**Peer Review and Communication History**

**MS Title**: Unfamiliar Contexts Compared to Familiar Contexts Impair Learning in Humans

**Author Names**: Alizadeh Asfestani M, Nagel J, Beer, S, Nikpourian, G, Born J, Feld GB

**Submitted:** Jun 21, 2022

**Editor First Decision**: Revise & Resubmit

Sep 23, 2022

Dear dr. Tkotz,

First, let me apologize for the long delay between your submission of your manuscript and this decision. Only one month after your submission, I was assigned as handling editor of your manuscript. Furthermore, it was hard to find reviewers perhaps due to the summer holidays. However, I have now received one expert review of your manuscript, “Unfamiliar Contexts Compared to Familiar Contexts Impair Learning in Humans”. I also independently read and reviewed your manuscript. Considering that the entire review process has already cost more than 3 months, I decided to make a decision based on the reviewer’s comment and my own evaluation. Both the reviewer and I agree that your manuscript addresses an interesting question. At the same, time we both came to the conclusion that the manuscript needs major revisions to warrant publication in Collabra. My decision is, therefore, to reject the current version of the manuscript and to invite you to submit a revised version in which you address the reviewer’s and my feedback.

You will find the reviewer’s feedback below. In think the reviewer has clearly formulated some excellent points and I will not reiterate them here. My own points are presented subsequently.

A major issue I have with your manuscript is that the main hypothesis is not sufficiently grounded in the memory literature. Some of the literature you mention in the Introduction seems not entirely relevant for your study. For example, the study of Smith and Handy (2014) is about repetitions, whereas your study is not. More importantly, however, is that the hypothesis is not linked to general theories of memory, like SAM, REM, MINERVA or others. It would be good to take these theories (or one of them) and use them as a basis for your main hypothesis. Also, I would urge you to take the type of context manipulation (i.e., multiple physical contexts), the type of processing at input (i.e., with these paired associates participants might engage in associative processing) and the memory task (i.e., cued recall) (see Smith and Vela (2001)) in your experimental set-up when formulating the hypothesis. For example, if participants are likely to engage in associative processing, the effects of context at the cued-recall test might be small. In any case, after reading the current version of the Introduction, I found your hypothesis not particularly compelling.

A second major issue is that your results might simply be due to a processing/encoding deficit; you provide this explanation in the introduction of the manuscript. Simply put, processing the new context might take processing/encoding efforts away from the word pairs. Such a processing/encoding account could perfectly explain the results in both of your experiments. However, there are no direct measurements of the processing of the word pairs. Such measurements would be needed to make a compelling case for the deficit processing/encoding account. One option to remedy this problem is to do a replication of experiment 1 including direct measurements of processing/encoding. A second option would be to address the processing/encoding account in the discussion, acknowledge that crucial measurements are lacking for compelling evidence and argue why an explanation according to an attention account is theoretically interesting. Personally, I would favor the first option because this would lead to a test of the processing/encoding account and would provide a replication of experiment 1.

A third major issue relates to the reference to metacognition in the introduction. I is not clear to me exactly what kind of metacognitive knowledge or skills, participants would employ in the current tasks and how this relates to the hypothesis. I propose to drop this part from the Introduction and focus on a more extensive grounding in the memory literature (see my first point) instead.

A fourth major issue is that the results are not entirely correctly reflected in the discussion. In the discussion, the focus is on the results of experiment 1, which shows a context effect. I think it is important to discuss the difference in outcomes between the experiments and to provide an explanation for that (see my second point). Furthermore, the part on the neuropsychological mechanisms underlying effects of contextual change on learning new information is highly speculative. I would suggest to drop this from the paper.

A minor issue is that you foreshadow the results of experiment 1 in the introduction. This is strange because the details and the results of experiment 1 are presented after the introduction. Hence, I would propose to formulate the hypothesis, present the method and results of experiment 1, discuss the results and provide a rationale for experiment 1.

In sum, I invite you to submit a revision of your manuscript. Please include a document with a point-by-point response to both the points I list here and the reviewer’s comments, outlining each change made in your manuscript or providing a suitable rebuttal.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This may be the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at editorialoffice@collabra.org.

