TITLE

Pupil dilation during recognition memory: Isolating unexpected recognition from judgment uncertainty.

RUNNING TITLE

Memory decision-making processes dilate pupil

AUTHORS

(1) Ravi D. Mill, Center for Molecular and Behavioral Neuroscience, Rutgers University, USA
(2) Akira R. O'Connor, School of Psychology & Neuroscience, University of St Andrews, UK
(3) Ian G. Dobbins, Department of Psychology, Washington University in Saint Louis, USA

CORRESPONDING AUTHOR

Ian G. Dobbins, Department of Psychology, Washington University in Saint Louis, One Brookings Drive, St Louis, MO 63130, USA. Email: idobbins@wustl.edu.

WORD COUNT: 9277

CONFLICT OF INTEREST: None

ABSTRACT

Optimally discriminating familiar from novel stimuli demands a decision-making process informed by prior expectations. Here we demonstrate that pupillary dilation (PD) responses during recognition memory decisions are modulated by expectations, and more specifically, that pupil dilation increases for unexpected compared to expected recognition. Furthermore, multi-level modeling demonstrated that the time course of the dilation during each individual trial contains separable early and late dilation components, with the early amplitude capturing unexpected recognition, and the later trailing slope reflecting general judgment uncertainty or effort. This is the first demonstration that the early dilation response during recognition is dependent upon observer expectations and that separate recognition expectation and judgment uncertainty components are present in the dilation time course of every trial. The findings provide novel insights into adaptive memory-linked orienting mechanisms as well as the general cognitive underpinnings of the pupillary index of autonomic nervous system activity.

KEYWORDS: recognition memory, decision-making, pupillometry, orienting, uncertainty.

1. INTRODUCTION

In addition to the pupillary light reflex, research spanning almost fifty years has established increased pupillary dilation (PD) as a correlate of diverse cognitive demands, including mental arithmetic (Hess & Polt, 1964), working memory (Kahneman & Beatty, 1966), and decision-making spanning perceptual (Kahneman & Beatty, 1967), semantic (Ahern & Beatty, 1981) and economic domains (Fiedler & Glockner, 2012).

The pupil is also sensitive to episodic memory judgments, dilating more for recognition probes identified as studied (*old*) versus unstudied (*new*) (Gardner, Mo, & Borrego, 1974; Goldinger & Papesh, 2012; Heaver & Hutton, 2011; Naber, Frassle, Rutishauser, & Einhauser, 2013; Papesh, Goldinger, & Hout, 2012; Vo et al., 2008). As an extension of earlier work linking pupillary dilation to cognition (Beatty, 1982; Kahneman, 1973), this 'pupil *old/new* effect' has been suggested to reflect the increased 'cognitive load' or voluntary effort required during the successful retrieval of episodic content. Below we briefly outline the Cognitive Load model, highlighting how its previous applications to the pupil old/new effect may have been strained. We then consider an alternative possibility in which the dilation reflects an involuntary response indicating attentional orienting, and conclude by explaining how our reported memory cueing paradigm pits the effort and orienting accounts of the pupil old/new response against one another.

1.1. The Cognitive Load model and its prior application to the pupil old/new effect

According to the Cognitive Load model of Kahneman and Beatty (Beatty, 1982; Kahneman & Beatty, 1966; Kahneman, 1973), pupil dilation is a peripheral marker of arousal that serves to transiently increase cognitive capacity as a result of either the ‘voluntary’ or ‘involuntary’ deployment of attention. When voluntary, the subject chooses to engage in a problem or decision task and the inherent ‘top-down’ demands of the task drive both the level of cognitive effort and pupillary dilation (Kahneman, 1973). For example, in a working memory digit span task, pupil dilation increases as the number of digits to be retained also increases (Kahneman & Beatty, 1966), tracking the greater effort expended for the larger memory set size. Critically, under voluntary attention, increasingly effortful tasks should yield increasingly slowed, erroneous and uncertain responding (Kahneman, 1973). These behavioral indices of voluntary attention/effort overlap with those in response conflict paradigms which also yield increased pupil dilation for conditions of heightened conflict and associated uncertainty; for example, when naming colour-incongruent words in the Stroop task (Laeng, Ørbo, Holmlund, & Miozzo, 2011; Stroop, 1935) and when making left/right button-presses to spatially incompatible locations in the Simon task (Simon, 1969 Steenbergen & Band, 2013).

Prior applications of the Cognitive Load model to the pupil old/new recognition effect have generally interpreted it as a marker of increased voluntary effort during successful retrieval (Goldinger & Papesh, 2012; Papesh et al., 2012; Vo et al., 2008). However, there is little to suggest that hits (correct responses to studied items) are subjectively more ‘effortful’ than correct rejections (correct responses to unstudied items) in the way effort is characterized during the working memory, Stroop and Simon tasks mentioned above. In all of these, increasingly effortful trials are rendered less accurately, more slowly, and often with reduced subjective confidence. Contrary to this characterization, recognition hits are generally rendered more confidently than correct rejections across a range of levels of processing manipulations (Dobbins & Han, 2007; Jaeger, Cox, & Dobbins, 2012). Hits are often also rendered more quickly (Wiese & Daum, 2006) and we replicate this convergent behavioral profile in the Supplementary Information section (see SI, section 1). Furthermore, increasing the depth of encoding also increases the pupil old/new dilation effect (Otero, Weekes, & Hutton, 2011), meaning that as studied materials become *easier* to identify, the size of the PD response to them also increases, which is the opposite of what should occur if the dilation indexed the voluntary expenditure of effort.

Overall, the behavioral characteristics of successful recognition judgments do not fit well with the notion that the pupil old/new effect occurs because hits are more effortful than correct rejections, as conceived by the voluntary component of the Cognitive Load model. However, this model also has an involuntary component that has been neglected in applications of cognitive load theory to pupillometry research in recognition memory. We next consider if this could explain the pupil old/new effect.

1.2. Involuntary Attention and the Orienting Response

The involuntary component of the cognitive load model is closely related to the orienting response (Kahneman, 1973), which is traditionally evoked in ‘bottom-up’ fashion by stimuli that perceptually violate an observer’s predictive model (Sokolov, 1963a, 1963b). The dilation response and other autonomic indices of orienting (such as the P300 event-related potential) have been well documented via the oddball paradigm, in which unexpected stimuli such as unpredictable (and rare) shifts in the intensity or frequency of tone pips (‘oddballs’) interspersed among regularly occurring tones (‘standards’) trigger prominent pupil dilation (Friedman, Hakerem, Sutton & Fleiss, 1973; Hillyard, Squires, Bauer & Lindsay, 1971). This involuntary response signals a rapid and involuntarily increased allocation of resources to the processing of the unexpected stimulus, and is hence not *directly* linked with voluntary ‘effort’ and its behavioral signatures of slowed, erroneous and uncertain judgment (Kahneman, 1973). This raises the alternative that the pupil old/new effect is driven by the involuntary rather than voluntary component of the cognitive load model.

However, while oddballs are often labelled as ‘novel’ in that they are unanticipated by the observer’s prior predictive model, this ‘perceptual novelty’ is fundamentally different from ‘novelty’ during episodic recognition memory. During the former, the stimulus stands out because it violates *perceptual expectations* given recent experiences. In contrast, during typical verbal recognition memory tasks, the unstudied items are not novel in the sense that they violate perceptual or linguistic expectations because both studied and unstudied items are drawn from equally known common words. Rather, unstudied items are *episodically* ‘novel’ because they fail to evoke memories of the current study context, and not because the linguistic features of the items themselves are unanticipated. Hence, the involuntary component of the Cognitive Load model does not seem to afford a straightforward explanation of the pupil old/new effect because in standard recognition memory paradigms there are no perceptual oddballs present. Nonetheless, even if one were to misapply the notion of perceptual novelty orienting to episodically ‘novel’ recognition stimuli, the Cognitive Load model makes an incorrect prediction; namely, greater dilation for correct rejections than hits, which is converse to the actual old/new pupil response evoked during recognition memory.

However, the observation that target oddballs across multiple sensory modalities are capable of eliciting a common neural signature has led to the suggestion that unexpected information in a more general sense (rather than *perceptual* information *per se*) might be the key driver of the orienting response (Corbetta & Shulman, 2002; Downar, Crawley, Mikulis & Davis, 2000). Thus if one expands the notion of the orienting response to encompass orienting to information that is generally unexpected, even when that information is recovered from long term memory, then the involuntary component of the Cognitive Load model may be an appropriate characterization of the pupil old/new response. Under this reconceptualization, the pupil old/new effect would reflect orienting to recovered episodic information and, as with other orienting phenomena, it would be potentiated by the degree this information is unexpected. This conceptualization is consistent with various recognition models inspired by functional neuroimaging that assume a role for bottom-up attention in the processing of unexpected memorial content (Cabeza, Ciaramelli, Olson & Moscovitch, 2008; O'Connor, Han & Dobbins, 2010) and it is also consistent with a recent conceptualization of the pupil dilation response as signalling the surprise value of diagnostic information during economic decision-making (Preuschoff, 't Hart & Einhauser, 2011). This characterization of the recognition pupil response as reflecting an involuntary orienting process has to date neither been considered nor tested.

