Title
Timing Is Everything: Effects Of Temporal Delay Of Confidence Judgments In Memory Decision-Making

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Klopukh, Isabella Rose Sky
Darby, Kevin P

Publication Date
2024

Timing Is Everything: Effects Of Temporal Delay Of Confidence Judgments In Memory Decision-Making

Isabella Klopukh (isabellaklopukh@gmail.com)
Department of Psychology, 777 Glades Road, Boca Raton, FL 33431 USA

Kevin P. Darby (darbyk@fau.edu)
Department of Psychology, 777 Glades Road, Boca Raton, FL 33431 USA

Abstract

Metacognitive confidence judgments are frequently adopted as a measure of certainty in decision-making tasks, but the mechanisms that underly these judgments have been long debated. In this work, we investigate the effect of the timing of confidence judgments in memory decisions by querying confidence immediately after, with a 3-second delay, or in a separate phase within an associative recognition task. An additional control condition did not probe confidence judgments at all to investigate how metacognitive monitoring may influence the memory decision-making process itself. The results indicate changes in memory performance and response times in conditions where confidence judgments were made, as well as a stronger association between confidence and accuracy when confidence was probed following a 3-second delay. We discuss the implications of these results regarding post-decision processing of metacognitive confidence and the bidirectional relationship between memory and metacognition.

Keywords: metacognition; confidence; decision-making; memory; associative memory; recognition memory

Introduction

Metacognitive confidence judgments measure how sure one feels in their decisions. Confidence can play a key role in decision-making, influencing how and why people make the decisions they do, but the underlying mechanisms of these judgments have been long debated. Here, we focus on behavioral evidence for two key issues of the confidence debate: the processing of metacognitive judgments over time and their influence on memory-based decision-making.

Metacognitive confidence is often studied in behavioral tasks where participants make a "primary" or initial decision (e.g., memory or perception related), as well as a judgment of confidence in that decision. Different theories concerning the latent processes underlying confidence judgments have been proposed, often in the form of computational models (Desender et al., 2021; Hellmann et al., 2023; Merkle & Van Zandt, 2006; Moran et al., 2015; Pleskac & Busemeyer, 2010; Yu et al., 2015). One prominent model (Pleskac & Busemeyer, 2010) assumes a two-stage process in which confidence judgements arise from a period of post-decision processing following the primary decision, which is thought to account for the ability to make accurate metacognitive assessments (i.e., higher confidence for correct responses and lower confidence for incorrect responses). To test this model, Pleskac and Busemeyer (2010) conducted an experiment in which confidence was always probed immediately after decisions in perceptual and general knowledge tasks (Pleskac & Busemeyer, 2010). However, without manipulating the time at which confidence judgments were probed, this paradigm did not allow for the behavioral investigation of post-decision processing. Similarly, other empirical data used to support dual-stage models come from paradigms where confidence judgments are made simultaneously with primary decisions, even though the models propose two phases of evidence accumulation corresponding to two separate decisions (Hellmann et al., 2023).

Thus, researchers have sought to behaviorally investigate the dynamics of post-decision processing of confidence by measuring metacognitive judgments at various times in the decision-making process. By comparing confidence judgements made simultaneously with, immediately after, or with a delay following perceptual decisions, studies can behaviorally probe how post-decision processing time can affect metacognitive confidence (Desender et al., 2021; Moran et al., 2015; Yu et al., 2015). Some researchers have found a stronger relationship between confidence and accuracy as the time between perceptual decisions and metacognitive judgments increases (Moran et al., 2015; Yu et al., 2015), which suggests that confidence may continue to be processed in this interjudgment time following primary decisions (but see Desender et al., 2021, for findings in which the confidence-accuracy relationship did not increase following a delay in another perceptual decision-making task). Similarly, behavioral evidence from a memory task has suggested that metacognitive confidence judgments have a stronger relationship with memory accuracy when made after as opposed to before memory decisions (Siedlecka et al., 2019), which is consistent with the dual-stage hypothesis that metacognitive processing continues after the primary decision has been made.

