previous findings (Hockley, 2008; Reder *et al.*, 2013), we expected recollection to play a negligible role in our design and this null effect seems at first blush consistent with our expectations. One could note that performance in the reinstated context condition was slightly better than in the re-paired context condition, although this was not reliable even with a one-tailed *t*-test, $t(49) = 1.378$, $SE = .029$, $p = .087$, $d = 0.196$.

Turning now to metacognitive measures, a one-way ANOVA on the mean of confidence judgments in forced-report recognition decisions revealed a significant effect of the context condition, $F(2, 98) = 28.032$, $MSE = 26.452$, $p < .001$, $\eta_p^2 = .364$. Pairwise comparisons showed that confidence was higher in the reinstated than in the re-paired context condition, $t(49) = 3.408$, $SE = 0.998$, $p = .001$, $d = 0.486$, and confidence in the re-paired context condition was in turn higher than confidence in the novel context condition, $t(49) = 4.701$, $SE = 0.912$, $p < .001$, $d = 0.668$.

A one-way ANOVA on the rate of DK responses in free-report recognition also revealed a significant effect of the context condition, $F(2, 98) = 16.122$, $MSE = .015$, $p < .001$, $\eta_p^2 = .248$. Pairwise comparisons showed that the rate of DK responses was higher in the re-paired than the reinstated context condition, $t(49) = 2.684$, $SE = .021$, $p = .01$, $d = 0.402$, and still higher in the novel versus re-paired context condition, $t(49) = 3.224$, $SE = .026$, $p = .002$, $d = 0.463$.⁴

The interpretation of the metacognitive results needs to be broken in two parts that are related to the issues context-dependent recognition and discrimination. Addressing first context-dependent recognition, we predicted that higher context familiarity in the re-paired than in the novel context condition would be attributed to a candidate response, inflating confidence and reducing DK responding compared to the novel context condition, in which context was unfamiliar. A comparison of the re-paired and novel context conditions supports our prediction. Thus, context familiarity, which does not allow participants to discern which of the test stimuli should be endorsed as previously studied, can induce in people more confidence that their candidate responses are indeed correct. In other words, the manipulation of context familiarity results in a confidence-accuracy dissociation by which context familiarity determines confidence in one's candidate responses even though it does not affect the accuracy of these responses.

⁴In the study phase of all experiments reported here, face photograph was presented to the right of the landscape photograph. In all tests, the target faces within each condition appeared equally often to the left and to the right of the context photograph. In this way, any differences between conditions cannot be explained by participants' bias towards the side on which faces were originally presented at study (i.e., right). To establish if participants were biased to choose faces presented on the right, we performed an additional analysis of forced-report accuracy with the side on which a target face was presented at test as a factor. If participants chose faces on the right more often, this should result in better accuracy when the target face was presented on the right. However, no such effect was observed as both the main effect of the side factor and the interaction with a context condition were not significant, $F_s < 1$. We also re-analysed the data for confidence and DK responding with the position of the target face as a factor but in these analyses the additional factor also failed to produce any reliable effects (lowest $p = .66$).

Turning now to the issue of context-dependent discrimination, a comparison of reinstated and re-paired context conditions produced reliable differences in metacognitive measures which remain inconsistent with our initial predictions. We predicted that metacognitive effects for this comparison will track the effects obtained in the memory measure of forced-report recognition accuracy, but despite equal accuracy between reinstated and re-paired context conditions, we still found reliably higher confidence and lower rate of DK responses for reinstated contexts. In other words, the pattern of results for reinstated and re-paired conditions revealed yet another confidence-accuracy dissociation. Because reinstated and re-paired context condition are equated in terms of context familiarity, the metacognitive differences between these conditions necessarily reflect the influence of the specific item-context association re-presented at test in the reinstated context condition but not in the re-paired context condition. Hence, the metacognitive effects of reinstating context seem to reflect augmented retrieval of item-context associations, which affects metacognitive measures but not the memory measure.

In formulating our initial predictions we assumed that the presence of context-dependent discrimination speaks directly to the issue of recollection of item-context associations: when recognition accuracy in the reinstated (vs. re-paired) context condition is enhanced, this shows that item-context associations are recollected, but when accuracy is equal between conditions, this shows that associations are not recollected. The results from metacognitive measures demonstrate, however, that our logic was wrong and that the lack of context-dependent discrimination in the recognition accuracy measure (the proportion of hits in forced-report 2AFC test in our case) cannot be taken as evidence that context reinstatement does not affect recollection of item-context associations. What we missed in our initial reasoning is that recognition accuracy measure may simply be insensitive to changes in memory processes. If recognition performance can be supported by memory information

other than recollection of item-context associations, for example by the relative familiarity of test stimuli, then augmenting recollection may fail to exert influence on recognition performance. At the same time, however, additional instances of recollection may induce more confidence in recognition choices. In this way, changes in recollection can be reflected in metacognitive measures, but at the same time be absent from memory measures.