The possibility that pupil old/new effects reflect orienting phenomena requires manipulating two elements; namely an expectation or prior belief and an information outcome with respect to that belief. It is the difference or distance between the expectation and outcome that regulates the degree of unexpectedness/surprise and hence the strength of involuntary orienting (e.g., Baldi & Itti, 2010). Unfortunately, standard recognition memory tests that evoke the pupil old/new response do not actively control the expectations of the observers. Thus even if episodic information is unexpectedly recovered on some trials (generating a modest dilation response in the trial average), standard recognition paradigms have no way of establishing when retrieval outcomes are more versus less unexpected because they do not manipulate memorial expectations at the level of individual trials. Here we use explicit memory cueing to do so and below we explain how this paradigm sets up competing predictions for voluntary and involuntary attentional accounts of the pupil old/new effect (see Supplementary Information section 3 for further discussion on how uncontrolled expectations operating in standard recognition paradigms might account for previous old/new effects).

1.3. Competing predictions for the old/new effect set up by explicit memory cueing

We collected pupillometry data during recognition using an Explicit Memory Cueing paradigm developed by O'Connor, Han & Dobbins (2010). In the cued phase of the paradigm, cues or 'hints' which are known to be 70% valid precede each recognition memory probe ('Likely Old' or 'Likely New'). In the uncued phase, uninformative cues are provided (Likely ???) in an analogue of typical recognition tests in which the observers' expectations go uncontrolled (see Figure 1). Thus the Explicit Memory Cueing paradigm instils expectations on every trial that can either converge or diverge with the subsequent recovered memory signal, which enables the testing of hypotheses of the functional underpinnings of the recognition old/new response that follow from the discussion above.

Recognition dilation as voluntary effort. If prior pupil old/new effects reflect the greater voluntary effort required to correctly conclude a recognition probe is old, then any manipulation that *generally* increases effort expended during recognition judgments should also increase the dilation response. Hence, one should see increased dilation during correct responses to both invalidly cued new materials and invalidly cued old materials because in both cases judgment uncertainty is heightened compared to the valid cue conditions, and thus the effort required to reach a correct conclusion will also increase. This is because invalid cueing for both old and new materials has been shown to adversely affect judgment accuracy, speed and confidence, given that it sets up a conflict between recovered memorial signals and expectations (Jaeger et al., 2012; Jaeger, Konkel, & Dobbins, 2013; Konkel, Selmeczy, & Dobbins, 2015; O'Connor et al., 2010; Selmeczy & Dobbins, 2012). Increased dilation for both invalid cue types would converge with demonstrations of increased dilation during Stroop (Laeng et al., 2011) and Simon conflict manipulations (Steenbergen & Band, 2013), and in turn would support the idea that the previously reported pupil old/new effect indeed reflected the amount of volitional effort expended, which just happens to be larger on average for old versus new recognition conclusions in standard, uncued recognition paradigms.

Recognition dilation as involuntary orienting. If the pupil old/new effect instead reflects involuntary orienting triggered by unexpected or surprising episodic information then the dilation should be modulated by observer expectations in a more *selective* way, such that invalidly cued hits should yield a prominent dilation response (due to the recovery of episodic content that is unexpected), whereas validly cued hits should yield a minimal response (because episodic content is expected). In contrast, the effect of invalid versus valid cueing on correct rejections should be more muted or perhaps absent altogether. This is because the involuntary orienting framework assumes that the response is triggered by the unexpected recovery of episodic content. Since new materials will infrequently, if ever, trigger such retrieval they will not trigger surprise or orienting even when they follow invalid cues. We expand upon this point when discussing dual process models of recognition in the general discussion.

To preview our findings, analyses of the trial-averaged dilation response (the standard method used to recover the pupil old/new effect) revealed a *highly specific* dilation response for ‘old’ recognition judgments that were unexpected, and a complete elimination of this response for ‘old’ recognition judgments that were expected; there was no analogous modulation of ‘new’ recognition judgments by expectations. The specificity of this response favors a link between previous old/new effects and the involuntary orienting rather than voluntary effort component of the Cognitive Load model. Critically, we also used a novel multi-level modeling approach to predict behavior at the level of individual trials and identified separate morphological markers of voluntary effort and involuntary orienting within each pupil dilation response that were obscured in the trial-averaged data. An ‘early amplitude’ component captured the involuntary orienting response observable in the trial-averaged data (which we term an ‘unexpected recognition’ response). A later ‘trailing slope’ component of the dilation time course was instead reliably tied to judgment uncertainty or effort regardless of the conclusion reached by the observers (‘old’ or ‘new’), and regardless of their cued expectations. This late response is a valid candidate for indexing voluntary effort and is also consistent with prior non-memory findings linking general judgment conflict and uncertainty to increased pupil dilation (e.g. Laeng et al., 2011). Overall, these findings demonstrate that there are two separable functional components within the dilation response at the level of individual pupillometry trials, which are linked to distinct voluntary effort and involuntary orienting processes.

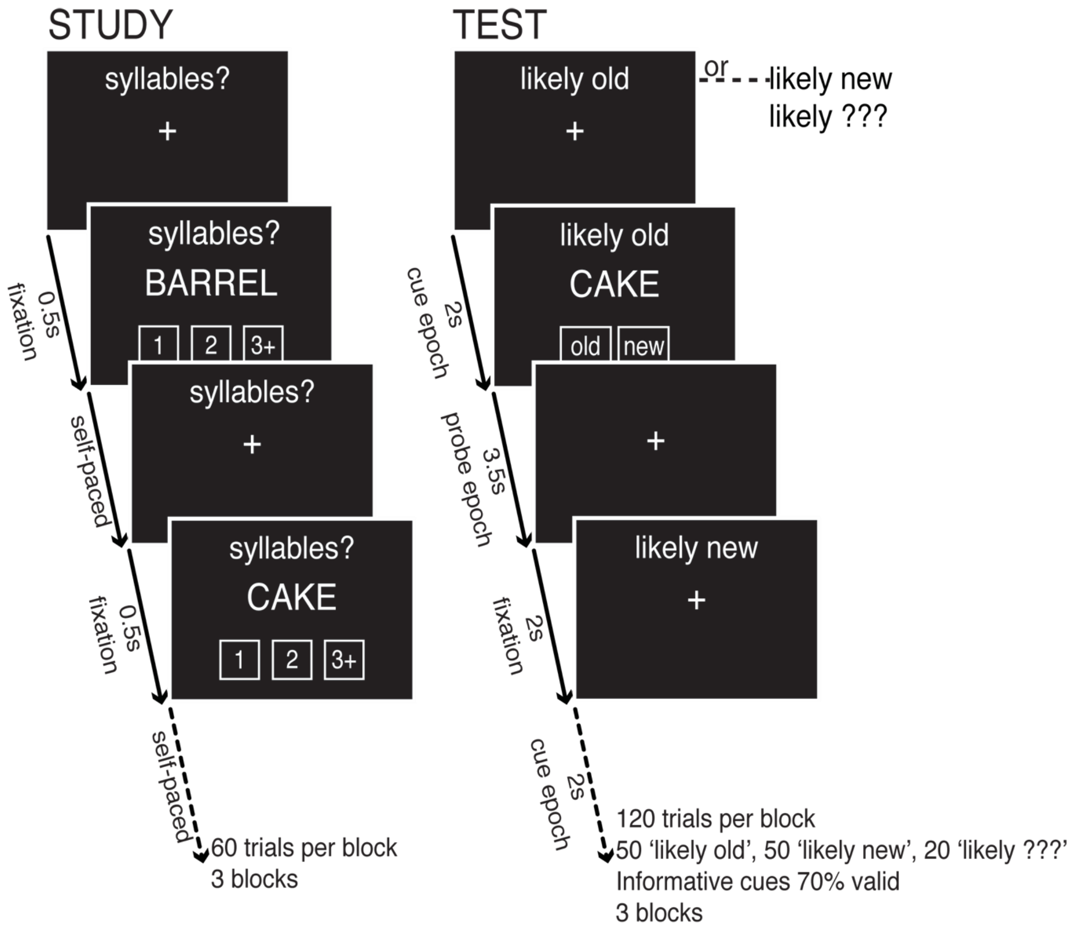


Figure 1. Design schematic for explicit memory cueing paradigm. Study comprised a syllable counting task for 60 presented words. Test required subjects to decide if presented words were studied i.e. "old" or unstudied i.e. "new", for a wordlist comprising 60 *old* and 60 *new* words. Each test trial was preceded by appearance of anticipatory cues of the form "Likely Old" and "Likely New", which were valid in predicting the status of ensuing test items on 70% of trials. Uninformative cues of the type "Likely ???" were also presented on a subset of trials in the uncued condition. Trial timings are in seconds (s).