In contrast to dual-stage theories of metacognitive confidence, other computational models only involve one phase of processing (Merkle & Van Zandt, 2006). These models posit that metacognitive processes terminate at the same time as the primary decision and are based off of the same information. The relative balance of evidence model, for example, stipulates that confidence judgments are calculated based on the difference in the amount of evidence between alternatives at the time of the primary decision (Merkle & Van Zandt, 2006). This model was behaviorally supported by a paradigm in which confidence judgments were made after perceptual decisions, which may have

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

allowed for post-decision processing that was not accounted for by the model. It is therefore difficult to conclude from such evidence whether metacognitive confidence arises from the same information as primary decisions.

Single- and dual-stage models differ on the timing of when confidence is processed relative to the primary decision. A related question is whether the processing of confidence may impact primary decision-making. Although less work has examined this issue, Petrusic and Baranski (2003) found a significant increase in response times (RTs) when participants were required to make confidence judgments as opposed to solely making perceptual decisions. The authors of this work interpreted this as evidence that confidence may be processed in parallel with the primary decision, at least to some extent, and may take up cognitive resources that would otherwise be dedicated to the primary decision. Interestingly, however, Petrusic and Baranski (2003) did not find an impact of confidence judgments on the *accuracy* of perceptual decisions.

Overall, previous research paints an unclear picture of the processing of metacognitive confidence judgments over time and how they interact with primary decision-making processes, particularly in memory tasks. Though most research has investigated the processing of confidence in perceptual decision-making, confidence can play a large role in memory decision-making. The current study seeks to investigate the issues surrounding the processing of metacognitive confidence in a memory task by manipulating the time between confidence judgments and memory decisions, as has been done in perceptual decision-making paradigms (Desender et al., 2021; Moran et al., 2015; Yu et al., 2015). Similar to this perceptual work, we examine confidence judgments made immediately following memory decisions (the Sequential Decision condition) and with an additional 3-second delay in interjudgment time (the Delayed Decision condition). In order to better understand the potential bidirectional relationship between metacognition and memory, we add two new manipulations: one condition in which there are no confidence judgments at all (the Memory-Only condition; Petrusic & Baranski, 2003), serving as a baseline for performance, and another condition in which the memory decisions and confidence judgments are made in separate task phases (the Separate Decisions condition), serving as a control for the interdependence of memory and confidence processes.

Hypotheses

To address our research questions, we analyzed associative memory performance, confidence judgments, and memory RTs. Memory RTs can provide valuable insights into the decision-making process and help us understand how it is impacted by confidence judgments in terms of cognitive load and allocation of neurocognitive resources (Sternberg, 1969). We expected that RTs would be longer in the Sequential and Delayed Decision conditions compared to the Memory-Only and Separate Decisions conditions, since in the former groups determining metacognitive confidence may take up additional time and resources. Examining memory performance allowed us to investigate the impact of metacognitive monitoring on the memory decision-making process and outcome. We expected better associative memory performance in the Sequential and Delayed Decision conditions compared to the Memory-Only and Separate Decisions conditions due to the possibility that actively assessing one's memory performance could allow for metacognitive control and strategy adjustment, thereby improving memory responses. Lastly, confidence judgments, and specifically the association between confidence and memory accuracy, can provide information on the temporal dynamics of metacognitive processing, as well as what information or evidence metacognitive judgments are based on. We expected a stronger relationship between confidence and accuracy (higher confidence for correct responses, lower confidence for incorrect responses) in the Delayed Decision condition compared to the Sequential Decision condition, which would indicate post-decision processing. We also expected a weaker confidence–accuracy relationship in the Separate Decisions condition compared to the Sequential Decision condition because the processes of memory and confidence may be less likely to impact each other when separated in time.

