

Memory in motion: Movement dynamics reveal memory strength

Megan H. Papesh · Stephen D. Goldinger

© Psychonomic Society, Inc. 2012

Abstract Recognition memory is typically examined as a discrete end-state, describable by static variables, such as accuracy, response time, and confidence. In the present study, we combined real-time mouse-tracking with subsequent, overt confidence estimates to examine the dynamic nature of memory decisions. By examining participants' streaming x -, y - mouse coordinates during recognition decisions, we observed that movement trajectories revealed underlying response confidence. More confident decisions were associated with shorter decision times and more linear response trajectories. Less confident decisions were made slowly, with increased trajectory curvature. Statistical indices of curvature and decision times, including area-under-the-curve and time to maximum deviation, suggested that memory strength relates to response dynamics. Whether participants were correct or incorrect, *old* responses showed a stronger correspondence between mouse trajectories and confidence, relative to *new* responses. We suggest that people subjectively experience a correspondence between feelings of memory and feelings of confidence; that subjective experience reveals itself in real-time decision processes, as suggested by sequential sampling models of recognition decisions.

Keywords Recognition memory · Confidence estimates · Mouse-tracking · Temporal dynamics

M. H. Papesh
Department of Psychology, 236 Audubon Hall,
Louisiana State University,
Baton Rouge, LA 70803, USA
e-mail: mpapesh@lsu.edu

S. D. Goldinger (✉)
Department of Psychology, Arizona State University,
Tempe, AZ 85287-1104, USA
e-mail: goldinger@asu.edu

In the study of recognition memory, a long-standing question has concerned the relationship of memorial accuracy to subjective feelings of confidence (e.g., Busey, Tunnicliff, Loftus, & Loftus, 2000; Dobbins, Kroll, & Yiu, 1998) and the associated concept of variations in *memory strength* (Wixted & Mickes, 2010). Although accuracy–confidence dissociations have been frequently observed, it is typically the case that participants are faster, more accurate, and have greater recollective detail when they express higher recognition confidence (Mickes, Wixted, & Wais, 2007; Ratcliff & Murdock, 1976). Such findings are consistent with both continuous signal detection theories (Wixted, 2007) and dual-process theories (Yonelinas, 2002). By either view, when a person correctly recognizes studied items as *old*, these memories range in relative strength, typically measured using metacognitive confidence estimates (e.g., 1 [*very sure new*] through 7 [*very sure old*]). The nature of confidence, as a proxy either for memory strength or separate processes, is central to theories of recognition memory (Parks & Yonelinas, 2007).

Although *very strong* memories are difficult to estimate using standard Likert scales (Mickes, Hwe, Wais, & Wixted, 2011), confidence ratings are almost exclusively used to gauge memory strength. However, participants' ultimate decisions, confidence estimates, and response times (RTs) may not reflect the same latent cognitive processes, as many standard memory theories assume (Pleskac & Busemeyer, 2010). Therefore, some researchers have proposed that all three components of recognition decisions may be described by *sequential sampling models* (e.g., Ratcliff & Starns, 2009), which suggest that perceivers continuously sample memorial strength from test items, comparing accumulated strength with a decision criterion. Recent models explain the relationship between decisions, confidence, and RTs by incorporating a *dynamic drift diffusion* process, suggesting that evidence for old/new responses accrues over time. The

present study was an attempt to make that time course observable, using *dynamic mouse-tracking* (Spivey, Grosjean, & Knoblich, 2005) to complement standard accuracy and RT measures. As described below, mouse-tracking provides rich, trial-level information during recognition (or other) decisions. In the present study, we report that, when people generated *old* recognition decisions, there was a close correspondence between physical movements of the computer mouse and subjective confidence. This pattern emerged for both hits and false alarms but was weaker (although still present) during *new* responses, either misses or correct rejections (CRs). The results are consistent with recent suggestions (e.g., Kello, Beltz, Holden, & Van Orden, 2007) that, when people engage in cognitive-behavioral tasks, they *self-organize* to produce stable performance. For example, Kloos and Van Orden (2010) described such “soft assembly,” suggesting that “[task] constraints self-organize as temporary dynamic structures that span the mind-body divide” (p. 19). Related to the present research, we suggest that, when people are given a “memory test,” they become sensitive to indications of memory strength (Cox & Dobbins, 2011), with greater appreciation for evidence of familiarity, relative to evidence of novelty. Under the hypothesis of soft assembly, when memory decisions must be indicated by moving a computer mouse, such movements will become tightly coupled to task-relevant cognitive activity. In the present case, the relevant cognitive activity is temporally accumulating evidence that a test item is *old*.