2. METHODS

2.1. SUBJECTS. 40 young adult subjects were recruited via flyers and subject pools maintained at Washington University in St Louis. This sample size was fixed with reference to prior pupillometry studies involving recognition memory (e.g. Naber et al., 2013; Otero et al., 2011). All subjects had normal or corrected-to-normal vision and abstained from caffeine in the hour immediately preceding their participation. Of these, 3 subjects were excluded for artifacts contaminating more than 20% of their eye-tracking data, 2 for falling asleep during the experiment, and 1 due to a computer malfunction, yielding a total of 34 subjects (22 female; mean age = 24.8 years, range = 18-37 years) for analyses. Informed consent was obtained in compliance with Washington University's human subjects guidelines, with compensation at a rate of $10/hour.

2.2. STIMULI. Stimuli consisted of words sampled from a pool of 1216 (Kucera-Francis corpus frequency = 8.85), yielding five, 360 item wordlists to which subjects were randomly assigned. Word presentation order within these five wordlists was randomised across subjects. Word length was controlled by excluding words less than four letters and greater than 10 letters in length, so as to minimise confounding luminance differences between presented words. Each subject completed three study-test blocks, with 60 words presented at study and 120 words presented at test (60 studied *old* words and 60 unstudied *new* words).

2.3. PROCEDURE. The experiment was conducted on a standard PC running E-Prime software (Psychology Software Tools, Pittsburgh, PA, USA) that synchronized with both the subject display PC and a dedicated PC running the eye-tracker recording software. Following the full setup of the eye-tracker (detailed below), subjects were presented with onscreen instructions and a brief practice phase. Encoding was governed by providing a self-paced syllable counting task, with response options ‘1’, ‘2’ and ‘3+’. This was succeeded immediately by the test phase, wherein subjects were asked to decide if words appearing onscreen were studied (*old*) or unstudied (*new*) for a randomized list containing equal proportions of *old* and *new* items. During each test phase, subjects were also provided with anticipatory cues as to the likely old or new status of ensuing items. These took the form of “Likely Old” (LO) and “Likely New” (LN) hints that were each valid on 70% of trials. Subjects were informed of the cues’ validity and encouraged to use them to guide their recognition judgments. On a subset of trials, subjects were provided non-informative cues of the form “Likely ???” (hereafter referred to as “uncued” trials, UC) that assessed behavioral and eye-tracking responses in the absence of explicit cues. All cues appeared in isolation for 2 s at the start of each trial (‘cue-only period’), followed by appearance of both the test item and a response prompt for 3.5 s (‘cue+probe’), which was followed by a 2 s fixation cross (See Figure 1). Hence, each test phase comprised a total of 120 cued test trials, of which 70 were validly cued, 30 invalidly cued and 20 uncued. Subjects completed three study-test block runs in total.

2.4. EYE-TRACKING DATA COLLECTION. Pupil and eye movement data were recorded using an Eyelink 1000 infrared eye-tracker (SR Research Ltd., Mississauga, ON, Canada) running Eyelink software (v 4.48), sampling at 1000 Hz and at spatial resolution <0.01° RMS. The camera was placed centrally under the presentation monitor at a fixed distance 60 cm from the subject. Subjects were seated on a comfortable chair that included an adjustable headrest to minimize head movement. The experiment was displayed in a room maintained at a constant low level of ambient illumination. Task stimuli were presented in white font on a black background in size 35pt font (the same for cues, probes and response prompts). Calibration and validation of gaze direction was conducted prior to each subject's experimental session. Pupil and eye movement data were acquired at study and test across all three blocks, with subjects instructed to maintain fixation throughout (as indicated by a central '+' fixation cross at the beginning of each trial). The data was pre-processed using in-house software written in Java (Oracle Corporation, Redwood Shores, CA, USA), which corrected for blinks and other sources of dropout by linear interpolation. Subjects with signal dropout in excess of 20% were excluded from the analysis.

The data were subsequently down-sampled to 20 Hz to decrease file sizes (by averaging across neighbouring timepoints), which was of particular concern considering the computationally intensive point-wise bootstrapping procedures adopted. Critically, as highlighted in prior literature (e.g. Nowak, Hachol & Kasprzak, 2008; Wang, Boehnke, Itti & Munoz, 2014), PD responses linked to perceptual orienting and cognitive phenomena occur in the 0.75-3Hz frequency range, and are hence a fairly slow temporal phenomenon. Thus the present down-sampled rate is approximately 20 times the periodicity of any known PD signal of interest, and is therefore sufficient for capturing evoked responses. This assumption was confirmed by the well-captured perceptual amplitude modulation in the trial-averaged PD response plots presented in the results, which is the pupil effect with the fastest latency currently known, taking the stereotyped form of an initial small dilation (bump) after onset of the visual cue, followed by a drastic constriction and return to baseline (see Figure 2; see Wang et al., 2014). Furthermore, extracting the same trial-averaged plots at a higher sampling rate of 100Hz yielded an identical PD response morphology as the 20 Hz downsampled plots (not shown). Prior to analysis, the PD response data was transformed into a percent signal change response by normalizing with respect to the 200 ms epoch immediately prior to the trial onset. Separate analyses were conducted for the mean trial-locked (i.e. locked to cue onset) and response-locked pupil dilation (see Results for further details).

3. RESULTS

3.1. BEHAVIOR

Discriminability, as measured by *d′* did not differ reliably across the three cueing conditions via one-way ANOVA (*F* < 1), although it was well above floor (*d′*s of 1.49, 1.48, 1.50 for LN, LO, & UC respectively)[1]. In contrast, decision bias, as indexed by *C*, differed considerably and in the appropriate direction given the cues ($F(2,64)=37.35$, $p\eta^2=.58$, $p<.001$). Subjects were most conservative following the LN cues (0.27) followed by the UC condition (0.01) and then the LO condition (-0.25), with all three differing reliably according to Tukey's HSD. Consistent with prior work using the Explicit Memory Cueing procedure, discriminability was unaffected by cue condition whereas decision bias was heavily influenced.

We also analysed reaction time via 2-way repeated measures ANOVAs with factors of Response ("old" and "new") and Cue (LO and LN), conducted separately for correct and incorrect trials (see Table 1 for RT means and standard errors). The 2x2 ANOVA on correct trials revealed

---

[1] One further subject who made no false alarms in one cueing condition was removed from the *d′* and *C* analyses, as these measures could not be calculated for that condition. This subject was excluded from all subsequent analyses involving incorrect responses.

significant main effects of response (F(33)=18.43 $p\eta^2$=.36, p<.001) and cue (F(33)=4.88, $p\eta^2$=.13, p=.034), which were tempered by a response by cue interaction (F(33)=48.79, $p\eta^2$=.60, p<.001). The interaction reflected an increase in RT when subjects made a correct response that was incongruent with the available cue, with Tukey's HSD revealing this effect to be significant for LO cues (CRs > Hits, p<.001) but not LN cues (Hits > CRs, p=.690). The same two-way repeated measures ANOVA on cued incorrect trials also revealed a significant response by cue interaction (F(32)=6.41, $p\eta^2$=.17, p=.016), this time in the absence of any main effect of response (F<1) or cue (F(32)=2.39, $p\eta^2$=.07, p=.132). Again, this interaction reflected increased RT for incorrect responses made in opposition to the available cue, with the effects found to be individually non-significant for the LO cue (Misses > FAs, p=.347) and LN cue (FAs > Misses, p=.145) conditions via Tukey's HSD. These findings suggest that group-averaged reaction time broadly indexes cue-incongruent responding, which engenders a slower, more effortful/uncertain decision process. The slowing of RT under cue-incongruent responding is consistent with prior cueing manipulations applied in recognition (e.g. O'Connor et al., 2010) and non-recognition domains (e.g. Posner, Synder & Davidson, 1980). Critically, these analyses demonstrate that both LO and LN cue conditions generated similar levels of response conflict, and this effect will be employed to make further functional inferences about the evoked pupil dilation response in the later multi-level modeling section.

Table 1. Reaction time in milliseconds across cueing (Likely Old and Likely New) and response conditions (“Old” and “New”). Standard errors are provided in parentheses.