Methods

Participants

Participants consisted of 164 college students (112 self-reported as Female, 46 as Male, and 2 as Other, with 4 non-responses; age: M = 18.75 years, SD = 1.99, Range 18–39, 15 non-responses) recruited from undergraduate introductory psychology classes. Participants received partial course credit as compensation. Two participants were unable to finish the experiment in the allotted time, and an additional 14 participants were excluded from analyses because their memory accuracy was not greater than chance based on a binomial test. This resulted in a final sample of 148 participants (100 Females, 43 Males, and 1 Other, with 4 non-responses; age: M = 18.78 years, SD = 2.08, Range 18–39, 14 non-responses). The study protocol was approved by the institutional review board of Florida Atlantic University.

Materials

Stimuli consisted of 160 scene–object pairs. These stimuli were randomly selected for each participant, with scenes extracted from the SUN Database (Xiao et al., 2010) and objects extracted from the Massive Memory Dataset (Konkle et al., 2010). Images with human faces, text, weapons, or religious symbols were excluded.

Procedure

We employed an associative recognition memory task where participants studied pairs of scenes and objects and were later tested on their ability to discriminate between "old" and "new" pairs (Castel & Craik, 2003; Naveh-Benjamin, 2000;

Ratcliff & McKoon, 2015). The State Machine Interface Library for Experiments (SMILE) Python library was used to present stimuli (https://github.com/compmem/smile).

The experiment comprised two blocks, each consisting of a study phase, retention interval, and test phase. In each study phase, 60 scene–object pairs were displayed one at a time for 3 seconds each, with a fixation cross presented in the center of the screen for 0.5 seconds in-between each pair. Following each study phase was a 60-second retention interval in which participants completed a Flanker inhibition task (Eriksen & Eriksen, 1974).

Participants then completed trials of the memory test, in which they were presented with a scene–object pair and asked to determine whether it was "old" or "new." A pair was considered "old," or *intact*, if the exact pair had been seen in the study phase. A pair was considered "new" if it was either: *recombined*, where both the scene and object appeared in the study phase but in different pairings; or *novel*, where neither the scene nor object had appeared in the study phase. A total of 80 pairs were presented in each test phase: 40 intact pairs, 20 recombined pairs, and 20 novel pairs (50% "old," 50% "new"). Unbeknownst to the participant, the first eight pairs of each test phase (which included four intact and two recombined pairs from the final six pairs of the previous study phase, as well as two novel pairs) were considered practice and excluded from analyses.

Participants were randomly assigned to one of four confidence judgment conditions: *Memory-Only*, where no confidence judgments were made (n = 40); *Sequential Decision*, where on each trial a confidence judgment was made immediately following a memory decision (n = 38); *Delayed Decision*, where on each trial a confidence judgment was made following a 3-second delay in which the stimulus pair was replaced by a fixation cross (n = 34); or *Separate Decisions*, where first memory decisions were made for all test pairs, before participants were asked to rate their confidence in each of those decisions in a second loop of all test pairs (n = 36). In the conditions where participants had to make confidence judgments, they were reminded of the memory decision previously made for each pair presented (i.e., "You decided this pair was 'old,' how confident are you in your decision?"). Confidence judgments were made on a 6-point scale, ranging from "Very Confident Correct" to "Very Confident Incorrect," allowing for participants to report memory errors or changes of mind (Desender et al., 2021).

Statistical Analyses

We fit linear and logistic mixed effects models using the package lme4 in the R Statistical Environment (Bates et al., 2015; R Core Team, 2023). For logistic models, regression output was used to analyze interactions, with Odds Ratios (OR) provided as measures of effect size. For the linear model, a Wald chi-square test using the Type III ANOVA function in R was performed to analyze the significance of model predictors. Tukey post-hoc tests were conducted to break down significant main effects using the package emmeans (Lenth, 2023), which evaluates the difference between estimated marginal means (EMmean) for the conditions in a comparison, as well as provides measures of effect size in terms of Cohen's d. Memory and confidence RTs above or below 2.5 standard deviations from individual subject means were considered outliers. Trials considered either memory RT outliers, confidence RT outliers, or both were excluded from analyses (a total of 7.49% of trials were excluded in this way).