Returning to our main topic of metacognitive effects, we have argued that clear effects of context familiarity on confidence and DK responses (re-paired vs. novel context condition) reflect attributions of context familiarity to chosen responses, not any changes in memory for the test stimuli, as evidenced by equated recognition accuracy between these conditions. However, it is quite clear now that this comparable accuracy cannot speak directly to whether novel and re-paired conditions were equated in recollection. Thus, providing familiar context at test could still potentially facilitate recollection of item-context associations in the re-paired condition, an effect reflected in higher confidence and fewer DK responses but not in forced-report accuracy. Such a mechanism would be obviously different from our favored mechanism of context familiarity attribution. Thus, since the exact mechanisms driving the metacognitive effects described in the present experiment are still uncertain, Experiment 2 was conducted to further assess the basis of increased confidence and reduced rate of DK responses due to a provision of familiar context at test.

Experiment 2

The purpose of the present experiment is to demonstrate that the effects caused by context familiarity and by context reinstatement can be dissociated. Documenting such a dissociation would support our claim that differences in metacognitive measures induced by changes in context familiarity and by context reinstatement are caused by two different

mechanisms: attribution of context familiarity to a candidate response and augmented recollection of item-context associations, respectively.

Our strategy in the present experiment was to further facilitate recollection of item-context associations by using specific associative instructions at study. Numerous studies have found that with associative study instructions the effects of context reinstatement are revealed in the measures of recognition accuracy (e.g., Hockley, 2008; Watkins et al., 1976), presumably because directing participants' attention to item-context associations enhances their encoding (Hockley & Cristi, 1996), at the same time increasing the probability of their subsequent retrieval. If our double-mechanism account of the results of Experiment 1 is correct, we would predict that stronger recollection would impact differently on a pattern of results caused by context familiarity and context reinstatement. In the case of context reinstatement (reinstated vs. re-paired context conditions), more recollection should result in better recognition accuracy in the reinstated context condition. In the case of context familiarity (re-paired vs. novel context conditions), more recollection should not lead to better recognition accuracy. At the same time, we predict that the pattern of results for metacognitive measures will replicate Experiment 1: recollection of item-context associations should create differences in confidence and DK responding between reinstated and re-paired context conditions, while attribution of context familiarity to a candidate response should still create differences in confidence and DK responding between re-paired and novel context conditions.

Method

Participants. Forty-six undergraduates from Cardiff University participated for course credit.

Materials, Design, and Procedure. All elements of the present experiment were the same as in Experiment 1, with one exception in the study phase. After the presentation of each pair of images, participants were asked to rate how well a presented face fits its context photograph on a scale from 1 to 6.

Results and discussion

The descriptive statistics are presented in Table 1. A one-way ANOVA on the accuracy of forced-report recognition yielded a significant effect of the context condition, $F(2, 90) = 3.639$, $MSE = .024$, $p = .03$, $\eta_p^2 = .075$. Pairwise comparisons showed that forced-report recognition accuracy was higher in the reinstated context condition than both in the novel context condition, $t(45) = 2.152$, $SE = .033$, $p = .037$, $d = 0.358$, and the re-paired context condition, $t(45) = 2.212$, $SE = .035$, $p = .032$, $d = 0.343$. There was no difference between the re-paired and novel context conditions, $t < 1$. These results confirm our prediction that aiding recollection of item-context associations by facilitating their creation at study would selectively impact on forced-report recognition accuracy in the reinstated context condition.

Turning now to the metacognitive measures, a one-way ANOVA on the mean of confidence judgments in forced-report recognition decisions revealed a significant effect of the context condition, $F(2, 90) = 49.185$, $MSE = 28.463$, $p < .001$, $\eta_p^2 = .522$. Pairwise comparisons showed that confidence judgments were higher in the reinstated context condition as compared both to the novel context condition, $t(45) = 8.28$, $SE = 1.24$, $p < .001$, $d = 1.233$, and to the re-paired context condition, $t(45) = 8.852$, $SE = 0.968$, $p < .001$, $d = 1.318$. The difference between the novel and re-paired context conditions failed to reach the level of statistical significance, $t(45) = 1.572$, $SE = 1.108$, $p = .123$, $d = 0.236$, although the judgments were numerically higher in the re-paired context condition.

A one-way ANOVA on the rate of DK responses revealed a significant effect of the context condition, $F(2, 90) = 24.408$, $MSE = .016$, $p < .001$, $\eta_p^2 = .352$. Pairwise comparisons showed that the rate of DK responses was higher in the re-paired versus reinstated context condition, $t(45) = 4.765$, $SE = .023$, $p < .001$, $d = 0.724$, and still higher in the novel versus re-paired context condition, $t(45) = 3.179$, $SE = .023$, $p = .003$, $d = 0.46$.⁵

The results for the measure of metacognitive control – DK responding – replicate the results of Experiment 1. Again, there are two effects here to discuss – a difference between reinstated and re-paired context conditions and a difference between re-paired and novel context conditions – which both need to be considered in light of the results of forced-report recognition accuracy. In Experiment 1 we could not conclude whether these two effects on metacognitive control stem from one or two mechanisms. The one-mechanism account postulated that our three context conditions differ in the strength of recollection which contributes to metacognitive processes. The double-mechanism account postulated that recollection determines the difference between reinstated and re-paired context conditions, while attribution of context familiarity to a candidate response determines the difference between re-paired and novel context conditions. The results of the present experiment settle this issue in favor of the double-mechanism account. In the present experiment, the difference in DK responding between reinstated and re-paired conditions occurred together with a parallel difference in forced-report recognition accuracy. This clearly points toward recollective underpinnings of this result. By contrast, the difference in DK responding