The mouse-tracking paradigm

Investigations into the relationship between bodily movements and cognitive processes have shown that the speed and force of buttonpresses is related to successful memory retrieval (Abrams & Balota, 1991) and that arm movements during response execution progressively index learning of paired associates (Dale, Roche, Snyder, & McCall, 2008). Two prominent methods are currently available to examine the continuous accumulation of information in cognitive processes: eye-tracking and mouse-tracking. Results from these techniques have supported an emerging view that cognition reflects the dynamic interplay of temporally continuous variables, as opposed to discrete, stage-like processes (Spivey & Dale, 2006). For example, Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995) developed an eye-tracking paradigm to observe real-time phonetic competition in spoken word recognition. When participants were shown two objects, eye

movements to targets (e.g., *candle*) were faster when the alternative object’s name did not overlap phonetically (e.g., *pickle*), relative to when it did (e.g., *candy*). Such results (see also Allopenna, Magnuson, & Tanenhaus, 1998) suggest that eye movements reveal dynamic cognitive operations, such as lexical competitions unfolding in real time.

Although eye-tracking has provided insight in many domains of cognitive science, it does not clearly differentiate discrete versus continuous processes: Because eye movements are ballistic, analyses must average across many trials in which the eyes were fixating relatively few locations (Magnuson, 2005). Thus, eye movements may reflect rapid transitioning of discrete mental states. Mouse-tracking may better reveal the continuous dynamics of cognition. Spivey et al. (2005) had participants view images of cohort competitors or noncompetitors (e.g., *candle* and *jacket*) and quickly click the images corresponding to spoken words. By recording the streaming *x*-, *y*- mouse coordinates as participants responded, Spivey et al. observed dynamic attraction from the competitor objects as target words unfolded in time. Participants’ mouse trajectories “curved” toward the phonological competitor during cohort trials but followed a more linear path during control trials.

Mouse-tracking has been broadly used in cognitive and social science to examine decision processes. In a moral-reasoning task, low-truth value statements (e.g., *Murder is sometimes justified*) were associated with curved response trajectories, relative to high-truth value statements (e.g., *Should you brush your teeth everyday?*; McKinstry, Dale, & Spivey, 2008). In race and sex judgment tasks, researchers typically find that category representations are simultaneously active when judged stimuli are less clearly defined, as when faces are racially ambiguous (Freeman & Ambady, 2009, 2011). Response trajectories track the time course of decisions, revealing response competition during more challenging decisions.

In the present study, we examined covert memory strength using mouse-tracking. Participants studied a series of words and, in a later recognition test, they made *old/new* decisions by clicking labeled boxes on the screen. During these decisions, we tracked the streaming *x*-, *y*-coordinates of their mouse movements (Spivey et al., 2005). After each decision, participants gave verbal confidence estimates. The questions of interest were (1) whether mouse trajectories during recognition would predict subsequent confidence estimates and (2) whether such a relationship would differ as a function of *new* versus *old* decisions. According to many theories (e.g., Wixted, 2007), the hypothetical target distribution is wider than the lure distribution. This may

reflect either a greater range of memory strength for old items or enhanced attention to familiarity signals in a recognition test. In either case, mouse trajectories may reveal that recognition decisions are not symmetric but that people selectively attend to evidence that items might be *old*.