Results

Memory Response Times

First, we investigated whether the timing of confidence judgments impacted the speed at which memory decisions were made in order to assess possible effects of confidence judgments on the memory decision-making process. A linear mixed effects model was run predicting log-transformed memory RTs from the between-subjects confidence judgment conditions (four levels: Memory-Only, Sequential Decision, Delayed Decision, and Separate Decisions) with random intercepts for each participant. There was a significant effect of confidence judgment condition ($X^2(3)$ = 141.35, p < .001). Tukey post-hoc tests comparing RTs between the conditions are shown in Table 1. Key findings include significantly longer response times in the Sequential Decision and Delayed Decision conditions compared to the Memory-Only and Separate Decisions conditions. A visualization of results is presented in Figure 1.

Table 1: Comparison of Memory Response Times Between Confidence Judgment Conditions

Comparison	EMmean	*SE*	*t*	*p*	*d*	95% CI
Memory-Only – Sequential	-0.39	0.05	-7.75	<.001***	-1.01	[-1.27, -0.75]
Memory-Only – Delayed	-0.50	0.05	-9.76	<.001***	-1.31	[-1.57, -1.04]
Memory-Only – Separate	-0.04	0.05	-0.77	.866	-0.10	[-0.36, 0.16]
Sequential – Delayed	-0.11	0.05	-2.21	.126	-0.30	[-0.57, -0.03]
Sequential – Separate	0.35	0.05	6.78	<.001***	0.91	[0.64, 1.17]
Delayed – Separate	0.46	0.05	8.77	<.001***	1.21	[0.93, 1.48]

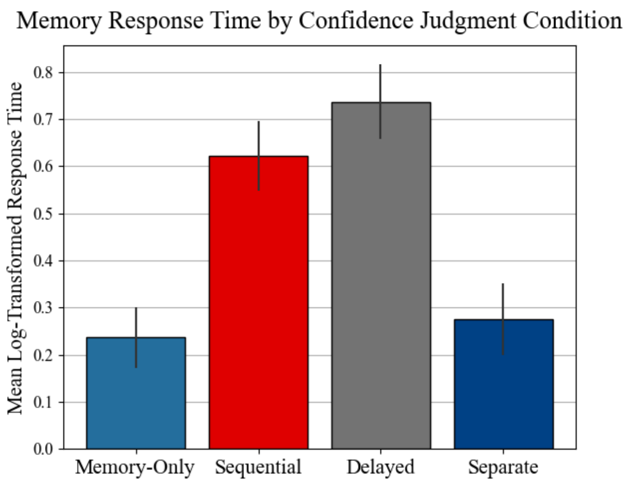

Figure 1: Mean log-transformed memory RTs by confidence judgment condition and memory pair type. To visualize results, RTs were averaged for each subject, then averaged overall for each between-subjects confidence judgment condition. Error bars represent 95% confidence intervals.

Associative Memory Performance

To assess whether memory processes may be impacted by the timing of confidence judgments, we compared memory performance between confidence judgment conditions. Memory performance was measured by comparing hit rates (correctly identifying an intact memory pair as "old") and false alarm rates (incorrectly identifying a recombined or novel memory pair as "old"). This comparison corresponds to a standard measure of recognition memory performance from signal detection theory, *d'* (DeCarlo, 1998), which probes how well participants can discriminate responses between pairs that are familiar and were associated together at study (intact), and pairs that are either comprised of equally familiar stimuli that were *not* associated together at study (recombined) or comprised of unfamiliar stimuli altogether (novel). A logistic mixed effects model was run predicting trial-level "old" response from confidence judgment condition (four levels: Memory-Only, Sequential Decision, Delayed Decision, and Separate Decisions) and memory pair type (three levels: Intact, Recombined, and Novel) with random effects of participant-specific intercepts and memory pair type coefficients. Model output is shown in Table 2, with base comparisons to the Memory-Only condition and Intact memory pairs. Key findings include a significant difference in "old" responses for intact vs. recombined pairs in the Sequential and Delayed Decision conditions compared to the Memory-Only condition, as well as a lack of significant difference between these conditions for intact vs. novel pairs. A visualization of these results is presented in Figure 2.