⁵ For Experiment 2, we repeated the analyses of whether the side on which the target face was presented at test affected forced-report recognition accuracy. Again, the side on which target face was presented did not affect accuracy in any way, as both the main effect of the side factor and the interaction with the context condition were not significant, $F_s < 1$. Additional analyses of confidence and DK responding with the side factor revealed that the side on which target was presented did affect confidence ($F(1, 45) = 6.47$, $p = .01$, for the main effect of the side, with higher confidence if the target face was presented on the right), but not DK responding. Other effects involving this factor were not significant (lowest $p = .34$). We offer no explanation for the effect of side of target presentation on confidence, which failed to replicate in Experiments 1 and 3.

between re-paired and novel context conditions occurred once again without a trace of a difference in forced-report recognition accuracy between these conditions. This indicates that recollection of item-context associations, successfully strengthened in this experiment, is not responsible for the discussed metacognitive effect. Instead, it should be assigned to the mechanism of attribution of context familiarity to a candidate response.

One unexpected feature of the present results is that a difference in the mean of confidence judgments between the re-paired and novel context conditions failed to reach the level of statistical significance. How should one interpret this apparent null result? One possibility is that participants were more confident about their responses in the re-paired context condition, but this was not picked up by the measure used in the confidence judgment task. We note here that just as recognition accuracy is only an imperfect measure of recognition processes, as the results of Experiment 1 clearly demonstrate, confidence judgments are only an imperfect measure of internal subjective confidence and thus insights from other measures are necessary to make inferences about these directly unobservable cognitive representations of confidence (cf. Hanczakowski, Zawadzka, Pasek, & Higham, 2013). The results concerning DK responding, an alternative measure of metacognitive processes included in our design, certainly give credence to the claim that the discussed null effect is spurious. Re-paired and novel context conditions clearly differed in the rate of DK responses, and this difference is difficult to explain unless these conditions differed also in participants' internal confidence assigned to candidate responses.

On the other hand, it is also possible that the encoding instructions introduced in Experiment 2 diminished the effects of context familiarity on metacognitive measures. This could occur because under conditions of strong recollection participants could have largely stopped relying on familiarity to choose a target face. If participants rely less on familiarity when recollection is more readily available (under associative study instructions), then

presumably the role of context familiarity in shaping metacognitive processes also becomes limited, and thus is reliably present only for DK responses but not confidence judgments. Experiment 3 was designed to unequivocally answer the question of whether the metacognitive effects of attributing context familiarity to candidate responses occur when participants rely on recollection in the final test.

Experiment 3

In the present experiment we wanted to create conditions in which correct responding in a recognition test would be highly dependent on recollection of item-context associations. The main question addressed under these conditions would be if people still use context familiarity as a basis for metacognitive processes. If such an effect was documented, it would suggest that the null effect of context familiarity on metacognitive monitoring in Experiment 2 was spurious. However, if again no effect of context familiarity emerged, consistent with Experiment 2 but contrasting with results of Experiment 1, it would suggest that the context familiarity effect on metacognitive monitoring is limited to conditions in which participants rely on familiarity as the basis of memory decisions.

A secondary objective of the present experiment was to investigate whether associative instructions at study, used in Experiment 2 but not in Experiment 1, are necessary for documenting the difference in the forced-report recognition accuracy between the reinstated and re-paired context conditions. To this end, we reverted to the non-associative instructions from Experiment 1, used in conjunction with our novel, recollection-driven test. In the discussion of Experiment 1 we have argued that recognition accuracy is not affected by context reinstatement under conditions of non-associative study instructions because recognition accuracy may not be sensitive to the effects concerning recollection. If this

reasoning is correct, then we would predict that in a recollection-driven test, accuracy would be affected by context reinstatement, even when non-associative study instructions are used.

To create a recollection-driven recognition test, a separate pre-study phase was added to the experimental procedure, in which all foils used subsequently in the recognition test were presented. Their presentation (without any context photographs) should more or less equate them on familiarity with faces presented in the actual study phase. In the test, participants were asked for each target-foil pair to endorse the face that was presented in the phase of the experiment in which faces were paired with context photographs. Because familiarity should be useless for this discrimination, the test required recollection of either temporal context of the experimental phase or recollection of a specific item-context association created at study (see Bodner & Lindsay, 2003, or Gruppuso, Lindsay, & Kelley, 1997, for similar procedures). We predicted that in such a recollection-driven test a difference in forced-report recognition accuracy between reinstated and re-paired context conditions would be obtained despite the use of non-associative instructions at study. The main question remained whether the metacognitive effects of context familiarity revealed in comparisons between the re-paired and novel context conditions (for confidence judgments in Experiment 1 and for DK responding in Experiments 1 and 2) would be preserved in this recollection-driven procedure.

Method

Participants. Thirty-six undergraduates from Cardiff University participated for course credit.