We hope you can submit your revision within the next 3 months. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

Peter Verkoeijen

# Reviewer 1

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

In this manuscript, the authors found that learning occurred in the same context benefit learning and change of context impair new learning in Experiment 1. These differences were diminished when participants were familiarized with the other context in Experiment 2. This finding is interesting and added findings to the understanding of the effect of learning occurred in novel context. Figures of illustration of procedures are clear and help understanding. Figures of reported results are very clear and informative. However, this manuscript would be benefited from revising the introduction and discussion session. Below are some suggestions:

1-In the introduction, the authors should have more clear predictions for changes across two learning sessions and the effect of same or different context and also potential interaction between them. It is important to clarify whether the comparison is between the same and different context condition or between two learning episodes.

2-On page 4, it is unclear what “additional changes to the context” means.

3-On page 5, hypothesis for experiment 2 is missing.

4-On page 14, participants in the same context group had better learning performance in the evening session compared to the morning session in Experiment 1. But this effect was not replicated in Experiment 2. The fact that this effect was found in Experiment 1 only should be unpacked and discussed.

5-On page 21, after summarizing findings in two experiments, it will be beneficial for readers to reinstate your hypothesis and elaborate on your findings further (e.g., the effect of context familiarization) before discussing the differences with other papers.

6-On page 23, it is unclear why pattern separation process will impair learning in new context. Why would not pattern separation enhance learning in a novel context by making the context distinct and reducing interference?

7-It might be interesting to look at learning performance across three blocks to see the acquisition of word pairs, e.g., different learning pattern between morning and evening session.

8-On page 4, one citation seems problematic. The paper by Cox and Fell’s (2020) is a review paper on methodological issues in human sleep EEG analysis and did not report empirical studies.

**Author Response**
Nov 10, 2022

Note: Based on another reviewer’s comment, we re-ordered the manuscript: It now starts with the introduction for experiment 1, followed by the methods, results and discussion for experiment 1. Afterwards, we provide the rationale for experiment 2, which is followed by the methods and results for experiment 2, and a general discussion. This means that some paragraphs have moved to different pages.

Reviewer comments are highlighted in bold, and changes to the manuscript are highlighted in red.

**Reviewer 1**

**1-In the introduction, the authors should have more clear predictions for changes across two learning sessions and the effect of same or different context and also potential interaction between them. It is important to clarify whether the comparison is between the same and different context condition or between two learning episodes.**

We predicted that the two experimental groups show the same performance in the first session, but differ in their performance in the second session, after the experimental manipulation has been introduced. Accordingly, the main comparison of interest is between the two conditions, which we have now clarified in the introduction (p. 5). A secondary prediction was that the same condition group would show improved performance when comparing sessions one and two:

“We predicted that learning in the same context would improve learning performance, i.e., that participants whose learning sessions took place in the same context would perform better in the second learning session than participants who spent their learning sessions in two different contexts. We also predicted that participants that learned twice in the same condition would improve their performance from the first to the second condition, whereas the performance would remain the same for participants who switched contexts.”

**2-On page 4, it is unclear what “additional changes to the context” means.**

We have provided examples of how contexts might differ in addition to a temporal separation:

“When the learning episodes take place shortly after each other, other contextual features might be needed to separate the episodes. When the learning episodes already differ in their temporal context, additional changes to the context (i.e., not only a different time, but also a different room, decoration etc.) might not provide any added value, but simply additional information that needs to be processed.”

**3-On page 5, hypothesis for experiment 2 is missing.**

Since we have split the manuscript into different parts for the two experiments, the rationale for experiment 2 is now more elaborate (p. 16). The last sentence elaborates on which pattern of results we would expect if familiarity was the explanation for the results observed in experiment 1:

“However, the benefit of matching contexts for new learning might be explained by the fact that participants in the same context condition were simply more familiar with the context by the second learning session. For example, when new learning occurs in the same context twice less attentional resources may be needed the second time and therefore explain why memory was improved for the same context condition. In this case, familiarisation to the context prior to the second learning should suffice to produce the context related learning benefit. Therefore, we decided to control for effects of familiarity, when manipulating the learning context. To do so, we conducted a second experiment that mirrored the first experiment, with the addition that participants were familiarised with the other context before the second learning session in the evening. Crucially, participants did not learn anything when they were exposed to the other context after the first learning session in the morning but performed a letter counting task instead that did not involve memory processes. This means that in the evening learning session, all participants were familiar with both contexts, but had only learned word-pairs in one of them. If familiarity is not a relevant moderator of the results observed in experiment 1, we would expect that in experiment 2, new learning in a different, but familiar context should be impaired compared to new learning in the same context.”