Method

Participants

Thirty-five right-handed Arizona State University students participated for partial course credit. All participants had normal or corrected-to-normal, vision and were native English speakers. One participant was excluded for failing to follow instructions, leaving 34 participants (20 men, 14 women; $M_{\text{age}} = 19$ years). Participants were tested individually in a sound-attenuated booth.

Materials

Words ($n = 128$) with concreteness ratings ranging from 550 to 700 were selected from the MRC Psycholinguistic Database (Coltheart, 1981). High- and low-frequency words were intermixed, with a goal of eliciting diverse confidence estimates (Glanzer, Adams, Iverson, & Kim, 1993). All words were pseudorandomly assigned to two lists and were used equally often as targets and foils. Stimuli were presented centrally on a 15-in. CRT screen (resolution, 1,024 × 768 pixels) in 28-point Courier font, with study procedures managed by E-Prime software (Psychology Software Tools, 2006) and test procedures managed by MouseTracker (Freeman & Ambady, 2010).

Procedure

Participants first completed 64 study trials. Each trial began with a 1,500-ms central fixation cross, followed by a 2,000-ms display of a randomly selected word. Participants were asked to memorize each word, in preparation for an upcoming memory test. Study trials were followed by a 2-min break, during which participants solved anagrams. Following the break, the researcher launched MouseTracker and gave test instructions. Test trials ($n = 128$, half old) began when participants clicked a “start” button at the bottom-center of the screen, which triggered the appearance of a centrally presented test word. Participants judged words *old* or *new* by clicking the corresponding options in the upper left and right corners of the screen (response mapping was counterbalanced across participants). During memory decisions, the x -, y -coordinates of

the mouse pointer were sampled at 70 Hz. To ensure that response trajectories revealed online decision making, participants were encouraged to begin their movements early and were warned after slow trials. If a response was not made within 3,000 ms, or if initial movements began more than 400 ms after stimulus onset, that trial was discarded from analysis.¹

Following each recognition decision, participants verbally estimated their confidence along a scale ranging from 1 (*not at all confident*) to 7 (*very confident*). The researcher recorded the estimate during a 1,000-ms intertrial interval.

Results

For all analyses, alpha was set to .05, and multiple comparisons were subject to Bonferroni corrections. All mouse-tracking data conversions were conducted via automated procedures in the MouseTracker software. Prior to analysis, overt confidence estimates were grouped into three categories: low (estimates of 1–4), medium (estimates of 5–6), and high (estimates of 7). The lower categories included wider ranges of estimates to provide balanced values and avoid missing data.²