Materials, Design, and Procedure. Materials and design were exactly the same as in Experiments 1 and 2. In the present experiment a pre-study phase was added at the beginning of the experimental procedure. In this phase, participants were presented with

faces which were subsequently used as foils in a recognition test. Each face was presented for 2 s with passive-viewing instructions, without any reference to memory. The following study phase used the non-associative instructions from Experiment 1. The test phase was the same as in Experiment 1 and 2, except that participants were clearly instructed to indicate faces studied in the second phase of the experiment, the one in which faces were presented with landscape photos.

Results and discussion

The descriptive statistics are presented in Table 1. A one-way ANOVA on the accuracy of forced-report recognition yielded a significant effect of the context condition, $F(2, 70) = 8.82$, $MSE = .021$, $p < .001$, $\eta_p^2 = .201$. Pairwise comparisons showed that forced-choice recognition accuracy did not differ between the re-paired and novel context conditions, $t < 1$, but it was higher in the reinstated context condition than both in the novel context condition, $t(35) = 3.876$, $SE = .029$, $p < .001$, $d = 0.641$, and in the re-paired context condition, $t(35) = 3.533$, $SE = .038$, $p = .001$, $d = 0.563$. These results closely follow the results of Experiment 2 and once again demonstrate that recollection of item-context associations is more successful in the reinstated compared to re-paired and novel context conditions. Importantly, these effects occurred despite using non-associative instructions at study. This once again indicates that participants encode item-context associations at study, even when they are not explicitly directed to do so, and that these associations can be retrieved at test when cued with reinstated contexts. The fact that item-context associations are automatically established at study is perhaps not surprising given the earlier literature speaking to this issue (e.g., Bancroft, Hockley, & Farquhar, 2013; Jou, 2010), although it should be pointed out that recently Reder et al. (2013) have argued that such associations may not arise for unfamiliar faces – a hypothesis which is refuted by our results. However, what our results mainly underscore is that even when such associations are created and their

retrieval is facilitated by context reinstatement, such facilitation may not always be detected by standard recognition tests (see Experiment 1) but it is revealed when participants are forced to rely on recollection (see Macken, 2002, for similar arguments).

Turning now to the metacognitive measures, a one-way ANOVA on the mean of confidence judgments in forced-report recognition decisions revealed a significant effect of the context condition, $F(2, 70) = 20.319$, $MSE = 36.995$, $p < .001$, $\eta_p^2 = .367$. Pairwise comparisons showed that confidence judgments were higher in the reinstated context condition than in the re-paired context condition, $t(35) = 3.861$, $SE = 1.465$, $p < .001$, $d = 0.645$, and in turn higher in the re-paired than in the novel context condition, $t(35) = 3.437$, $SE = 0.986$, $p = .002$, $d = 0.574$.

A one-way ANOVA on the rate of DK responses revealed a significant effect of the context condition, $F(2, 70) = 15.166$, $MSE = .017$, $p < .001$, $\eta_p^2 = .302$. Pairwise comparisons showed that the rate of DK responses was higher in the re-paired versus reinstated context condition, $t(35) = 2.794$, $SE = .031$, $p = .008$, $d = 0.489$, and still higher in the novel versus re-paired context condition, $t(35) = 2.934$, $SE = .029$, $p = .006$, $d = 0.405$.⁶ The results for metacognitive measures exactly replicate the results of Experiments 1 but differ from the results of Experiment 2 inasmuch as the difference in mean confidence judgments between re-paired and novel conditions was again reliable. Recognition performance in the present

⁶ For Experiment 3, we repeated the analyses of whether the side on which the target face was presented at the test affected forced-report recognition accuracy. Interestingly, in this experiment a factor of side reliably interacted with the context condition, $F(2, 70) = 5.15$, $p = .01$. This interaction arose because participants' recognition performance in the re-paired context condition was markedly better when face was presented on the right ($M = .62$) rather than on the left ($M = .44$), whereas no such differences emerged for novel (.55 and .56, respectively) and reinstated (.69 and .64, respectively) context conditions. This difference for the re-paired context condition could suggest that merely presenting a target face on the same side of a familiar context photograph as at study aided participants' recollection. However, it remains unclear why such beneficial effect of the position reinstatement would be limited only to familiar contexts as the lack of a parallel difference for the novel context condition would indicate. The additional analyses of confidence and DK responding with ANOVAs including the factor of side failed to produce any significant effects involving this factor (lowest $p = .28$).

experiment necessitated the use of recollection and the results concerning the accuracy of forced-report recognition indicated that participants did rely on recollection. The fact that under these conditions differences in metacognitive measures (both confidence and DK responding) were reliably present between re-paired and novel context conditions indicates that context familiarity underlies metacognitive processes even in recollection-driven tests. In other words, the present results join the results for DK responding in Experiment 2 in suggesting that a null effect for confidence judgments observed in Experiment 2 was spurious.