**4-On page 14, participants in the same context group had better learning performance in the evening session compared to the morning session in Experiment 1. But this effect was not replicated in Experiment 2. The fact that this effect was found in Experiment 1 only should be unpacked and discussed.**

We agree with the reviewer that this pattern of results is interesting and have added the necessary analysis to compare the difference between morning and evening session across the two experiments (a Session x Experiment ANOVA for the same context condition) on page 19. Of note, there was no Session by Experiment interaction, which means we cannot make strong inferences from this different pattern:

“Even though participants in the same context condition in experiment 1 showed significantly better performance in the evening than in the morning session, and no such effect was seen in experiment 2, there was no Experiment x Session interaction in an ANOVA run on the same context condition (F(1, 38) = 1.78, p = .190, ηp² = 0.045).”

We nevertheless elaborate on this result in the discussion (p. 23):

“While participants in the same context condition of experiment 1 even showed significantly improved performance in the second learning session, no such effect was observed in experiment 2. However, statistically a direct comparison between the same context condition across the two experiments shows that participants in experiment 1 do not show a significantly stronger improvement across sessions than participants in experiment 2. This means we can neither conclude that the familiarisation procedure disturbed the effect of learning in the same context, nor that it did not. However, it suggests that the effect of the familiarisation procedure on the context change condition was stronger. Nonetheless, it would be interesting to follow up this research to investigate whether the context familiarisation procedure interferes with the context consolidation in some way.”

**5-On page 21, after summarizing findings in two experiments, it will be beneficial for readers to reinstate your hypothesis and elaborate on your findings further (e.g., the effect of context familiarization) before discussing the differences with other papers.**

We have elaborated the first paragraphs of the general discussion (p. 22):

“We investigated the effect of context on learning in two experiments. In the first, participants learned word pairs in the morning and then learned a new set of word pairs in either the same or a different context in the evening. We found that participants performed better when context was kept constant across learning sessions. While participants in the same context condition improved across learning sessions, participants in the different context condition learned less in the second session. In a second experiment, participants were additionally familiarized to the other context before the second learning session. This was done by letting participants complete a task without learning in the other context before or after learning in the morning. The aim of experiment 2 was to test whether different contexts have the same impact on new learning when they are familiar. The familiarization with the other context wiped out the differential effects found for the two context conditions in experiment 1. That is, no significant differences between the same and the different context condition emerged in experiment 2.

When comparing the two experiments the results supported a role of the familiarisation process to reduce the detrimental effect of changing contexts. For the first run of the evening learning session in the same context condition, performance in experiment 1 and experiment 2 does not differ. However, in the different context condition, learning performance during the first run of the evening learning session in experiment 1 was significantly lower than in experiment 2. While participants in the same context condition of experiment 1 even showed significantly improved performance in the second learning session, no such effect was observed in experiment 2. However, statistically a direct comparison between the same context condition across the two experiments shows that participants in experiment 1 do not show a significantly stronger improvement across sessions than participants in experiment 2. This means we can neither conclude that the familiarisation procedure disturbed the effect of learning in the same context, nor that it did not. However, it suggests that the effect of the familiarisation procedure on the context change condition was stronger. Nonetheless, it would be interesting to follow up this research to investigate whether the context familiarisation procedure interferes with the context consolidation in some way.”