	Correct responses		Incorrect responses	
	Likely Old	Likely New	Likely Old	Likely New
“Old” response	1333.06	1443.93	1700.41	1742.35
	(41.13)	(40.77)	(59.62)	(62.50)
“New” response	1613.17	1424.27	1780.07	1624.80
	(51.68)	(45.71)	(63.02)	(60.67)

3.2. MEAN TRIAL-LOCKED PUPILLARY DILATION

We restrict our main trial-averaged analyses to cued recognition in which the subjects' expectations are actively controlled with respect to the upcoming memoranda. The data from uncued recognition trials are considered in the supplementary section, which also examines seven eye-movement measures to rule out any confounds in the interpretation of the PD effects (see SI, sections 3 and 4 respectively). Figure 2 illustrates the trial-locked mean PD response for correct trials (upper panels) and incorrect trials (lower panels) across the Likely Old and Likely New cue conditions (left and right columns), with both horizontal and vertical markers delimiting key periods within the trial (viz., pre-trial baseline, cue-only period, and cue+probe period). As alluded to in the methods section, all four panels demonstrate that the appearance of the cue was followed by a highly stereotyped perceptual dilation response, consisting of an initial small dilation, marked constriction, and then a return towards baseline (as indicated in Figure 2 between the dashed lines; Wang et al., 2014). Following the appearance of the recognition probe (cue+probe), the dilation timecourse begins to differentiate the cue by response combinations after approximately 750-1000ms. The dissociation that occurs in this cue+probe period is easy to summarize – there is no discernible difference in pupil size for "old" and "new" judgments, regardless of their correctness, under the LO cue (left columns) when subjects expect materials to be recognized. In other words, *the pupil old/new effect is eliminated when observers are cued to expect episodic information from the environment.* In contrast, there is a robust old/new dilation effect under the LN cueing condition, with a greater response for "old" than "new" judgments (right columns) when subjects are biased to expect new materials. This differential response occurs regardless of whether the “old” judgment is correct (i.e. a “hit”) or incorrect (i.e. a “false alarm”, FA), and lasts throughout the duration of the trial.

To confirm this impression we considered the mean PD response during the probe epoch. Focusing on correct responses (top panels in Figure 2), a two-way ANOVA with factors of Response (“old” or “new”) and Cue (LO and LN) demonstrated main effects of Response ($F(1,33) = 10.39$, $p\eta^2=.24$, $p = .003$) and Cue ($F(1,33) = 5.14$,$p\eta^2=.13$, $p = .030$) that were conditioned by a Response X Cue interaction ($F(1,33) = 8.21$, $p\eta^2=.20$, $p = .007$). The interaction resulted because there was a reliable increase in PD for Hits compared to CRs under the LN cue ($p =.001$), but not the LO cue ($p \sim 1$) via Tukey’s HSD, consistent with Figure 2. Turning to incorrect responding (Figure 2, bottom), the two-way repeated measures ANOVA yielded only an interaction of Response X Cue ($F(1,32) = 6.09$, $p\eta^2 = .16$, $p = .019$), which again occurred because the PD for incorrect “old” responses (FAs) was greater than that for “new” responses (misses) under the LN cue ($p = .023$), but not the LO cue ($p > .95$) via Tukey’s HSD.

These findings demonstrate a highly selective PD response that isolates “old” from “new” judgments, regardless of correctness, *but only when recognition is not expected in the environment*. Given that the PD response follows the observers’ subjective conclusions and not the actual item status, it does not signal veridical episodic retrieval (i.e. retrieval success) but instead the subjects’ perceived sense of recognition. As it fails to occur following the LO cue, but does so following the LN cue, it is a marker of orienting to unexpected episodic information in the environment. Hence we refer to the increased diameter of the pupil in the right panels of Figure 2 as the ‘unexpected recognition’ response.

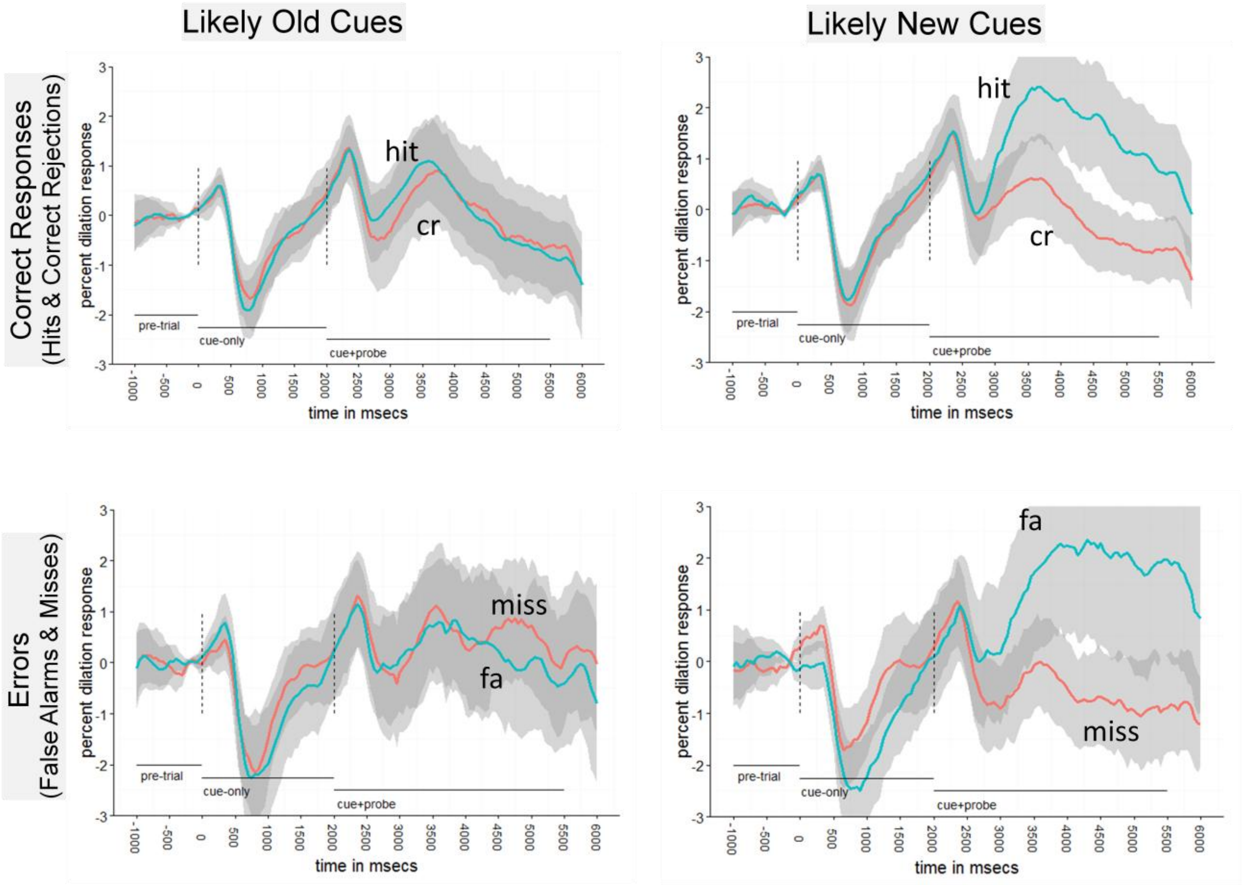


Figure 2. Trial-locked pupillary dilation (PD) across cued trial types (left column = “Likely Old” cue, right column = “Likely New”) and response accuracy (upper panel = correct responses, lower panel = errors). For all plots, blue lines indicate PD response during “old” judgments and red lines indicate PD during “new” judgments. The timing of events within each trial (pre-trial, cue-only and cue+probe) is depicted with solid horizontal lines and vertical dashed lines. Light grey areas denote the bootstrapped 95% confidence interval of each response across subjects, established at each time point using 1000 replications. ‘Hit’ = correct ‘old’ response; ‘cr’ = correct rejection or correct ‘new’ response, ‘fa’ = false alarm or incorrect ‘old’ response; ‘miss’ = incorrect ‘new’ response.

Figure 2 plots the data for ‘old’ and ‘new’ judgments under fixed cue conditions (i.e. separately for LO and LN cues), to demonstrate that the pupil old/new effect is heavily contingent upon observer expectations. Figure 3 provides another way to look at the phenomenon, by plotting the PD response for invalid and valid cueing under fixed response conditions (i.e. separately for hits and correct rejections). The left panel demonstrates that the dilation response during correct rejections does not depend upon cue validity. This is not because the cueing was ineffective in instilling expectations because ‘new’ judgments slowed considerably under invalid LO cues (see behavioral RT results in Table 1) and they were significantly less accurate (‘new’ decision accuracy: LN = .83, LO = .69, $p < .001$). In contrast, the right panel depicts an increased dilation response for invalidly versus validly cued hits and thus it is only when information specifically supporting hits is unexpectedly recovered that the increased PD occurs. Critically, the behavioral results also highlighted that RT for hits slowed during invalid versus valid cueing (see Table 1) and responses were considerably less accurate (‘old’ decision accuracy: LO = .81, LN = .67, $p < .001$). As slowed, less accurate performance occurred under invalid cueing for both hits and correct rejections, yet increased pupil dilation occurred only for invalidly cued hits, the increased effort/uncertainty of judgments made during invalid cueing cannot be the cause of the observed mean PD response.