General Discussion

In the present study we set out to assess the context effects on metacognitive processes. Our results can be summarized as follows. First, in the process of memory reporting in the 2AFC recognition test, context familiarity is attributed to a candidate response, inflating confidence judgments (Experiments 1 and 3) and reducing DK responding (Experiments 1, 2, and 3). This occurs without a change to forced-report recognition accuracy, and thus constitutes a clear example of a confidence-accuracy dissociation. Second, the effect of context familiarity on metacognitive processes is robust and occurs even when familiarity plays a negligible role as the basis of responding in a recollection-driven recognition task (Experiment 3). Third, metacognitive processes in the presence of reinstated contexts are also affected by recollection of item-context associations. When the exact study context is reinstated at test, it facilitates retrieval of item-context associations, further boosting confidence and reducing DK responding (Experiments 1, 2, and 3). Fourth, the effects of context reinstatement on recollection of item-context associations can be masked in the measure of recognition accuracy when the recognition task can be accomplished solely by the means of other memory information, like relative familiarity of targets and foils (Experiment 1). The contribution of recollection can be revealed either by strengthening

associations that could be recollected (Experiment 2) or by rendering performance in a recognition task reliant on recollection (Experiment 3).

Confidence-accuracy dissociations

Previous research on confidence judgments in a memory setting often touched upon the problem of a confidence-accuracy relationship. The question is to what extent confidence in memory choice tracks the accuracy of this choice. Numerous dissociations between confidence and accuracy have been documented, suggesting that confidence and accuracy partially depend on different processes. For example, in their seminal paper Busey, Tunnicliff, Loftus, and Loftus (2000) showed how low luminance of a tested face deflated retrospective confidence judgments for ‘old’ decisions, while simultaneously increasing memory performance when the studied face was also of low luminance. This result indicates that the higher the ease of processing of a face at test, presumably a function of luminance, the higher the confidence in ‘old’ decisions. Our results from the re-paired versus novel context conditions parallel these findings. Just as participants attributed the fluency of processing of a test face to fluency expected from the previous study episode for a given face in the experiments by Busey *et al.*, participants in our study attributed the feeling of familiarity evoked by context to a test face. Both types of attribution inflated retrospective confidence, producing a confidence-accuracy dissociation, as in our study recognition performance was independent of context familiarity, whereas in the study by Busey *et al.* recognition performance depended on study-test match in luminance, not on test luminance alone.

When discussing factors shaping metacognitive judgments such as retrospective confidence, it is important to distinguish between experience-based and theory-based processes (cf. Koriat, Bjork, Sheffer, & Bar, 2004). In short, a certain factor may affect

metacognition either because it changes people's subjective feelings about stimuli, or because people have a certain theory of how this factor affects their memory processes. Applied to the confidence-accuracy dissociation under discussion, we have argued here that the effects of familiar context on confidence stem from attribution of context familiarity to a candidate response, which would be an experience-based process. However, it is also possible that people have a particular theory according to which their memory responses are by and large more likely to be correct when made in a recognizable context. After all, the results of our experiments clearly show that context often does convey information that increases memory accuracy. Although we believe that the manipulation of familiarity is very close to the notion of experience-based processes, we do not wish to dismiss here the possibility that the effects of context familiarity may originate in participants' lay beliefs about context-dependency of memory. The methods for distinguishing between experience-based and theory-based processes in metacognition are currently being developed (see e.g., Mueller, Tauber, & Dunlosky, 2013) and these methods may be used in the future to address this issue.

Turning now to a comparison of reinstated and re-paired context conditions, it is worth evoking a recent study by Reinitz, Séguin, Peria, and Loftus (2012; see also Reinitz, Peria, Séguin, & Loftus, 2011). This study demonstrated that although recognition performance for faces depends both on familiarity of a face and recollection of distinctive features of this face, recollection affects confidence to a greater extent than it does affect recognition performance, again leading to a confidence-accuracy dissociation. This observation is paralleled by our second example of a confidence-accuracy dissociation, documented in Experiment 1. In this experiment, a comparison of the reinstated versus re-paired context conditions revealed that recollection strongly contributes to retrospective confidence, while leaving recognition performance mostly unaffected.

Together, our results concerning confidence confirm that recognition accuracy and retrospective confidence can often be based on at least partially distinct psychological processes, resulting in confidence-accuracy dissociations. We wish to stress, however, that by no means should this be taken to imply that confidence is always or even most commonly dissociated from accuracy. Our results from Experiments 2 and 3 show how recollection of item-context associations can at the same time shape confidence and recognition accuracy. This, contrasted with the results of Experiment 1, shows how the nature of confidence-accuracy relationship is affected by the conditions of both encoding and testing.

Response withholding in memory tests

In our study we assessed not only metacognitive monitoring in the form of confidence judgments but also metacognitive control in the form of DK responding. The results for DK responding largely mirrored the effects observed with retrospective confidence judgments, with lower rates of DK responses when confidence judgments were higher and higher rates of DK responses when confidence judgments were lower. This pattern is consistent with the framework developed by Koriat and Goldsmith (1996), in which metacognitive control in the form of DK responding is dependent on the products of metacognitive monitoring, reflected in confidence judgments.

Critically, the framework of Koriat and Goldsmith (1996) describes how metacognitive control in the form of DK responding is related to memory performance. When participants are allowed to withhold responses in a free-report test, the ultimate number of correct responses reported (what Koriat and Goldsmith refer to as an input-bound measure of memory accuracy) depends not only the efficacy of memory processes but also on participants' willingness to volunteer the products of retrieval. Thus, differences in metacognitive processes between conditions can contribute to effects obtained with the

measures of the proportion of correct responses volunteered, sometimes even when actual memory processes are equated between these conditions. To illustrate this issue, we will now use the combined data from Experiments 1-3 to present additional analyses of novel and re-paired context conditions.