**6-On page 23, it is unclear why pattern separation process will impair learning in new context. Why would not pattern separation enhance learning in a novel context by making the context distinct and reducing interference?**

Since the word lists in the two sessions do not overlap (A-B in session 1, C-D in session 2), reducing interference is not a major issue for this task – instead, the cued recall procedure likely benefits from pattern completion processes. We have elaborated on this in the respective paragraph (p. 26):

“Cholinergic models of hippocampal function (Easton et al., 2012; Hasselmo et al., 1996; Meeter et al., 2004) suggest that higher levels of cholinergic activity, following exposure to novel contextual cues, may bias the hippocampus toward forming distinctive memories by prioritizing encoding over retrieval (Douchamps et al., 2013). Whereas lower cholinergic levels occurring in familiar contexts (Giovannini et al., 2001), may bias the hippocampus toward memory reactivation aiding pattern completion processes (Duncan et al., 2012). In our cued recall word-pair task, pattern completion processes were likely beneficial, because participants had to generate the matching target to the cue presented at test. Furthermore, the word lists across the two sessions did not overlap (A-B, C-D), so that reducing interference via pattern separation was less important. This explains why new learning in a familiar context is superior to new learning in an unfamiliar context. Future experiments should resolve, whether in an A-B, A-C paradigm these effects reverse. However, participants in the different context condition of experiment 1 decreased their performance across sessions, in both of which the context was equally unfamiliar. A possible explanation is that pattern completion and separation processes do not only affect the learned material itself, but also metacognitive knowledge about the learning task itself (such as strategies). In the evening session of the unfamiliar context condition of experiment 1, pattern separation processes might have prevented access to neuronal processing steps relevant for task execution that were associated to the morning context. In the morning, when the context was also unfamiliar, but no pattern separation processes were initiated because the task was entirely novel as well, this access would not have been blocked, leading to the performance measured in the morning session.”

**7-It might be interesting to look at learning performance across three blocks to see the acquisition of word pairs, e.g., different learning pattern between morning and evening session.**

We are unsure, if we understand this point correctly. Does the reviewer mean the three runs of learning? We already report the analysis across the three runs per session. We did not analyse the three blocks per run individually, since retrieval was only measure after each run completed (i.e., after all three blocks had been seen) and thus no higher frequency data are available.

**Experiment 1 (p. 11):**

“When extending the analysis to all three runs in each session (Figure 2 C), an ANOVA likewise showed that participants learned significantly more word-pairs in the same context in the evening learning session (Session × Context interaction: F(1,38) = 6.38, p = .016, ηp² = .144). In the same analysis we found that word pair learning improved across the three runs (F(2,76) = 857.05, p ≤ .001, ηp² = .958). All other factors did not significantly affect performance (all F ≤ 2.77, p ≥ .104). An analysis of the separate sessions revealed that the benefit of the same context was only evident in the evening learning session (F(1, 38) = 4.43, p = .042, ηp² = 0.104), but not in the morning learning session (F(1, 38) = 1.04, p = .313, ηp² = 0.027).”

**Experiment 2 (p. 19):**

“As in experiment 1, an ANOVA across all runs (Figure 2D) revealed that participants improved their learning performance across the three runs (F(1.47,55.80) = 676.10, p ≤ .001, ηp² = 0.947). However, in contrast to experiment 1, there was no significant difference between the two contexts (Context main effect: F(1, 38) = 0.10, p = .759, ηp² = 0.003; Context x Session interaction effect: F(1, 38) = 0.09, p = .769, ηp² = 0.002).”

**8-On page 4, one citation seems problematic. The paper by Cox and Fell’s (2020) is a review paper on methodological issues in human sleep EEG analysis and did not report empirical studies.**

Thank you for catching that, we accidentally cited the wrong paper and have corrected the reference to the correct paper:

Cox, W. R., Dobbelaar, S., Meeter, M., Kindt, M., & Ast, V. A. van. (2021). Episodic memory enhancement versus impairment is determined by contextual similarity across events. Proceedings of the National Academy of Sciences, 118(48). <https://doi.org/10.1073/pnas.2101509118>

**Reviewer 2**

**Some of the literature you mention in the Introduction seems not entirely relevant for your study. For example, the study of Smith and Handy (2014) is about repetitions, whereas your study is not.**

We deleted the part in the introduction about Smith & handy (2014).