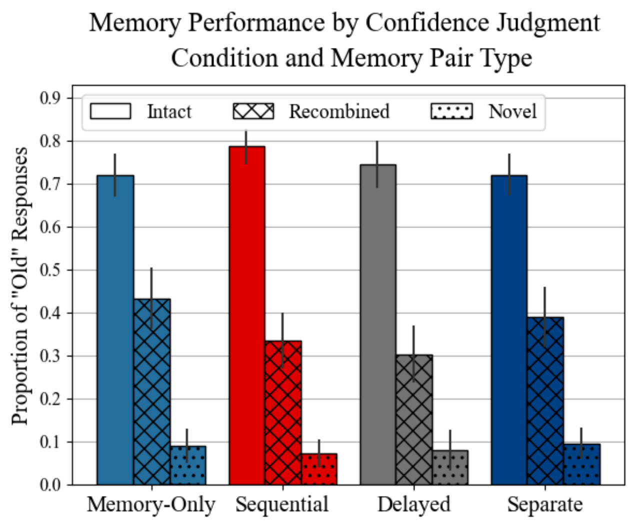

Figure 2: Mean proportion of "old" responses by confidence judgment condition and memory pair type. To visualize results, the proportion of "old" responses was calculated for each subject within each memory pair type, then averaged overall and sorted by between-subjects confidence judgment condition. Error bars represent 95% confidence intervals.

Confidence–Accuracy Relation

We also investigated whether the timing of confidence judgments impacted metacognitive accuracy, which was defined as the strength of the linear relationship between accuracy and confidence judgments (Siedlecka et al., 2019). Confidence judgments were coded numerically based on the ideal or expected level of accuracy for that rating, as follows: "Very Confident Incorrect": 0.0 (no correct responses), "Confident Incorrect": 0.2, "Somewhat Confident Incorrect": 0.4, "Somewhat Confident Correct": 0.6, "Confident Correct": 0.8, and "Very Confident Correct": 1.0 (perfect accuracy). The Memory-Only condition was excluded from

Table 2: Output from Logistic Regression Mixed Model Predicting Memory Performance

Coefficient	Estimate	*SE*	*z*	*p*	OR	95% CI
Intercept	1.11	0.14	7.69	<.001***	3.05	[2.29, 4.05]
Recombined: Sequential	-0.85	0.35	-2.41	.016*	0.43	[0.21, 0.85]
Recombined: Delayed	-0.83	0.37	-2.27	.023*	0.44	[0.21, 0.89]
Recombined: Separate	-0.15	0.36	-0.42	.677	0.86	[0.43, 1.74]
Novel: Sequential	-0.71	0.59	-1.21	.226	0.49	[0.16, 1.55]
Novel: Delayed	-0.74	0.62	-1.20	.229	0.48	[0.14, 1.59]
Novel: Separate	0.18	0.59	0.30	.764	1.19	[0.38, 3.80]

this analysis as there were no confidence ratings made. A logistic mixed effects model was run predicting trial-level accuracy from confidence judgment condition (three levels: Sequential Decision, Delayed Decision, and Separate Decisions) and numerical confidence rating (as a continuous variable) with random effects of participant-specific intercepts and confidence rating coefficients. Model output is shown in Table 3, with base comparisons to the Sequential Decision condition. Key findings include a significantly better relationship between confidence and accuracy for the Delayed Decision condition, and a significantly worse relationship between confidence and accuracy for the Separate Decisions condition. To better understand these differences in metacognitive accuracy, separate models were run predicting accuracy from confidence judgment condition for each half of the confidence scale, the "correct" side and the "incorrect" or error reporting side. For the correct side, there was no significant difference in the relationship between confidence and accuracy for the Delayed Decision condition compared to the Sequential Decision condition (Estimate = 0.33, *SE* = 0.59, *z* = 0.55, *p* = .058, *OR* = 1.39, 95% CI = [0.44, 4.40]). However, for the incorrect side there was a significantly better relationship between confidence and accuracy for the Delayed Decision condition compared to the Sequential Decision condition (Estimate = 4.67, *SE* = 1.35, *z* = 3.46, *p* < .001, *OR* = 106.38, 95% CI = [7.59, 1491.10]). A visualization of these results is presented in Figure 3.