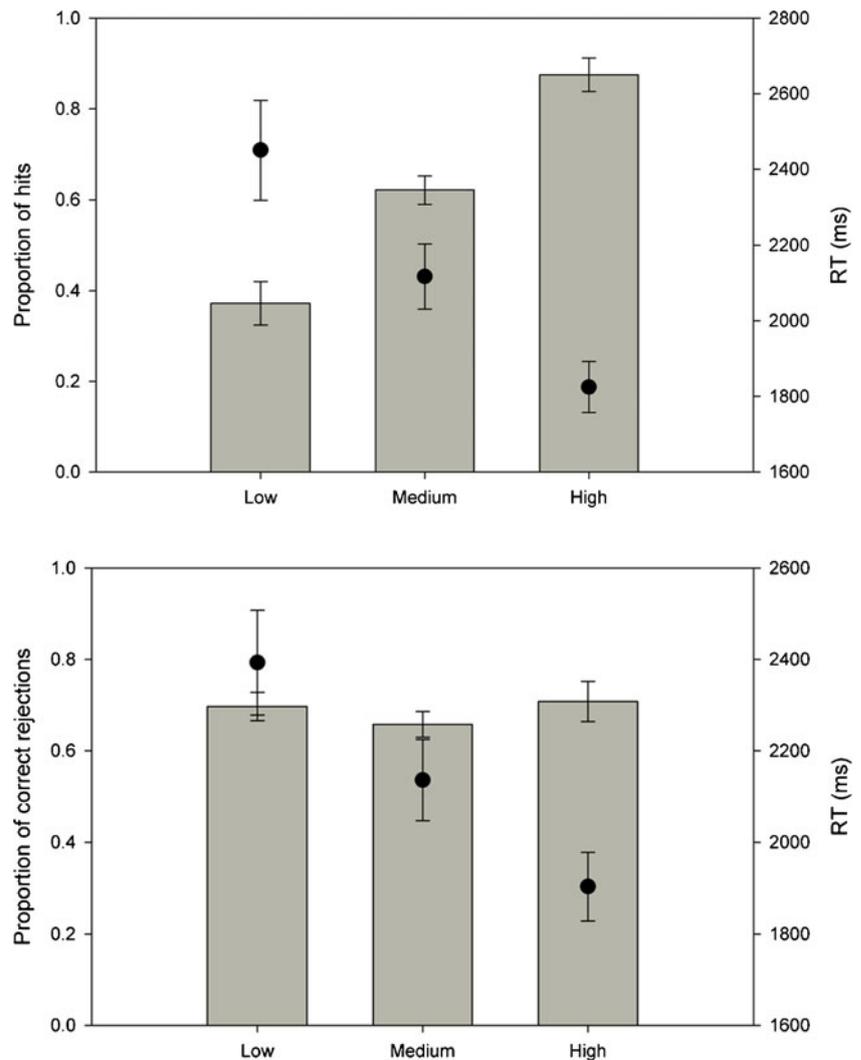
Recognition accuracy and RT

Standard signal detection indices were computed for each participant, including d' and c (a bias measure centered at zero; higher values represent more conservative biases). Hit and CR rates were analyzed as a function of confidence (low, medium, high) in a three-level within-subjects ANOVA. As is shown in the upper panel of Fig. 1, hits increased linearly with confidence, $F(2, 31) = 38.7, p < .01, \eta_p^2 = .71$. In pairwise comparisons, all differences were reliable [low–medium, medium–high, low–high, all $t(32) > 4.3, ps < .001$]. RTs followed a complementary pattern; increases in confidence were associated with decreases in RT, $F(5, 24) = 22.3, p < .01, \eta_p^2 = .65$. In pairwise comparisons, all differences were again reliable [low–medium, medium–high, low–high, all $t(32) > 4.4, p < .001$].

¹ As was noted by Freeman and Ambady (2010), it is necessary to have participants begin moving the mouse quickly, to ensure that decision processes are still occurring while mouse movements are being recorded. Across all participants, fewer than 3 % of trials were excluded.

² As was noted by Mickes et al. (2011), participants are typically biased toward higher-confidence judgments. In the present case, several participants never gave estimates of 1, which would have excluded too much data. We selected confidence intervals that evenly divided the ratings into three categories. To ensure that the uneven binning procedure did not unduly affect the results, all comparisons involving confidence were tested using both ANOVAs, which assume linearity across the bins, and simple pairwise comparisons, which did not assume linearity.

Fig. 1 Proportion of hits (upper panel) and correct rejections (lower panel) as a function of confidence. Overlaid circles represent average response times (RTs, in milliseconds). Error bars represent standard errors



As is shown in the bottom panel of Fig. 1, correct rejections were equivalent across levels of confidence, $F(2, 31) = 0.89, p = .41$, and all pairwise comparisons were null ($p > .48$). Although accuracy did not increase with confidence, CRs were issued faster when participants were more confident, $F(2, 31) = 30.9, p < .001, \eta_p^2 = .51$. In pairwise comparisons, all differences were reliable (low–medium, medium–high, low–high, all $ps < .001$)

Mouse-tracking

Using MouseTracker, we rescaled response trajectories into a standard coordinate space and normalized them into 101 time-steps; this allows all responses to be compared, regardless of absolute RTs. To examine the association of confidence and response dynamics, trials were again categorized by confidence, and mouse trajectories were examined by area-under-the-curve (AUC). AUC represents the area between the observed and ideal (i.e., linear) trajectories and is used to measure the dynamic attraction to response

competitors (for details and validation, see Freeman & Ambady, 2010, p. 230).