First, we focus on forced-report recognition accuracy, which provides the measure of memory processes that is uncontaminated by the metacognitive process of response withholding. A 2 (context condition: re-paired vs. novel) x 3 (Experiment: 1, 2, 3) mixed ANOVA on forced-report recognition accuracy yielded a significant effect of Experiment, $F(1, 129) = 24.75$, $MSE = .037$, $p < .001$, $\eta_p^2 = .277$, which arose because recognition accuracy was best in Experiment 2 with associative instructions ($M = .75$, $SD = .15$, collapsed across context conditions), somewhat worse in Experiment 1 with non-associative instructions ($M = .67$, $SD = .14$), and much worse in Experiment 3 in a test that required contextual discrimination between study and pre-study phases ($M = .54$, $SD = .12$). More importantly, a main effect of context condition and the interaction were both not significant, $F_s < 1$, once more testifying to the fact that presenting at test context which is merely familiar but is not reinstated does not help recognition memory beyond presenting a novel context.

Second, we focus on the measure of the proportion of correct answers volunteered out of all questions asked in the free-report recognition test. A 2 (context condition: re-paired vs. novel) x 3 (Experiment: 1, 2, 3) mixed ANOVA yielded again a significant main effect of Experiment, $F(1, 129) = 10.526$, $MSE = .083$, $p < .001$, $\eta_p^2 = .14$, paralleling the results of the forced-report recognition accuracy. The performance was again best in Experiment 2 ($M = .48$, $SD = .20$, collapsed across context conditions), slightly worse in Experiment 1 ($M = .42$, $SD = .22$) and worst in Experiment 3 ($M = .28$, $SD = .19$). Crucially, for the measure of free-report recognition a main effect of context condition was also significant, $F(1, 129) = 12.736$, $MSE = .017$, $p = .001$, $\eta_p^2 = .09$, as participants volunteered more correct responses in the re-

paired context condition ($M = .43$, $SD = .23$, collapsed across experiments) than in the novel context condition ($M = .37$, $SD = .24$). This difference should not be taken to imply that memory was somehow better in the re-paired context. In the 2AFC recognition test any change in the rate of DK responses necessarily results in the change in the proportion of correct responses volunteered because participants can get the correct answer half of the time simply by guessing. In summary, although forced-report recognition data clearly points to equal recognition accuracy between re-paired and novel context conditions, the analysis of the proportion of correct responses in free-report recognition shows better performance in the re-paired context condition. These results clearly demonstrate that in a test in which response withholding is possible, the analysis of the proportion of correct responses can produce a reliable difference between conditions which are actually perfectly equated in terms of memory processes.

The analyses presented above may well have important consequences for our understanding of context effects in memory. At the beginning of the present paper we mentioned that the effects of context reinstatement seem to be more readily observed in recall than in recognition procedures. It is important to note, however, that recall procedures almost invariably allow for response withholding. When comparing reinstated and novel context conditions in recall procedures in terms of the proportion of recalled items, one is unable to separate memory and metacognitive processes. More items volunteered in the reinstated context condition as compared to the novel context condition may reflect better memory access due to context reinstatement, but may also at least partially reflect a greater propensity to volunteer retrieved answers when reinstated and thus familiar context is presented at test. Of course this metacognitive origin of a difference in proportions of recalled (and actually volunteered) answers is possible only if participants in recall procedures are prone to withholding some of the correct answers. It is worth pointing out, however, that previous

comparisons of free- and forced-report recall revealed that at least under some conditions people are quite likely to withhold answers that are correct (Higham, 2002; Higham & Tam, 2005, 2006). If this occurs in studies examining context effects in recall, then the contribution of metacognitive processes to what have been described as purely memory effects of context reinstatement cannot be dismissed. This issue awaits further research with recall procedures.

Context effects in memory

Although our study was mainly designed to investigate context effects on metacognitive processes, it also provides insights into basic memory processes. This is so because metacognition in memory tasks builds on memory information. However, as we discussed earlier, this memory information may not be necessarily of the same type as information underlying memory performance. We see this occasional discrepancy between memory and metacognition as quite beneficial for memory researchers because it provides a window of opportunity for examining memory processes that are difficult to reveal by standard memory tests (see also Jaeger, Cox, & Dobbins, 2012, for a similar approach).

The literature on context effects in recognition is mostly focused on the discussion of context-dependent discrimination. Researchers focused on recognition tests have argued whether reinstating the exact study context at test facilitates discrimination of targets and foils. The results in this respect are not entirely consistent, with some studies revealing context-dependent discrimination (e.g., Macken, 2002) and some failing to reveal it (e.g., Hockley, 2008). This discrepancy is also visible in our own study. Context-dependent discrimination was absent in Experiment 1, where forced-report 2AFC recognition was equal in all three context conditions, but was reliably present in Experiments 2 and 3, where forced-report 2AFC recognition was clearly best for the reinstated context condition. These results, coupled with the pattern already observed in the literature, demonstrate that the question

concerning context-dependent discrimination should not be *whether* it occurs but *when* it occurs. Macken suggested that such effects are most likely to occur when recognition responding is based on recollection of specific associations, as revealed in ‘remember’ responses. Our own results support this conclusion as context-dependent discrimination occurred when the role of recollection was promoted by specific study instructions (Experiment 2), presumably increasing the proportion of recollection-based responses, and when the nature of the recognition test was changed to directly induce more reliance on recollection (Experiment 3).