Discussion

Our results show that manipulating the time between metacognitive confidence judgments and memory decisions can impact the accuracy of those confidence judgments, as well as some aspects of the memory decision-making process itself. Memory RTs were longer for conditions where confidence judgments were made following each memory decision (the Sequential and Delayed Decision conditions). Associative memory performance was also improved in these conditions – specifically, participants were most sensitive to differences between intact and recombined stimulus pairs in the Sequential and Delayed Decision conditions. Additionally, the relationship between memory accuracy and confidence was strongest in the Delayed Decision condition. We discuss the implications of all of these findings below.

The findings of slower and more accurate memory decisions when these decisions were followed by confidence judgments could be interpreted in multiple ways. One possibility is that participants were processing information about their memory and confidence at least partially in parallel, such that participants began to make their memory decisions and determine their metacognitive confidence at the same time. This would presumably require more neurocognitive resources than processing memory alone, which could slow down the decision-making process, resulting in increased RTs (Petrusic & Baranski, 2003). At the same time, parallel processing could improve memory accuracy if participants leverage metacognitive monitoring to gain greater control over memory recognition processes. Another possibility is that the presence of confidence judgments impacted participants' speed-accuracy tradeoff, or their willingness to spend more time to make a more accurate decision. The requirement of making confidence judgments may have motivated participants to use additional memory processing, thus lengthening the decision process while also improving accuracy. Unfortunately, the analyses we have presented here cannot distinguish between these possibilities, and future research, along with computational modeling, is needed to better understand how the timing of confidence judgments impacts the memory decision process.

The other, and perhaps most theoretically important, findings were the changes in the confidence–accuracy relationship (i.e., metacognitive accuracy) between the different confidence judgment timing conditions. There was a stronger relationship between confidence and accuracy when the time between memory decisions and confidence judgments was delayed by 3 seconds. This suggests that processing of confidence continued after memory decisions had been made, allowing confidence judgments to better reflect participants' accuracy, providing behavioral evidence in support of dual-stage theories (Pleskac & Busemeyer, 2010). It also replicates findings from perceptual decision-making studies that manipulate interjudgment time (Moran et al., 2015; Yu et al., 2015), but in a memory decision-making paradigm, which suggests there are likely shared confidence processes across task type. Further analyses revealed that this increase in metacognitive accuracy stemmed from improved error detection, or a stronger relationship between confidence and accuracy when participants responded with the "incorrect" half of the confidence scale, indicating a perceived error. in their memory response. This provides

Table 3: Output from Logistic Regression Mixed Model Predicting Accuracy from Confidence

Coefficient	Estimate	*SE*	*z*	*p*	OR	95% CI
Intercept	-0.18	0.25	-0.71	.479	0.84	[0.51, 1.37]
Delayed	-0.75	0.36	-2.10	.036*	0.47	[0.23, 0.95]
Separate	1.02	0.33	3.10	.002**	2.77	[1.45, 5.29]
Confidence Rating	2.04	0.31	6.65	<.001***	7.66	[4.21, 13.97]
Confidence Rating: Delayed	0.93	0.44	2.21	.035*	2.54	[1.07, 6.04]
Confidence Rating: Separate	-1.59	0.41	-3.89	<.001***	0.20	[0.09, 0.45]