Considering first *old* responses, the main portion of Fig. 2 shows average mouse trajectories corresponding to hits, with separate functions for low-, medium-, and high-confidence responses. As is shown, curvature increased with decreases in confidence. The smaller, inset panel shows the trajectories for false alarms, which followed the same overall pattern. However, because few participants contributed false alarms to all three confidence bins, only the hits were statistically analyzed. For the hits, high-confidence responses were associated with more linear trajectories, relative to medium- and low-confidence decisions. AUC values were first compared across confidence categories in a three-level within-subjects ANOVA, with Greenhouse–Geisser corrected degrees of freedom. The main effect of confidence was reliable, $F(1.5, 46.6) = 8.6, p = .002, \eta_p^2 = .22$, supporting the subjective impression of the trajectory plot. In pairwise comparisons, curvature differed between high- and medium-confidence hits, $t(32) > 2.7, p < .05$, and also differed between high- and low-confidence

hits, $t(26) > 2.9, p < .02$. The difference between medium- and low-confidence hits was marginal, $t(27) > 1.9, p < .09$.

The foregoing analysis suggested that curvature of the mouse trajectories was related to underlying confidence. One concern, however, is that creating artificial confidence bins (low, medium, high) may have unduly imposed such structure on the data. To ensure that the apparent relationship was not an artifact of the binning procedure, we conducted a second set of analyses. For each participant, we calculated the correlation (Pearson's r) between AUC and expressed confidence, across trials. For trials producing hits, the correlation was negative (i.e., higher confidence was associated with lower curvature) and reliably differed from zero, $r = -.202, SE = .04, t(33) = -5.38, p < .001$. Therefore, the relationship suggested by Fig. 2 still held, even when separate functions were not derived for different confidence bins.

It is also hypothetically possible that the results reflect an artifact of the averaging process. Specifically, a concern with mouse-tracking is that, in a subset of trials, the participant may make discrete movements in one direction and then reverse course toward the eventual choice. Given a handful of such trials, the average trajectories may have a misleading curvilinear form. To address this, we computed the bimodality coefficient, b , from the distribution of z -normalized AUC values provided by MouseTracker. If b exceeds .555, the distribution is bimodal; otherwise, it is unimodal (see Freeman & Ambady, 2010, p. 231). In our case, b was less than .0001, allowing us to confidently reject the bimodality hypothesis.³ This is important for data validation, and it helps clarify the likely direction of effect between mouse trajectories and confidence. That is, instead of implicit confidence affecting curvature, people might be aware of their own mouse trajectories, then heuristically infer confidence levels after the fact. Although such an interpretation cannot be ruled out in the present study, it seems likely that people would mainly be aware of large variations in mouse trajectories, such as direction reversals, rather than the rather subtle differences we observed.

Considering next *new* responses, the main portion of Fig. 3 shows average mouse trajectories corresponding to CRs, with separate functions for low-, medium-, and high-confidence responses. As is shown, there was little systematic relationship of curvature and confidence. The smaller, inset panel shows the trajectories for misses, which showed the same pattern. Again, because few participants contributed misses in all three confidence bins, only the CRs were statistically analyzed. For the

³ The time necessary to reach the maximum deviation (MD) from a linear path was also examined. During old test trials, MD time was reliably influenced by confidence, $F(2, 29) = 16.07, p < .01, \eta_p^2 = .53$; as confidence increased, MD decreased. This pattern was also evident in new test trials, $F(2, 29) = 13.64, p < .01, \eta_p^2 = .49$.

CRs, although the relationship was relatively weak, there was a reliable main effect of confidence, $F(1.8, 55.4) = 5.67, p = .009, \eta_p^2 = .15$. Despite this main effect, none of the pairwise comparisons were reliable (all t s $< 1.6, p$ s $> .11$). As with the hits, we next conducted a correlation analysis on the CRs, without separating trials into confidence bins. The correlation was again negative and reliably differed from zero, $r = -.127, SE = .03, t(33) = -4.39, p < .001$.