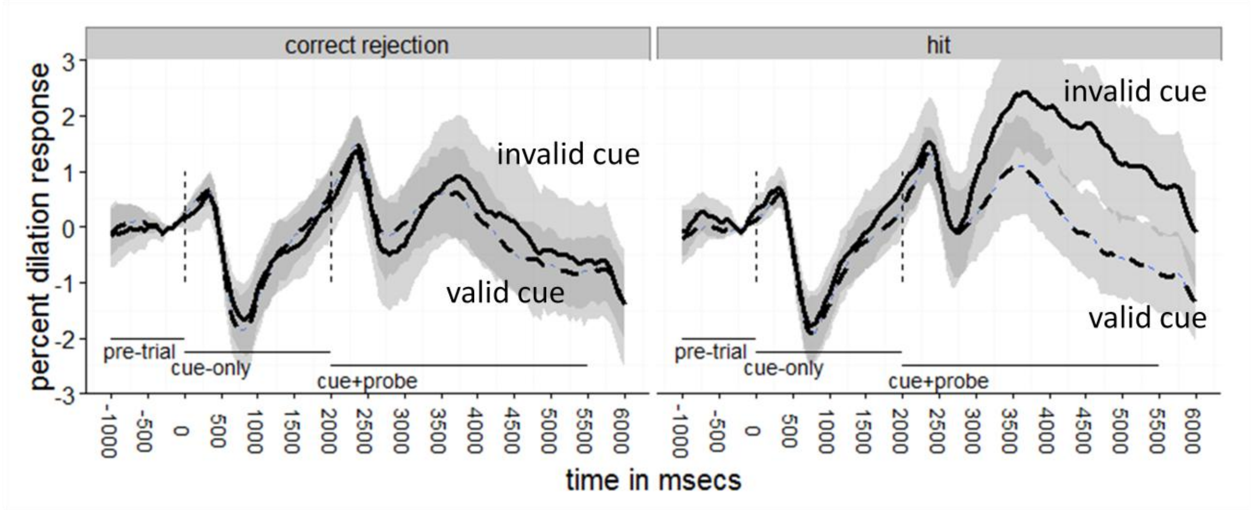


Figure 3. Dilation response for hits and correct rejections as a function of whether the judgment followed valid (dashed) or invalid (solid) preparatory cues. The timing of events within each trial (pre-trial, cue-only and cue+probe) is denoted with solid horizontal lines and vertical dashed lines. Light grey areas denote the bootstrapped 95% confidence interval of each response across subjects, established at each time point using 1000 replications.

Overall, Figures 2 and 3 show that the most prominent PD effect in the trial-averaged data, which begins about 1000ms following onset of the recognition memory probe, is consistent with a form of involuntary orienting and not voluntary effort, and more specifically reflects a response tied to the recovery of unexpected episodic information. Below we consider whether this unexpected recognition dilation precedes overt responding, which would corroborate that it is tied to the recovery of evidence that precedes the response.

3.3. MEAN RESPONSE-LOCKED PUPILLARY DILATION

Figure 4 plots pupil dilation data under the Likely New cue time-locked to subjects' response times. The figure captures the 1500 ms before and after subjects responded, and clearly shows that the unexpected recognition PD precedes responding. Given that the PD occurs regardless of whether the "old" report is correct or not (see Figure 2), we collapsed the data across correct and incorrect responses to increase power. Following this, we contrasted the "old" and "new" PD responses for all of the 61 time points collected at 50ms intervals. To control family-wise error the p-values were adjusted using the false discovery rate procedure (Benjamini & Hochberg, 1995). The first reliable difference in pupillary dilation amplitude occurred 450 ms before the overt response (see Figure 5). We also examined the slope of the dilation time course prior to the subject's response. For each subject we fit a separate simple linear regression of the dilation response as a function of time, yielding a separate slope for each subject's "old" and "new" dilation responses up to the point of responding. These were then contrasted via a paired t-test, which demonstrated a reliably more positive slope for "old" than “new” decision trials prior to the execution of an overt response ($t(33) = 4.33$, $p < .001$). Additionally, whereas the average slope for “old” trials differed from zero ($t(33) = 3.00$, $p = .005$), it did not for “new” trials ($t<1$). Collectively, these data demonstrate that the unexpected recognition PD precedes the subjects’ overt responses by almost half a second, consistent with a role in indexing memory decision-making processes that precede response commission.

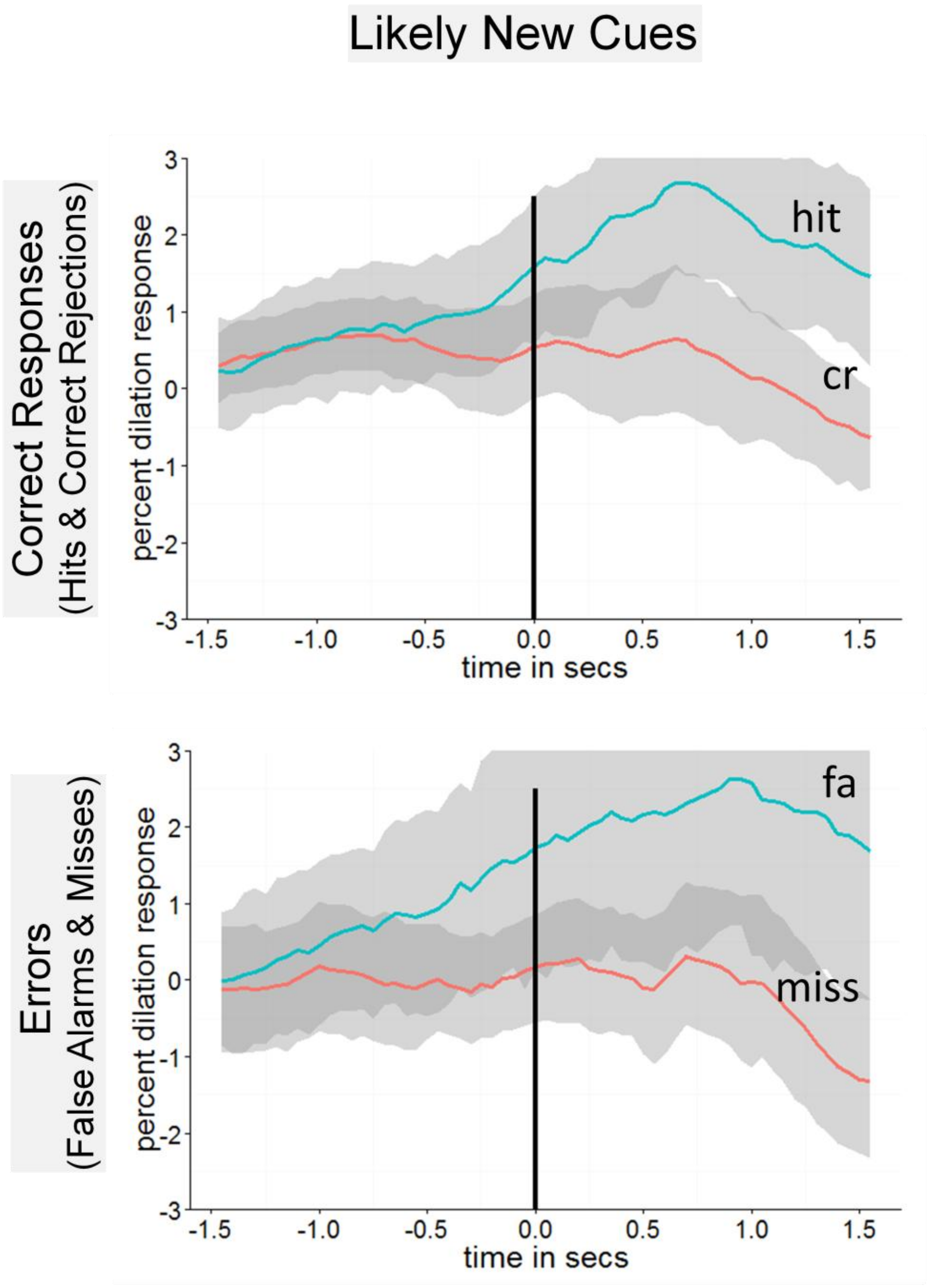


Figure 4. Response-locked pupillary dilation under the Likely New cue for correct responses (upper panel) and errors (lower panel). Blue lines indicate PD response during "old" judgments whereas red lines indicate PD during "new" judgments. The vertical line marks the response of the subjects. Light grey areas denote the bootstrapped 95% confidence interval of each response across subjects, established at each time point using 1000 replications.