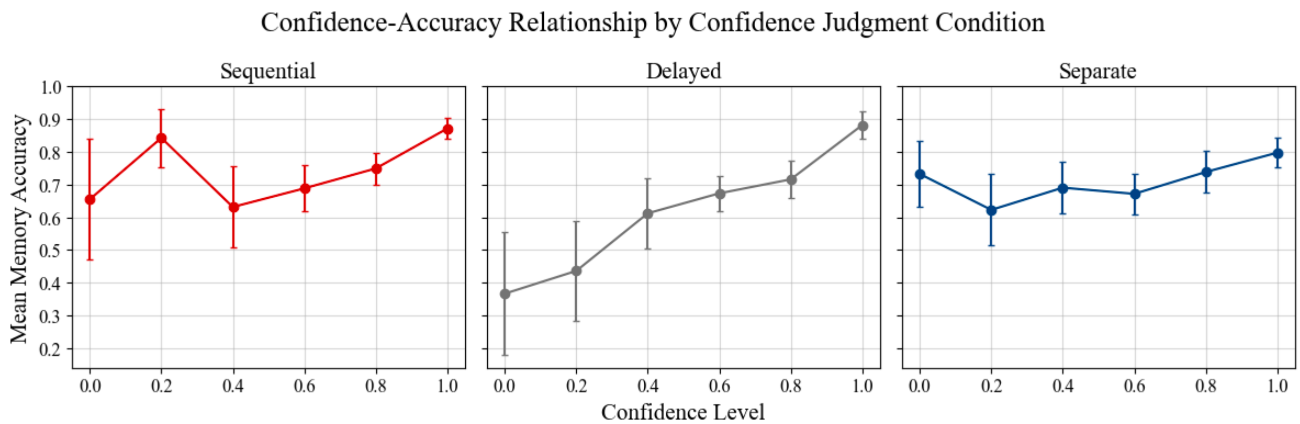

Figure 3: Mean memory accuracy as a function of confidence rating level. To visualize results, the proportion of correct responses for each confidence level was calculated for each subject, then averaged overall for each confidence level and sorted by confidence judgment condition. Error bars represent 95% confidence intervals.

additional support for dual-stage models, which posit that confidence for incorrect responses decreases as evidence is gained in favor of the unchosen decision alternative during the post-decision processing period (Pleskac & Busemeyer, 2010; Yu et al., 2015). Given that participants could not see the stimulus during the delay period, this improvement in error detection must have been due to additional time to process information extracted from memory or additional information gathered from making the memory decision itself (Siedlecka et al., 2019).

There was also a weaker relationship between confidence and accuracy when memory decisions and confidence judgments were separated into different task phases. This suggests that metacognitive confidence depends on memory processing. In the Separate Decisions condition, the two processes are disconnected, either by the length of time between memory decisions and confidence judgments or by interference from other memory decisions made in the interim. This poses significant changes to the mental context in which confidence judgments are made, decreasing access to memory decision-making processes, which could account for the weaker relationship between confidence and accuracy.

Limitations and Future Directions

The present work did not involve computational modeling, which limits our ability to formally test different mechanistic explanations of the results (see Darby & Sederberg, 2022 for discussion). Computational exploration and formal model comparison is needed to distinguish between different theories of how metacognitive confidence is processed in memory decisions, and how the timing of the two decisions could affect these underlying processes. For example, models could help determine to what extent memory and confidence are processed serially or in parallel to account for the differences in RTs and accuracy between confidence conditions. Another area which warrants further investigation is delays in interjudgment time; it is unclear whether post-decision processing of confidence terminated at some point on its own or continued for the entirety of the 3-second delay. This may be partially answered using computational models, but future work should further investigate these issues by varying interjudgment delays, to better understand the time course of memory decision-making and confidence, as well as optimize the confidence–accuracy relationship. Additionally, the current work did not include metacognitive confidence judgments made simultaneously with memory decisions, which would be based off of the same information as primary memory decisions. Prompting confidence judgments following memory decisions means there is some room for additional post-decision processing or information used to make those judgments, though less so compared to a 3-second delay. Future work should address this issue, incorporating both simultaneous and sequential confidence judgments to further explore the effects of post-decision processing.