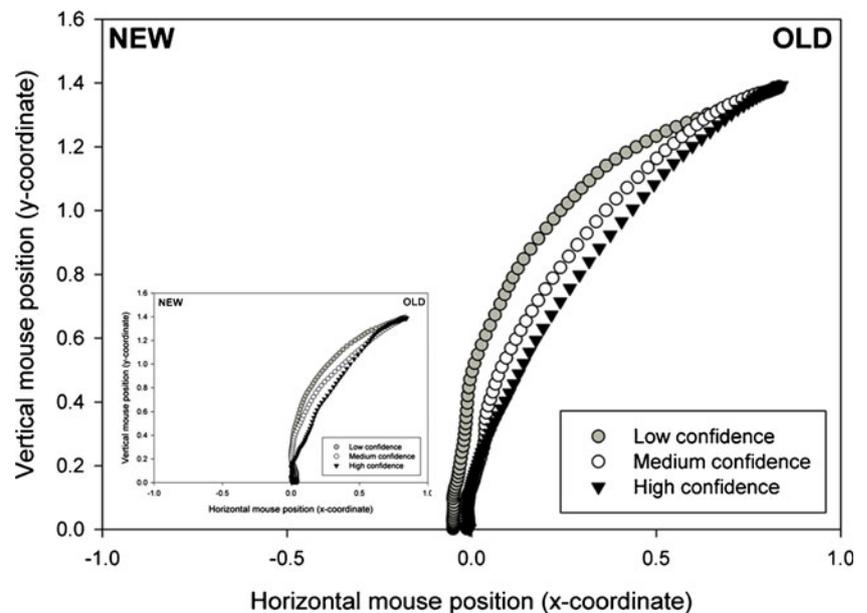
Although we observed a relationship between trajectory curvature and confidence for both hits and CRs, the association appeared more robust for *old* trials. To assess this, we first conducted a 2×3 ANOVA with factors of response type and confidence level. The key interaction was marginal, $F(2.7, 109.8) = 3.03, p = .06, \eta_p^2 = .07$, suggesting a stronger relationship in the hit trials. To test this more directly, the correlation coefficients derived for hit and CR trials were Fisher-transformed and compared, showing a reliable difference, $t(33) = 2.18, p < .05$. Taken together, the results suggest that confidence was related to response trajectories across trials, with a slightly stronger relationship for recognition hits.

Discussion

The present results suggest that subjective memory strength accrues in real time and is observable in behavioral response dynamics.⁴ Participants in this study made *old/new* decisions while their mouse coordinates were tracked, followed by confidence estimates along a 1–7 scale. By examining response trajectories and subsequent confidence, we observed a reliable correspondence of apparent memory strength and movement dynamics, primarily in trials that participants (correctly or incorrectly) judged as *old*. Although confidence estimates were issued after recognition decisions were complete, there was a clear relationship: Stronger memories yielded fast, linear movements, while weaker memories yielded slower, curvilinear movements. We suggest that such continuous, dynamic information may offer researchers new insight, complementing the standard accuracy and RT measures in recognition memory. In the present case, metacognitive confidence predicted both hit rates and RTs, as has been frequently reported (e.g., Johnson, 1939). By virtue of mouse-tracking, we also

⁴ Rather than variations in memory strength, an alternative interpretation is that response criteria varied across trials (Benjamin, Diaz, & Wee, 2009). Although this is plausible, whether the findings reflect variations in memory strength or criteria, the deeper interpretation is similar, since the hypothetical “distance” from the memory signal to the criterion is revealed by curvature of the mouse dynamics. Notably, Dale et al. (2008) observed increased linearity as paired-associate learning became stronger, in a task that required no *old/new* decisions.