**[T]he hypothesis is not linked to general theories of memory, like SAM, REM, MINERVA or others. It would be good to take these theories (or one of them) and use them as a basis for your main hypothesis.**

When we designed the experiment and formulated our hypothesis, we did not take these models into account. We feel that it would be misleading to give the impression our hypotheses were motivated by these memory models. Instead, we decided to add a paragraph about the interesting question how our findings would fit into REM in the discussion (p. 23):

“It is intriguing to speculate how our findings relate to formal memory models such as REM (retrieving effectively from memory; Shiffrin & Steyvers, 1997). According to REM, each item is stored in memory as a vector consisting of different feature values. These features refer both to the content and the context of the item. During recognition, a probe vector (of the item to be recognised) is compared to the vectors stored in memory to determine which item to retrieve. In a cued recall procedure, as in our experiments, paired items (e.g., word-pairs) are thought to be stored in a common vector. When the cue is presented during recall testing, it will partially match the vector representing the pair of items, which enables us to retrieve the target associated with the cue. Items that were learned in a certain context will have similar context features, so context cues can be used to group items into different lists. Against this background, one might expect that new learning in different contexts is advantageous: When participants in the different context condition retrieve items from the second learning session, the probability of accidentally activating items from the first learning session would be low, because the items from the second and first session differ in their context features. Participants in the same context condition on the other hand, might erroneously match items from the first session during retrieval in the second session, because the items learned in different session share some context features. However, our results show the opposite pattern. One might argue that this is because the physical context features are not relevant and are de-weighted during memory retrieval: There is no overlap between the lists studied in the two sessions (A-B; C-D), so participants do not need to rely on context information to disentangle the two lists. Furthermore, when the items studied fall into the domain of associative processing (e.g., word-pairs), context manipulations may be less effective (Smith & Vela, 2001). However, while this explains why new learning in the same context is not a disadvantage, it does not explain why new learning in a different, unfamiliar context is detrimental.

**Also, I would urge you to take the type of context manipulation (i.e., multiple physical contexts), the type of processing at input (i.e., with these paired associates participants might engage in associative processing) and the memory task (i.e., cued recall) (see Smith and Vela (2001)) in your experimental set-up when formulating the hypothesis. For example, if participants are likely to engage in associative processing, the effects of context at the cued-recall test might be small. In any case, after reading the current version of the Introduction, I found your hypothesis not particularly compelling.**

As with your previous point, we are hesitant to alter our hypothesis (or the foundation for it) post hoc. Changing it now that we know the results could be considered a questionable research practice. Of note, Smith and Vela (2001) report that they did not find a significant effect of memory test on context effects. Only, when averaging across studies, they found an effect size close to zero for cued recall procedures, while effect sizes for free recall and recognition were larger than zero. Nevertheless, we have added some information about multiple physical contexts to the introduction. And we have added thoughts on associative processing to the discussion, where we address the fact that context cues might not have been that important for retrieval in the newly added paragraph about REM. We also added a paragraph about other types of memory test to the discussion.

(p. 3) “In their meta-analysis of context effects, Smith & Vela (2001) report learning in different contexts to be beneficial – however, only seven experiments in their meta-analysis were about new learning across several contexts. Furthermore, in all but two experiments, the final memory test occurred in an entirely new, unfamiliar context. It is possible that people whose learning sessions were spread across several contexts only had an advantage because at the time of testing, they were more accustomed to changing contexts. Furthermore, different methodological choices increased the variance even within studies.“

(p. 24) “One might argue that this is because the physical context features are not relevant and are de-weighted during memory retrieval: There is no overlap between the lists studied in the two sessions (A-B; C-D), so participants do not need to rely on context information to disentangle the two lists. Furthermore, when the items studied fall into the domain of associative processing (e.g., word-pairs), effects of context manipulations may be less pronounced (Smith & Vela, 2001).”

 (p. 27) “Furthermore, it would be interesting to see whether our results generalise to other forms of memory testing, other than cued recall. The meta-analysis by Smith and Vela (2001) suggests that context memory effects might be larger for free recall or recognition procedures (even though there was no meta-analytic effect of the type of memory test used).”