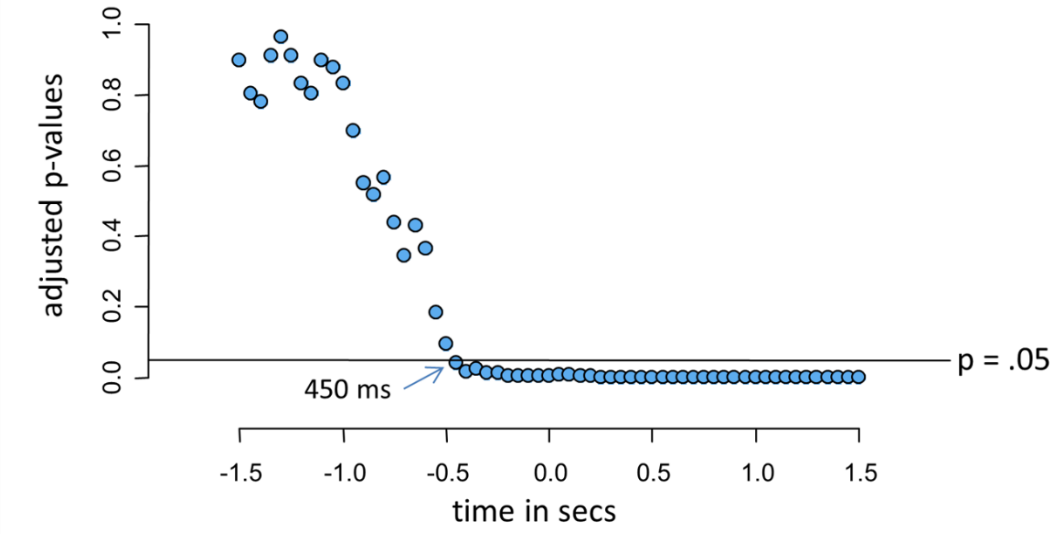


Figure 5. False Discovery Rate (FDR) adjusted p-values comparing response-locked pupil dilation for “old” versus “new” responses under the Likely New cue condition. Subjects’ responses onset at 0 seconds on the x axis and the horizontal line denotes the significance threshold on the y axis of *p* = .05.

3.4. MULTI-LEVEL MODELING OF TRIALWISE PUPILLARY DILATION

Figure 6 contrasts the PD for correct and incorrect “old” reports under the Likely New cueing condition, so as to illustrate the rationale for our ensuing decomposition of the trial-wise pupillary responses. Although both correct and incorrect “old” judgments show a prominent early dilation response at about 1.5 s into the probe period (base of grey arrows; see also Figure 2 right panels), they appear to differ in the rate at which the dilation response returned to baseline; a characteristic we refer to as the ‘trailing slope’. During correct responses the early dilation quickly returned toward baseline, whereas during incorrect responding the pupil remained dilated for a longer duration. Aside from the difference in accuracy across these trials (errors versus correct responses), they also differed appreciably in reaction times, such that RT under the LN cue was slower for incorrect “old” compared to correct “old” judgments (t(33) = 3.19, p =.003; see also Table 1). Considered alongside the typical finding that false alarms are rendered much less confidently than hits (e.g. Ratcliff & Murdock, 1976), the overall pattern suggests that the slope of the dilation response may index judgment effort or uncertainty, such that a sustained dilation (viz., a more positive slope) accompanies a slower, more uncertain judgment whereas a quicker return towards baseline (viz., a more negative slope) indicates a confident, less uncertain judgment. If correct, *then the early amplitude of the response and its trailing slope convey different cognitive information*, with the former signalling unexpected recognition and the latter the level of judgment effort/uncertainty. Moreover, if the trailing slope component of the dilation response is sensitive to judgment effort/uncertainty, it should also track uncertainty during the conditions not illustrated in Figure 6, since varying degrees of uncertainty should be present within all of the cueing conditions studied and not only the LN condition that gives rise to the unexpected recognition effect.

To test this hypothesis, we obtained the early amplitude and trailing slope values for each individual pupillary dilation response across all the conditions in Figure 2 by fitting a simple linear regression to the PD response beginning at 1450 ms into the probe period and ending at the offset of the probe. This yields an early amplitude estimate (intercept of the regression) and trailing slope estimate (slope of the time predictor) for each trial for each subject. These components were then directly pitted against one another in a multilevel model (MLM), to contrast their ability to predict three separate trial-wise behavioral outcomes within each cueing condition; namely, reaction time, accuracy, and old/new response. The use of an MLM allowed us to jointly model subject- and trial-level variation in these three dependent variables, treating the former as a random effect as implemented in the lme4 and lmerTest packages in the R statistical language (R Project for Statistical Computing, http://www.r-project.org/). Additionally, because we modeled at the level of each individual dilation response, any findings linking components of the time course of the dilation response to behavior cannot be the result of averaging artefacts; instead, they would indicate that the unfolding morphology of the dilation response on individual trials contains separable psychological components, which are differentially predictive of three selected behavioral measures (as outlined in the ensuing sections). Figure 7 illustrates the distributions of trial-wise early amplitude and trailing slope components used in the MLMs. Critically, it demonstrates that the trailing slopes are not always negative (as would be suggested by the trial-averaged plots in Figure 6), which means that on some individual trials there is an increase in pupil diameter during the probe phase.

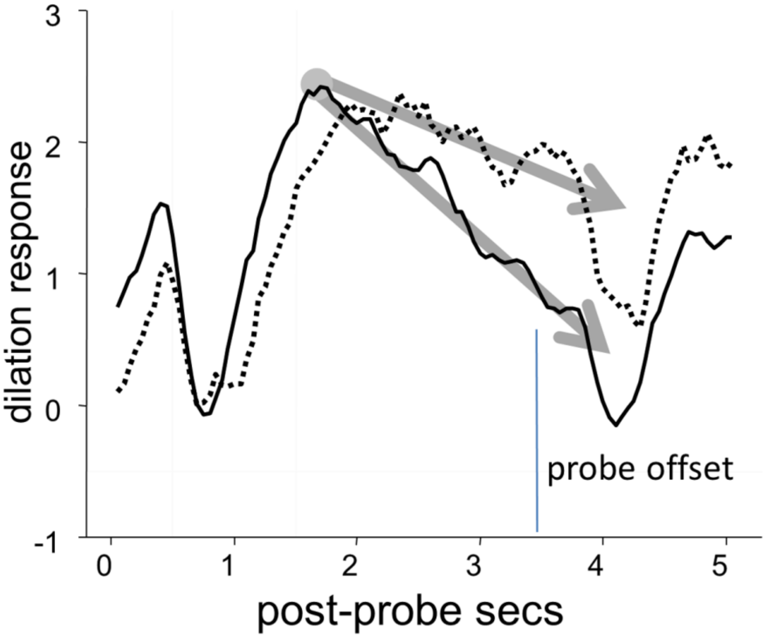


Figure 6. Comparison of correct (solid) and incorrect (dashed) “old” dilation responses under the Likely New cue condition, locked to the onset of the probe at 0 seconds. The responses appear to differ in the rate at which the initial dilation, triggered by unexpectedly recognised stimuli, returns to baseline. If such shape differences were preserved in the individual trials of subjects, with errors reflecting more positive-going slopes, then the early amplitude and trailing slope of each response should predict different aspects of behavior at the trial level.

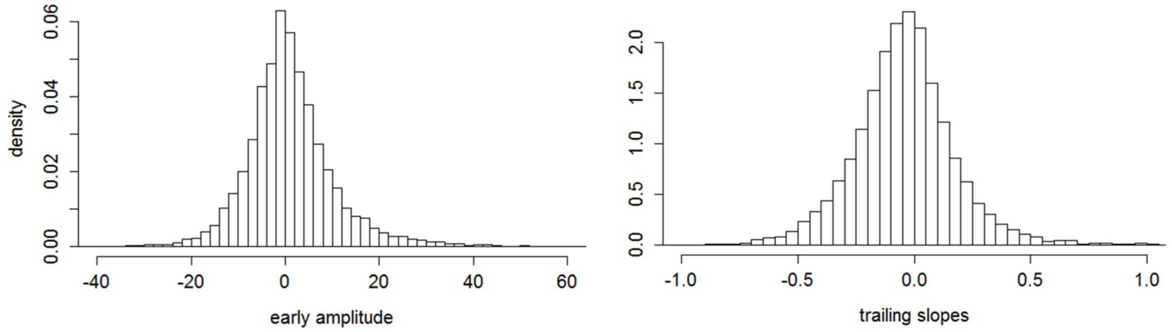


Figure 7. Distribution of trial-wise "early amplitude" and "trailing slope" values obtained from simple linear decomposition of the pupillary dilation response into regression intercept (i.e. amplitude) and slope components respectively.