Fig. 2 Time-normalized mouse trajectories during *old* decisions as a function of subsequent confidence ratings. All responses are remapped to the right. The main panel shows hit responses; the inlaid plot shows false alarm responses

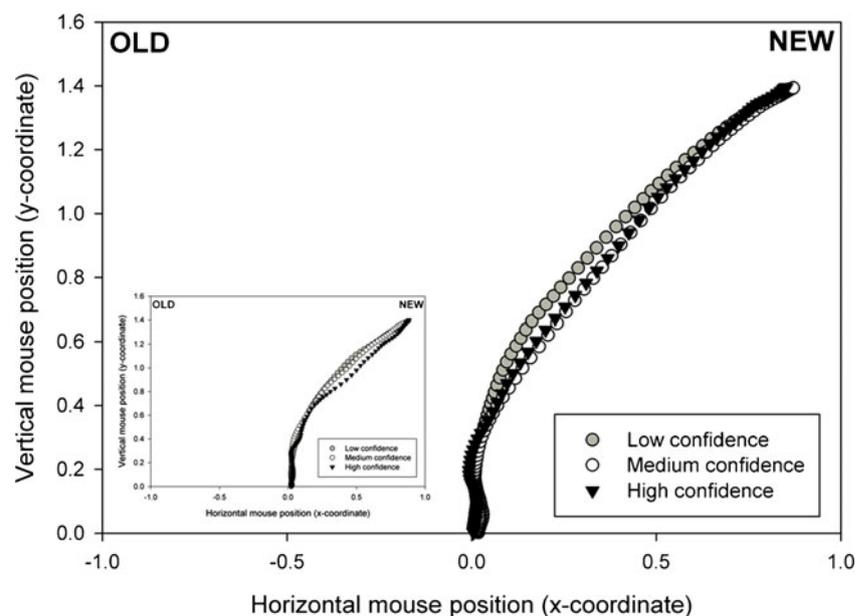


observed that less confident *old* decisions were characterized by slow-resolving competition between the *old* and *new* response options, giving greater insight into the processes leading to eventual decisions and RTs.

Our results are (at least metaphorically) compatible with sequential sampling models, which propose that recognition decisions are based on the accumulated strength of stimulus information. For example, Pleskac and Busmeyer (2010) recently developed a *two-stage dynamic signal detection theory* (2DSD), which combines signal detection theory (Green & Swets, 1966) with random-walk/diffusion theory (see Ratcliff & Starns, 2009; Vickers, 1979). During the first stage of processing, information (strength) is sequentially

sampled to favor one alternative over the other. Perceivers make a choice when enough evidence has accumulated that the trajectory “drifts” toward one of the alternatives and a criterion is crossed. Critically, however, perceivers do not stop collecting evidence once the random walk process has surpassed an *old* or *new* criterion; cognitive states do not discretely transition to a second stage. Instead, evidence continues to accrue, and an eventual confidence estimate is based on the total strength of accumulated information. The present results accord nicely with the theoretical premises of such diffusion models, since they suggest continuous, strength-driven evidence evaluation. Participants (literally) moved toward recognition decisions faster and more directly

Fig. 3 Time-normalized mouse trajectories during *new* decisions as a function of subsequent confidence ratings. All responses are remapped to the left. The main panel shows correct rejection responses; the inlaid plot shows miss responses



when they had higher confidence, in a manner that resembles information accumulation in the 2DSD theory. Of course, it will take considerable theoretical development to establish a valid “linking hypothesis” (Alloppenna et al., 1998) between mouse trajectories and theoretical constructs such as diffusion processes.

Beyond offering a potential connection to dynamic recognition theories, the mouse-tracking results also revealed an asymmetry: When people believed items were *old* (either correctly or incorrectly), there was a tight correspondence between confidence and mouse-trajectory curvature. When people believed items were *new*, this relationship was observed, but weaker. This pattern supports several interpretations. First, it is often theorized that recognition memory reflects *unequal-variance signal detection* (Wixted, 2007). In theory, when a person studies a list of words, each one receives some increment in memory strength, giving rise (during test) to a distribution of targets with greater average “strength” than a corresponding lure distribution. If every word received an equivalent boost from study, we would expect these distributions to have roughly equal variance. But, under the more likely assumption that study items vary in their “memory boosts,” we expect the target distribution to have greater variance than the lure distribution, a prediction that has been frequently supported by ROC analyses (e.g., Mickes et al., 2007). With respect to the present study, we may hypothesize that mouse trajectories correspond to sampled memory strength: Greater variation arises in dynamic decision movements because greater variation exists in implicit strength.