**A second major issue is that your results might simply be due to a processing/encoding deficit; you provide this explanation in the introduction of the manuscript. Simply put, processing the new context might take processing/encoding efforts away from the word pairs. Such a processing/encoding account could perfectly explain the results in both of your experiments. However, there are no direct measurements of the processing of the word pairs. Such measurements would be needed to make a compelling case for the deficit processing/encoding account. One option to remedy this problem is to do a replication of experiment 1 including direct measurements of processing/encoding. A second option would be to address the processing/encoding account in the discussion, acknowledge that crucial measurements are lacking for compelling evidence and argue why an explanation according to an attention account is theoretically interesting. Personally, I would favor the first option because this would lead to a test of the processing/encoding account and would provide a replication of experiment 1.**

While it seems plausible that the resources used to process the unfamiliar context reduce the amount of resources that can be allocated to process the learning task, we do not think that this explanation can adequately account for the observed pattern of results. This is because the first context the participants encounter in the morning is equally unfamiliar in both conditions, in experiment 1, and thus using it as a baseline in our analyses is sufficient to account for such a processing account. Improvement in the same context condition may indeed be because of improved processing, but impairment in the changed context condition cannot. If that were the explanation the performance level should be the same as in the morning, since that context at the time was equally unfamiliar. In addition, our control measures did not give any evidence that there was a processing deficit. Further evidence against the attention account comes from and additional the analysis of the time point when the context familiarization occurred in experiment 2 (either before or after the first learning session). Originally, we balanced the order of the first learning session and the familiarization with the other context to control for effects of task order. In response to the authors comment we realised that an analysis of the task order effects would provide a means of investigating whether a change of context impacts subsequent learning per se. We found no significant difference in learning performance for participants who were familiarised with the other context before or after the first learning session. Descriptively, participants who were exposed to the other context before learning even performed better.

Of note, we decided not to perform an additional experiment for several reasons. First, as explained above we do not believe there is a strong argument for a processing account to explain the pattern of our results. Second, the primary researchers performing this research are no longer at the institution that the research was performed at and the context labs are no longer available to us. Third, the current research was laborious and expensive (more so when adding e.g. eye tracking) and the additional insight does not seem commensurate to the costs. Instead, we elaborate on this further in the discussion, including how eye tracking could be used to investigate how much attention is drawn by the task and/or the context (p. 25).

“Unfamiliar contexts might lead to the allocation of attentional resources to the processing of the novel environment – resources that are then missing for the processing of the learning task. While this may explain the improved encoding in the same context condition, it is not sufficient to explain impaired encoding in the other context condition, of experiment 1. In the morning, both context conditions encoded in equally unfamiliar environments thus any processing bound by the changed context in the evening should also have been bound by processing the context in the morning. By statistically comparing encoding in the morning and the evening, we could show impaired performance in the changed context in the evening compared to the morning thus ruling out the processing account as sufficient explanation. Furthermore, if the unfamiliar context would have drawn attentional resources, this should also have been reflected in control measures like the verbal fluency task. However, there was no difference between the same and different context group in the control tasks. A further argument against an attentional processing deficit as explanation for the negative effects of a context change arises from the analysis of context familiarisation order in experiment 2 (before or after the first learning session). Exposure to the other context before learning did not affect learning performance. Descriptively, we even observed the opposite pattern: Participants who were exposed to a different context before the first learning session performed better than those who encountered the different context afterwards. This speaks against a deficit in attentional resources induced by a change of context. Nevertheless, it would be interesting to incorporate physiological measures of processing such as eye tracking in future studies to investigate how much attention participants devote to the task at hand and/or their environment.”

**A third major issue relates to the reference to metacognition in the introduction. I is not clear to me exactly what kind of metacognitive knowledge or skills, participants would employ in the current tasks and how this relates to the hypothesis. I propose to drop this part from the Introduction and focus on a more extensive grounding in the memory literature (see my first point) instead.**

We have deleted the part about meta-cognition from the introduction.

**A fourth major issue is that the results are not entirely correctly reflected in the discussion. In the discussion, the focus is on the results of experiment 1, which shows a context effect. I think it is important to discuss the difference in outcomes between the experiments and to provide an explanation for that (see my second point).**

In addition to the paragraphs we added in response to your second point (see above), the discussion now presents the results of both experiments in greater detail (p. 22):

“We investigated the effect of context on learning in two experiments. In the first, participants learned word pairs in the morning and then learned a new set of word pairs in either the same or a different context in the evening. We found that participants performed better when context was kept constant across learning sessions. While participants in the same context condition improved across learning sessions, participants in the different context condition learned less in the second session. In a second experiment, participants were additionally familiarized to the other context before the second learning session. This was done by letting participants complete a task without learning in the other context before or after learning in the morning. The aim of experiment 2 was to test whether different contexts have the same impact on new learning when they are familiar. The familiarization with the other context wiped out the differential effects found for the two context conditions in experiment 1. That is, no significant differences between the same and the different context condition emerged in experiment 2.