Table 2 shows the results of all MLM analyses. Each pair of columns from left to right reflect the three different behavioral outcomes examined; namely Reaction Time, Accuracy (1=correct, 0=error) and Response (1='old',0='new'). For each outcome a separate model was fit for the 'Likely Old' and 'Likely New' cueing conditions, leading to six models in total, enumerated within the third row of the table (2 cueing conditions across three behavioral indices). To reiterate, within each model, the early amplitude and trailing slope components were entered as competing predictors of the behavioral outcomes at the level of each trial within each subject, with the subject variable treated as a random effect. For each predictor, the table lists the t-value of its contribution and associated standard error (in parentheses).

Beginning with the reaction time MLMs, an increasing slope was associated with slowed responding during both LO and LN cue conditions (columns 1 and 2 of Table 2), whereas the early amplitude of the dilation response played no appreciable role in predicting reaction times. The link between the slope of the pupil dilation response and the reaction times of the observers is robust, as demonstrated by t-values in excess of 10 for both the LO and LN cueing conditions. To help visualize this relationship, Figure 8 plots the trial-averaged pupil dilation responses as a function of the upper and lower quartiles of reaction time (i.e. 'slow' and 'fast' RT trials) separately for the two cueing conditions (collapsed across old/new and correct/incorrect responses). The left panel depicts the dilation responses for the LO cueing condition, which is the condition where there was no modulation of the trial-averaged dilation response as a function of whether the items were judged old or new (see left panels Figure 2). Thus the observed variation in the dilation response in this condition cannot be linked to whether the subject's recognition judgment violated or confirmed the expectation of episodic recovery instilled by the LO cue. Figure 8 shows that the quickest responses (the solid dilation time-courses) rapidly return to baseline during the probe period and hence have negative slopes. In contrast, the slowest quartile of reaction times (the dashed time-course) demonstrates a sustained dilation throughout the cue+probe period and thus has a more positive going slope. The same pattern recurs when the dilation responses are considered during the LN cueing condition (Figure 8 right panel). This demonstrates that the duration of recognition memory decisions influences the dilation response, independent of whether observers classify memoranda as old or new and independent of whether those classifications are expected or unexpected**.** Furthermore, the fact that this phenomenon replicates across the LO and LN cue conditions, whereas the unexpected recognition response is entirely restricted to the LN cueing condition (see Figure 2), underscores the fact that *there are two functionally separable components operating during the dilation response on each trial*. Next we consider the remaining MLMs to further buttress this point.

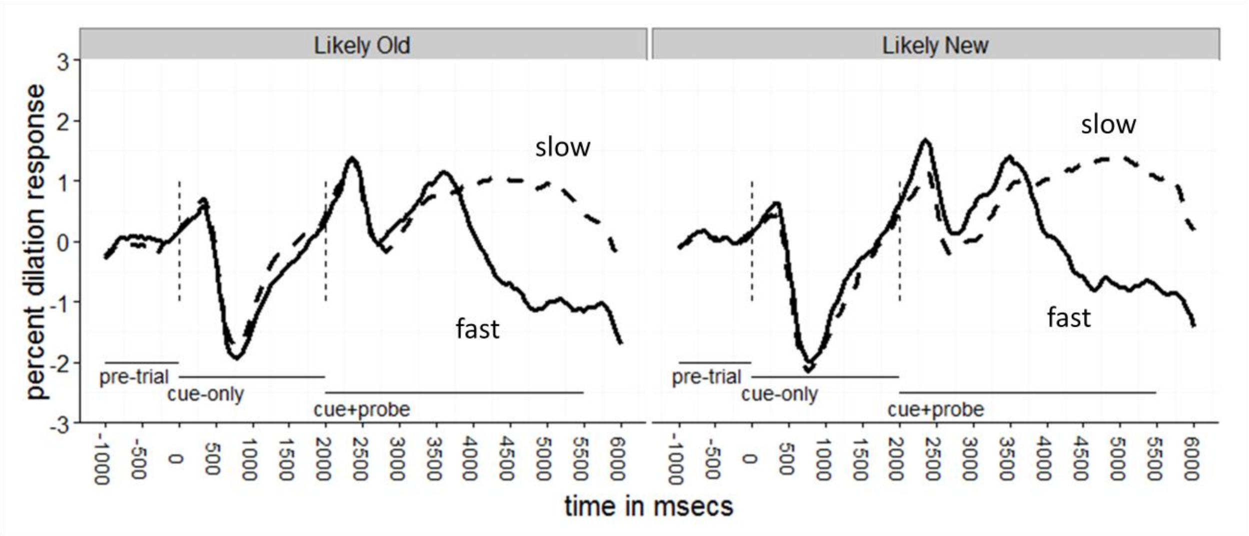


Figure 8. Averaged dilation responses broken down by the slowest (dashed lines) and fastest (solid) reaction time quartiles under the two cueing conditions. Data is collapsed across response type (old/new) and accuracy (correct/erroneous).

Turning to the MLM of recognition decision accuracy, an increasing slope was associated with incorrect responding under both cue conditions, with the negative sign of the t-value reflecting the coding of correct responses as 1 and errors as 0 (columns 3 and 4 in Table 2). This finding corroborates the link between PD slope and judgment effort/uncertainty outlined in the preceding RT MLM, as errors are associated with more uncertainty than correct responses. As with the reaction time MLM analysis, the early amplitude of the dilation response was again unrelated to this measure of effort/uncertainty, showing no relationship with the accuracy of the recognition memory judgments. Figure 9 visualizes the increased (i.e. more positive) slope with erroneous compared to correct responding across both cue conditions.

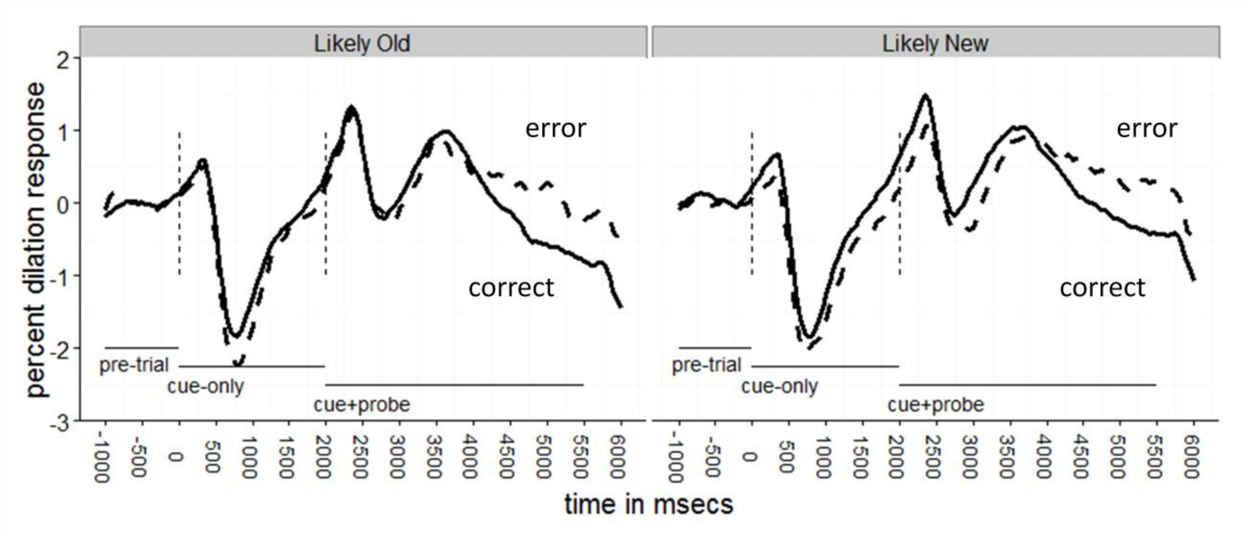


Figure 9. Averaged dilation responses broken down by erroneous (dashed lines) versus correct (solid) judgments under the two cueing conditions. Data is collapsed across response type (old/new).