Second, the results may reflect occasional recollection. In *high-threshold* models, recognition decisions may reflect gradations of familiarity or qualitatively separate episodes of recollection. We may hypothesize that more linear mouse trajectories reflected moments of recollection and that more curved trials reflected gradations of familiarity, a concrete hypothesis for future research. Third, the results may reflect participants’ orientation toward the recognition task. Cox and Dobbins (2011) suggested that, when people are given a “memory test,” they may place a premium on feelings of memory, with less focus on feelings of novelty. They reported that memory decisions were distributed nearly identically, even in tests without lures or without targets. When asked to evaluate whether items feel familiar, people may dynamically assemble a correspondence between feelings of memory and, in this case, the movements used to indicate those feelings. This ability to self-organize cognition and action (Kloos & Van Orden, 2010) forms the underlying logic of the mouse-tracking paradigm (Magnuson, 2005). In the present case, an implicit orientation toward “memory signals” may support a tight correspondence between latent memory strength and observable action.

Author Note This research was supported by NIDCD Grant R01-DC04535-11, awarded to S. D. Goldinger. We thank Tresa Marchi,

Melissa Miola, Suhani Mehrotra, and Rachele Friedman for their assistance with data collection.

References

- Abrams, R. A., & Balota, D. A. (1991). Mental chronometry: Beyond reaction time. *Psychological Science*, 2, 153–157.
- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439.
- Benjamin, A. S., Diaz, M. L., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review*, 116, 84–115.
- 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.
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A, 497–505.
- Cox, J. C., & Dobbins, I. G. (2011). The striking similarities between standard, distractor-free, and target-free recognition. *Memory & Cognition*, 39, 925–940.
- Dale, R., Roche, J., Snyder, K., & McCall, R. (2008). Exploring action dynamics as an index of paired-associate learning. *PLoS ONE*, 3, e1728.
- Dobbins, I. G., Kröll, N. E. A., & Liu, Q. (1998). Confidence–accuracy inversions in scene recognition: A remember–know analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 24, 1306–1315.
- Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, 20, 1183–1188.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, 42, 226–241.
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118, 247–279.
- Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review*, 100, 546–567.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford: Wiley.
- Johnson, D. M. (1939). Confidence and speed in the two-category judgment. *Archives of Psychology*, 34, 1–53.
- Kello, C. T., Beltz, B. C., Holden, J. G., & Van Orden, G. C. (2007). The emergent coordination of cognitive function. *Journal of Experimental Psychology: General*, 136, 551–568.
- Kloos, H., & Van Orden, G. C. (2010). Voluntary behavior in cognitive and motor tasks. *Mind & Matter*, 8, 19–43.
- Magnuson, J. S. (2005). Moving hand reveals dynamics of thought. *Proceedings of the National Academy of Sciences*, 102, 9995–9996.
- McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science*, 19, 22–24.
- Mickes, L., Hwe, V., Wais, P. E., & Wixted, J. T. (2011). Strong memories are hard to scale. *Journal of Experimental Psychology: General*, 140, 239–257.
- Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal-detection model of recognition memory. *Psychonomic Bulletin & Review*, 14, 858–865. doi:10.3758/BF03194112
- Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal detection models: A reply to Wixted (2007). *Psychological Review*, 114, 188–202.

- Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal-detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*, 864–901.
- Psychology Software Tools. (2006). E-Prime (Version 1.2). Retrieved from <http://www.pstnet.com>
- Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*, 190–214.
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*, 59–83.
- Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, *15*, 207–211.
- Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, *102*, 10393–10398.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.
- Vickers, D. (1979). *Decision processes in visual perception*. New York, NY: Academic Press.
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*, 152–176.
- Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*, 1025–1054.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.