When comparing the two experiments the results supported a role of the familiarisation process to reduce the detrimental effect of changing contexts. For the first run of the evening learning session in the same context condition, performance in experiment 1 and experiment 2 does not differ. However, in the different context condition, learning performance during the first run of the evening learning session in experiment 1 was significantly lower than in experiment 2. While participants in the same context condition of experiment 1 even showed significantly improved performance in the second learning session, no such effect was observed in experiment 2. However, statistically a direct comparison between the same context condition across the two experiments shows that participants in experiment 1 do not show a significantly stronger improvement across sessions than participants in experiment 2. This means we can neither conclude that the familiarisation procedure disturbed the effect of learning in the same context, nor that it did not. However, it suggests that the effect of the familiarisation procedure on the context change condition was stronger. Nonetheless, it would be interesting to follow up this research to investigate whether the context familiarisation procedure interferes with the context consolidation in some way.”

**Furthermore, the part on the neuropsychological mechanisms underlying effects of contextual change on learning new information is highly speculative. I would suggest to drop this from the paper.**

We believe that behavioural studies are important to generate neuropsychological theories (see e.g. Krakauer et al. (2017), Neuron). This is why we think it is good practice to offer a neuropsychological explanation for behavioural results. However, we have shortened the description of the neuronal details (p 26; note that we also elaborated pattern separation/completion due to the comments of the other reviewer).

**A minor issue is that you foreshadow the results of experiment 1 in the introduction. This is strange because the details and the results of experiment 1 are presented after the introduction. Hence, I would propose to formulate the hypothesis, present the method and results of experiment 1, discuss the results and provide a rationale for experiment 1.**

We re-ordered the manuscript according to your feedback.

**Editor First Decision**: Revise & Resubmit

Dec 16, 2022

Dear Dr. Nagel,

I have now read your revised manuscript. I appreciate your careful attention to the concerns the reviewers raised. I am happy to provisionally accept your manuscript for submission. However, I found a few relatively small things I would like you to address.

Minor revisions:

* In the Results section of Experiment 2, you conduct an extra analysis to examine the Experiment-Context-Session three-way interaction on the performances in the first run. This interaction turned out to be significant, presumably because the crucial Context-Session interaction differs between the experiments. This result - and the results that unpack the three-way interaction are important, but for clarity, I would advise you to first report the results of Experiment 2, and then to report the exploratory between-experiment comparisons under a separate header.
* In the Discussion section, you argue that there is no convincing evidence that the familiarization procedure in Experiment 2 had an effect on the Context-Session interaction on the first-run results. The significant three-way interaction in combination with a significant Context-Session interaction in Experiment 1, but not in Experiment 2, however, suggests that familiarization had played a role (although the results are not entirey in line with a " familiarization hypothesis" ). Given your results, and considering that your central hypothesis gave no reason to expect an effect of familiarization, I would like to see some brief speculation in the Dicscussion as to why familiarization might have had an impact.
* Please check the text on small errors. For example, on page 3: Cox et al., (2021) report should be Cox et al., (2021) reported. Page 18: shouldn’t " …. on the same context condition" be " … on the different context condition" . In the first full paragraph on page 22, the present tense is sometimes used where the past tense would have been appropriate.

I look forward to receiving your final revision and accepting it for publication in Collabra: Psychology.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at editorialoffice@collabra.org.

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

Peter Verkoeijen

**Author Response**
Jan 19, 2023

Editor comments are highlighted in bold, and changes to the manuscript are highlighted in red.

**In the Results section of Experiment 2, you conduct an extra analysis to examine the Experiment-Context-Session three-way interaction on the performances in the first run. This interaction turned out to be significant, presumably because the crucial Context-Session interaction differs between the experiments. This result - and the results that unpack the three-way interaction are important, but for clarity, I would advise you to first report the results of Experiment 2, and then to report the exploratory between-experiment comparisons under a separate header.**

We restructured the result section of experiment 2 according to your feedback.