Conclusion

The present study examined the temporal dynamics of metacognitive confidence in memory decision-making to better understand the mechanisms of metacognitive confidence. By varying the time at which confidence is queried, we gained insights into the processes that underly it. Results provided some support for the prominent dual-stage theory of metacognitive confidence, where processing of confidence continues after a primary decision been indicated. Improvements in metacognitive confidence with a delay has theoretical implications as well as potential practical implications for improving metacognition and memory decision-making in everyday life. Researchers should also consider that simply being asked to make confidence judgments may impact memory decision-making, namely memory performance and RTs. In the future, we hope to extend this work to investigate different delay lengths and use computational models to test theories of the mechanisms underlying confidence in memory decisions.

Acknowledgements

This work is derived from a project funded by Florida Atlantic University's Office of Undergraduate Research and Inquiry Summer Undergraduate Research Fellowship (2023).

References

Bates D., Mächler M., Bolker B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging, 18*(4), 873–885. https://doi.org/10.1037/0882-7974.18.4.873

Darby, K. P., & Sederberg, P. B. (2022). Transparency, replicability, and discovery in cognitive aging research: A computational modeling approach. *Psychology and Aging, 37*(1), 10–29. https://doi.org/10.1037/pag0000665

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods, 3*(2), 186–205. https://doi.org/10.1037/1082-989X.3.2.186

Desender, K., Donner, T. H., & Verguts, T. (2021). Dynamic expressions of confidence within an evidence accumulation framework. *Cognition*, *207*, 104522. https://doi.org/10.1016/j.cognition.2020.104522

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143-149. https://doi.org/10.3758/BF03203267

Hellmann, S., Zehetleitner, M., & Rausch, M. (2023). Simultaneous modeling of choice, confidence, and response time in visual perception. *Psychological Review, 130*(6), 1521–1543. https://doi.org/10.1037/rev0000411

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139*(3), 558–578. https://doi.org/10.1037/a0019165

Lenth, R. (2023). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.9.0. https://CRAN.R-project.org/package=emmeans

Merkle, E. C., & Van Zandt, T. (2006). An application of the Poisson race model to confidence calibration. *Journal of Experimental Psychology: General, 135*(3), 391–408. https://doi.org/10.1037/0096-3445.135.3.391

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive psychology*, *78*, 99-147. https://doi.org/10.1016/j.cogpsych.2015.01.002

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1170–1187. https://doi.org/10.1037/0278-7393.26.5.1170

Petrusic, W. M., & Baranski, J. V. (2003). Judging confidence influences decision processing in comparative judgments. *Psychonomic bulletin & review*, *10*(1), 177-183. https://doi.org/10.3758/BF03196482

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review, 117*(3), 864–901. https://doi.org/10.1037/a0019737

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Ratcliff, R., & McKoon, G. (2015). Aging effects in item and associative recognition memory for pictures and words. *Psychology and Aging, 30*(3), 669–674. https://doi.org/10.1037/pag0000030

Siedlecka, M., Skóra, Z., Paulewicz, B., Fijałkowska, S., Timmermans, B., & Wierzchoń, M. (2019). Responses improve the accuracy of confidence judgements in memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(4), 712–723. https://doi.org/10.1037/xlm0000608

Slane, C. R., & Dodson, C. S. (2022). Eyewitness confidence and mock juror decisions of guilt: A meta-analytic review. *Law and Human Behavior, 46*(1), 45–66. https://doi.org/10.1037/lhb0000481

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, *57*(4), 421-457. http://www.jstor.org/stable/27828738

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). Sun database: Large-scale scene recognition from abbey to zoo. In *2010 IEEE computer society conference on computer vision and pattern recognition*, 3485–3492. https://doi.org/10.1109/CVPR.2010.5539970

Yu, S., Pleskac, T. J., & Zeigenfuse, M. D. (2015). Dynamics of postdecisional processing of confidence. *Journal of Experimental Psychology: General, 144*(2), 489–510. https://doi.org/10.1037/xge0000062