Perhaps more importantly, our results demonstrate how the absence of context-dependent discrimination should not be taken to imply the absence of recollection of item-context associations. Based on the previous results and our incorrect logic, we initially assumed that the lack of effects of context reinstatement in recognition accuracy can be straightforwardly used to infer the nature of recognition processes. Our results reveal the precarious nature of such reasoning by demonstrating how recognition accuracy can be insensitive to memory processes occurring in a recognition test. This observation underscores the necessity of distinguishing between a cognitive process and the measure used to tap into this process. Although it may sound quite counterintuitive, it appears that recognition tests are not necessarily sensitive to processes involved in recognition. To the extent to which cognitive researchers are interested in cognitive processes, and not merely in their reflections in particular laboratory tasks, we argue that they should strive to use a variety of measures to assess cognitive processes. We believe that the examination of metacognitive measures may prove of further use in this endeavour.

Context outside laboratory

We will devote the final words of the present paper to possible applied consequences of our results. The issue of context reinstatement plays a prominent role in research on eyewitness memory. Indeed, the context reinstatement manipulation constitutes a vital part of the widely used and investigated protocol of eliciting eyewitness testimony – the Cognitive Interview (Fisher & Geiselman, 1992). Our results, although obtained in a quite simplified laboratory procedure, may thus convey important information for this type of applied research. What our results contribute is a word of caution when considering the effects of context reinstatement techniques. It is certainly true that context reinstatement benefits memory, as demonstrated in numerous studies examining the effectiveness of different techniques included in the Cognitive Interview (e.g., Dando, Wilcock, & Milne, 2009; Milne & Bull, 2002), as well as in our laboratory procedure. However, an issue that, at least to our knowledge, did not catch researchers' attention is that familiar context may inflate people's confidence concerning their memory reports even when it has no power of augmenting memory. This is particularly problematic in the eyewitness context as confidence expressed by an eyewitness is a major clue that people use to assess the credibility of the eyewitness (e.g., Sporer, Penrod, Read, & Cutler, 1995). Reinstating context can thus have an effect of increasing the perceived credibility of an eyewitness quite independently of any possible beneficial effect on memory.

The present experiment involved faces as to-be-remembered stimuli. This kind of stimuli is important for the eyewitness literature, which is often concerned with the processes involved in identifying culprits from line-up procedures. Interestingly, there are some indications that line-up identifications can be related to the issue of context-dependent memory. For example, Evans, Marcon, & Meissner (2009) showed that the accuracy of line-up identifications can be improved if context reinstatement techniques are implemented. The present study chimes with these conclusions by showing that contextual associations are

created for faces even without voluntary encoding, which are the conditions of most relevance to eyewitness scenarios, and that retrieval of these associations can be facilitated by context reinstatement. However, our results can be used to pose yet another question for the eyewitness identification research. Specifically, if context familiarity inflates participants' confidence in candidate responses and makes them more likely to volunteer these responses, then the question arises whether context familiarity would also make people more willing to choose from a line-up. In other words, would context reinstatement techniques affect people's bias to identify their 'best choice' in the line-up? Despite the fact that our laboratory procedure differs in number of ways from the line-up studies (e.g., number of faces used, the nature of a memory choice), this seems to be an important avenue for follow-up investigations, given the importance of identification bias in real-life scenarios (cf. Mickes, Flowe, & Wixted, 2012).

References

- Bodner, G. E., & Lindsay, D. S. (2003). Remembering and knowing in context. *Journal of Memory and Language, 48*, 563-580.
- Bancroft, T. D., Hockley, W. E., & Farquhar, R. (2013). The longer we have to forget the more we remember: The ironic effect of postcue duration in item-based directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 691-699.
- Bulevich, J. B., & Thomas, A. K. (2012). Retrieval effort improves memory and metamemory in the face of misinformation. *Journal of Memory and Language, 67*, 45-58.
- Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review, 7*, 26-48.
- Clark, S. E. (1997). A familiarity-based account of confidence-accuracy inversions in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 232-238.
- Dando, C., Wilcock, R., & Milne, R. (2009). The Cognitive Interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology, 23*, 138-147.

- Dodson, C. S., & Shimamura, A. P. (2000). Differential effects of cue dependency on item and source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1023-1044.
- Eich, E. (1985). Context, memory and integrated item/context imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 764-770.
- Eich, E., & Metcalfe, J. (1989). Mood dependent memory for internal versus external events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 443-455.
- Evans, J. R., Marcon, J. L., & Meissner, C. A. (2009). Cross-racial lineup identification: Assessing the potential benefits of context reinstatement. *Psychology, Crime & Law*, *15*, 19-28.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.
- Geiselman, R. E., & Bjork, R. A. (1980). Primary versus secondary rehearsal in imagined voices: Different effects on recognition. *Cognitive Psychology*, *12*, 188-205.
- 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.
- Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 259-278.
- Gruppuso, V., Lindsay, D. S., & Masson, M. E. J. (2007). I'd know that face anywhere! *Psychonomic Bulletin & Review*, *14*, 1085-1089.