**In the Discussion section, you argue that there is no convincing evidence that the familiarization procedure in Experiment 2 had an effect on the Context-Session interaction on the first-run results. The significant three-way interaction in combination with a significant Context-Session interaction in Experiment 1, but not in Experiment 2, however, suggests that familiarization had played a role (although the results are not entirey in line with a " familiarization hypothesis" ). Given your results, and considering that your central hypothesis gave no reason to expect an effect of familiarization, I would like to see some brief speculation in the Dicscussion as to why familiarization might have had an impact.**

Thank you for this additional advice regarding clarity of our arguments. We are not entirely sure which part of the discussion this feedback refers to specifically, since we do state that familiarization has an effect at several points in several instances of the discussion (namely that it wipes out the differences between the same and different context condition we found in experiment 1).

At the end of the second paragraph, we say that we can neither confirm nor rule out that familiarization had an effect and we assume this is the section in question. Here, we refer to the benefit of repeated contexts across new learning. That is, we elaborate on the finding that participants improve across sessions in the same context condition of experiment 1, which is not the case in experiment 2, raising the interesting question whether familiarization with the different context disrupted the benefits of learning in the same context otherwise would have had. To clarify this, we rephrased this part of the discussion to emphasize which pattern of results we are referring to (p. 23):

“While participants in the same context condition of experiment 1 even showed significantly improved performance in the second learning session, no such effect was observed in experiment 2. However, statistically a direct comparison between the same context condition across the two experiments showed that participants in experiment 1 did not show a significantly stronger improvement across sessions than participants in experiment 2. This means we can neither conclude that the familiarization procedure disturbed the effect of improved learning in the same context, nor that it did not. However, it suggests that the effect of the familiarization procedure on the context change condition was stronger, where the negative effect of changing contexts was abolished. Nonetheless, it would be interesting to follow up this research to investigate whether the context familiarization procedure interferes with the context consolidation in some way.”

Additionally, as you advised, we have added a more speculative section about the mechanisms by which familiarization could have caused the observed effects to the paragraph about the REM model:

“Within the REM framework, the “environmental base rates” of the feature values within each item vector play a role during retrieval. That is, some feature values are more common than others, which is taken into account, when the match between the memory probe and the stored vectors is calculated. When a context is more familiar, this might mean that the features representing the familiar context are treated as more common. Importantly, the errors during the storage process are not independent of the environmental base rate. When a feature is not stored correctly, the feature is drawn randomly according to the environmental base rates of potential feature values. This makes it more likely that more plausible common feature values are stored. When a context is more familiar (i.e., the environmental base rate is higher), it might be more likely that the correct feature values are stored by accident in the case of a storage error. This would both explain why participants in the same context condition in experiment 1 improve across sessions, and why the familiarization procedure in experiment 2 abolishes the differences between the same and different context condition. However, while this explains the advantage of new learning in the same context, it does not explain why new learning in a different, unfamiliar context is detrimental. Furthermore, it is unclear whether a single exposure to the different context is enough to influence the environmental base rates of item features.”

**Please check the text on small errors.**

* **For example, on page 3: Cox et al., (2021) report should be Cox et al., (2021) reported.**
* **Page 18: shouldn't " .... on the same context condition" be " ... on the different context condition"**
* **In the first full paragraph on page 22, the present tense is sometimes used where the past tense would have been appropriate.**

We fixed the issues you raised, and additionally checked our manuscript for further errors. However, on page 18 (page 19 in the non-anonymised version), “on the same context condition” is correct.

**Editor Final Decision:** Accept

Feb 15, 2023

Dear Juliane Nagel,

I have now had a chance to read over your manuscript “Unfamiliar Contexts Compared to Familiar Contexts Impair Learning in Humans”, along with the letter describing the changes you made. Thank you for your responsiveness to the minor issues I raised. I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it will make an important contribution to the literature and I look forward to seeing it published! I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point.

You will be receiving separate correspondence regarding any production and technical comments, data deposits, as well as publication charges. We work with the Copyright Clearance Center to process any applicable APC charges. Please note that your APC transaction must be completed before your article gets published.

You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely,
Peter Verkoeijen