- 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.
- Hanczakowski, M., Zawadzka, K., Pasek, T., & Higham, P. A. (2013). Calibration of metacognitive judgments: Insights from the underconfidence-with-practice effect. *Journal of Memory and Language, 69*, 429-444.
- Higham, P. A. (2002). Strong cues are not necessarily weak: Thomson and Tulving (1970) and the encoding specificity principle revisited. *Memory & Cognition, 30*, 67-80.
- Higham, P. A., & Tam, H. (2005). Generation failure: Estimating metacognition in cued recall. *Journal of Memory and Language, 52*, 595-617.
- Higham, P. A., & Tam, H. (2006). Release from generation failure: The role of study list structure. *Memory & Cognition, 34*, 148-157.
- 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.
- Hockley, W. E., Bancroft, T. D., & Bryant, E. (2012). Associative and familiarity-based effects of environmental context on memory. *Canadian Journal of Experimental Psychology, 66*, 81-89.
- Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition, 24*, 202-216.

- Jaeger, A., Cox, J. C., & Dobbins, I. G. (2012). Recognition confidence under violated and confirmed memory expectations. *Journal of Experimental Psychology: General, 141*, 282-301.
- Jou, J. (2010). Can associative information be strategically separated from item information in word-pair recognition? *Psychonomic Bulletin & Review, 17*, 778-783.
- Kerr, N. H., & Winograd, E. (1982). Effects of contextual elaboration on face recognition. *Memory & Cognition, 10*, 603-609.
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609-639.
- Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133*, 643-656.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517.
- 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.
- Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18*, 361-376.

- Milne, R., & Bull, R. (2002). Back to basics: A componential analysis of the original Cognitive Interview mnemonics with three age groups. *Applied Cognitive Psychology, 16*, 743-753.
- Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review, 20*, 378-384.
- 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.
- Murnane, K., & Phelps, M. P. (1994). When does a different environmental context make a difference in recognition? A global activation model. *Memory & Cognition, 22*, 584-590.
- Murnane, K., & Phelps, M. P. (1995). Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 158-172.
- Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128*, 403-415.
- Reder, L. M., Victoria, L. W., Manelis, A., Oates, J. M., Dutcher, J. M., Bates, J. T., Cook, S., Aizenstein, H. J., Quinlan, J., & Gyulai, F. (2013). Why it's easier to remember seeing a face we already know than one we don't: Preexisting memory representations facilitate memory formation. *Psychological Science, 24*, 363-372.
- Reinitz, M. T., Peria, W. J., Séguin, J. A., & Loftus, G. R. (2011). Different confidence-accuracy relationships for feature-based and familiarity-based memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 507-515.

- Reinitz, M. T., Séguin, J. A., Peria, W., & Loftus, G. R. (2012). Confidence-accuracy relations for faces and scenes: Roles of features and familiarity. *Psychonomic Bulletin & Review, 19*, 1085-1093.
- Smith, S. M. (1985). Environmental context and recognition memory reconsidered. *Bulletin of the Psychonomic Society, 23*, 173-176.
- Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*, 342-353.
- Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*, 203-220.
- Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*, 315-327.
- Tibon, R., Vakil, E., Goldstein, A., & Levy, D. A. (2012). Unitization and temporality in associative memory: Evidence from modulation of context effects. *Journal of Memory and Language, 67*, 93-105.
- Watkins, M. J., Ho, E., & Tulving, E. (1976). Context effects in recognition memory for faces. *Journal of Verbal Learning and Verbal Behavior, 15*, 505-517.
- Winograd, E., & Rivers-Bulkeley, N. T. (1977). Effects of changing context on remembering faces. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 397-405.

Table 1. Recognition accuracy on a forced-report test, mean of retrospective confidence judgments for forced-report recognition decisions and the rate of ‘don’t know’ (DK) responses on a free-report test as a function of the context condition in Experiments 1, 2, and 3. Standard errors of the means are given in parentheses.

	Context condition		
	Novel	Repaired	Reinstated
Experiment 1			
Forced-report accuracy	.66 (.02)	.67 (.02)	.71 (.02)
Confidence	66.68 (1.37)	70.97 (1.25)	74.37 (1.30)
DK responding	.46 (.04)	.38 (.03)	.32 (.03)
Experiment 2			
Forced-report accuracy	.75 (.02)	.75 (.03)	.83 (.02)
Confidence	68.44 (1.63)	70.18 (1.44)	78.75 (1.45)
DK responding	.43 (.04)	.36 (.03)	.25 (.03)
Experiment 3			
Forced-report accuracy	.55 (.03)	.53 (.03)	.66 (.02)
Confidence	61.87 (1.75)	65.26 (1.62)	70.91 (1.64)
DK responding	.57 (.05)	.49 (.04)	.40 (.04)