Finally, we consider whether either the early amplitude or slope components can predict the individual responses of the subjects ('old'=1 or 'new'=0). In this analysis, the slope component was a significant predictor of responses under both cue conditions. However, the sign of the predictor changed depending upon the cueing condition, being negative for the LO cueing condition and positive for the LN (columns 5 and 6, Table 2). Thus increases in slope are predictive of the response option that is incongruent with the cued expectations (i.e. when responding ‘new’ for LO cues and ‘old’ for LN cues) which is why the sign changes across the two models. This reversal should occur if PD slope is a marker of effort/uncertainty because it is established that subjects respond more effortfully when their recognition judgment conflicts with the cued expectation than when it confirms it (as demonstrated by the behavioral accuracy and RT analysis, Table 1). These slope differences can be seen in Figure 10, although they are not numerically large in these averaged time-courses. This is commensurate with the comparatively moderate t-values accompanying the slope predictor in the MLMs of response types, which nonetheless are statistically reliable at the individual trial level (Table 2). In the left panel for the LO cue, the slope is more positive going for ‘new’ responses than ‘old’ responses. In contrast, in the right panel depicting the LN cue condition, the slope is more positive going for ‘old’ responses than ‘new’ responses; a difference that is difficult to see due to the large response effect in this particular cueing condition evoked in the early portion of the dilation time-course.

Indeed, turning to the early amplitude component, there is no reliable relationship with responding under the LO cue (column 5) but a robust effect under the LN cue (column 6). As shown in Figure 10, there is a marked increase in the dilation response for ‘old’ relative to ‘new’ judgments in the LN condition that begins early in the trial (and which makes the subtle trailing slope differences difficult to see). This matches the trial-averaged patterns in the right panels of Figure 2 and supports the interpretation that the early amplitude of the dilation response is a highly selective marker of encountering unexpected evidence supporting conclusions of oldness (viz. an unexpected recognition response).

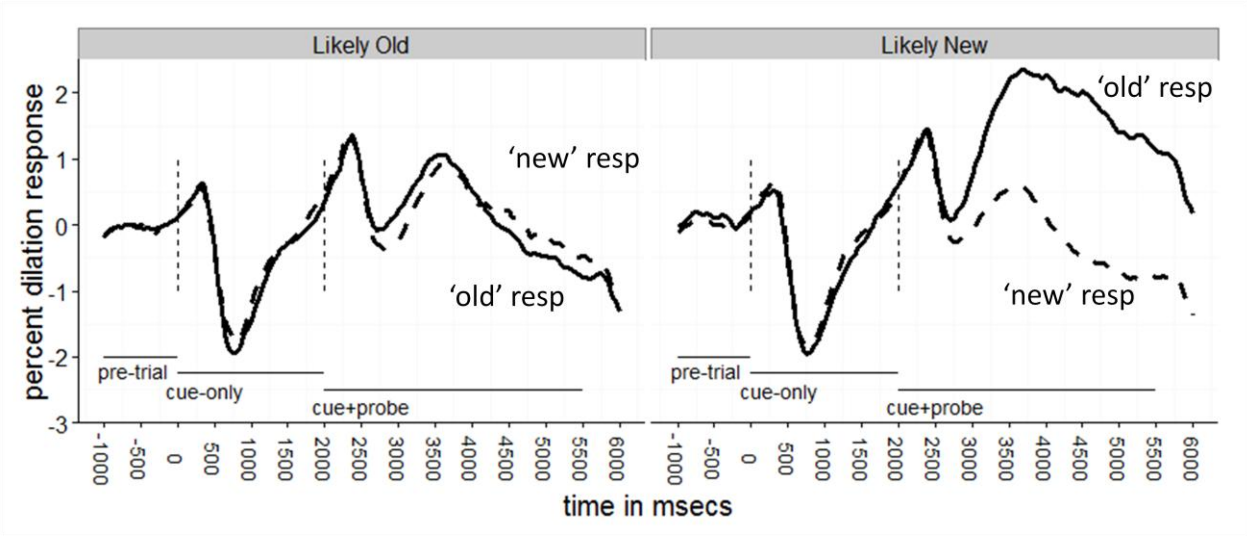


Figure 10. Dilation responses broken down by 'old' and 'new' responses under the two cueing conditions. Data is collapsed across accuracy (correct/erroneous).

3.5. SUMMARY OF MLM OUTCOMES

The multi-level modeling demonstrates that there are two behaviorally dissociable components present in the pupil dilation time-course of each individual recognition judgment trial. These components are isolable into an early amplitude component signalling the unexpected encounter of episodic information (viz. an unexpected recognition response) and a trailing slope component that more generally tracks the effort/uncertainty of recognition judgment. Thus an increasing dilation slope predicts slow (Figure 8), erroneous (Figure 9) and cue-incongruent (Figure 10) recognition judgments (see also SI section 2 for a confirmatory behavioral analysis of confidence collected in a previous explicit memory cueing study). All three cases provide convergent evidence that this component reflects a general marker of judgment uncertainty. In contrast, the early amplitude component of the PD response only yielded a positive effect in the MLMs for 'old' decisions made under LN cues (see right panel, Figure 10; column 6, Table 2) – a finding that corroborates the specificity of the earlier trial-averaged pupil response in signifying the experience of unexpected recognition (see Figure 2). In the Discussion, we further link these two separable PD components to functionally distinct cognitive processes, such that the slope reflects voluntary effort expenditure whereas the early amplitude reflects involuntary orienting to unexpected memory evidence.

Aside from the theoretical importance of demonstrating that two dissociable cognitive processes separately influence the temporal morphology of the dilation response on every trial, the current findings also have methodological significance in showing that simple amplitude averages or peak response measures run the risk of misinterpretation because they miss or conflate valuable information. To date, these types of summary measures have been the only characterization of recognition-linked dilation responses, and hence the failure to isolate trailing slope and early amplitude components may have obscured or weakened earlier effects in the literature.

Table 2. Multi-level modeling of behavioral variables by pupillary dilation components. Note that accuracy is coded as ‘1’=correct, ‘0’=error; response is coded as ‘1’=old, ‘0’=new.

Linear MLMs contrasting Early Intercept and Slope of PD

Dep. Vars	Reaction Time		Accuracy		Response	
Cues	Likely Old	Likely New	Likely Old	Likely New	Likely Old	Likely New
	(1)	(2)	(3)	(4)	(5)	(6)
Early Amplitude	t = 1.668	t = 1.843	t = -0.139	t = -1.247	t = -0.525	**t = 7.166*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Trailing Slope	**t = 10.184*****	**t = 10.040*****	**t = -2.514***	**t = -3.913*****	**t = -2.049***	**t = 3.928*****
	(0.039)	(0.038)	(0.028)	(0.028)	(0.031)	(0.031)
Observations	4,824	4,834	4,824	4,834	4,824	4,834
Note:	*p<0.05; ***p<0.001					
	Subject as random factor					
	Intercepts omitted					
	Parentheses: SE					

4. GENERAL DISCUSSION

The current data demonstrate an important dissociation in the PD response during episodic recognition and suggest a number of refinements to the functional interpretation of dilations during recognition and cognitive judgments more broadly. As noted in the Introduction, the prevailing view is that the pupil old/new dilation response reflects the increased effort required for hits versus correct rejections, as envisioned by the voluntary component of the Cognitive Load model (Papesh et al., 2012; Vo et al., 2008; although see Otero et al., 2011). The primary problem facing this interpretation is that the extant literature indicates that hits are *less* not more effortful than correct rejections, in the sense they are rendered more confidently, often more quickly, and more accurately with deepening levels of processing. Thus to claim that the pupil old/new effect reflects the increased voluntary effort of hits versus correct rejections amounts to abandoning the established connection between mental 'effort' and these convergent behavioral indices (Kahneman, 1973).

To further test the voluntary effort account, and contrast it with one based on involuntary orienting that has yet to be applied to recognition memory, we examined PD during the Explicit Memory Cueing paradigm. Here the effort expended for both correct rejections and hits is directly altered by whether or not the predictive recognition cues are valid or invalid, with the latter expected to heighten effort by slowing performance, increasing errors and generating conflict between expectations and memory signals (as well as reducing subjective confidence, see SI section 2). In other response conflict paradigms, such as the Stroop and Simon tasks, this type of heightened conflict produces greater pupil dilation, which has also been interpreted as reflecting increased effort in line with the Cognitive Load model (Laeng et al., 2011; Steenbergen & Band, 2013). Despite the observed decrement in performance under invalid cueing for both ‘old’ and ‘new’ decisions, suggesting that conflict was indeed instantiated in both conditions, the typical analysis of the trial-averaged pupillary data only yielded a reliable dilation increase for invalid ‘old’ decisions, and not invalid ‘new’ decisions (see Figures 2 and 3). The specificity of this response challenges the link between recognition dilations and general effort/uncertainty implied by previous voluntary load accounts, and instead favors the involvement of an involuntary orienting process sensitive to the recovery of unexpected episodic information. Indeed, a voluntary effort effect on PD was only conclusively isolated from the involuntary orienting component via our multi-level models predicting different behavioral outcomes at the individual trial level. Below we detail the functional significance of these two components, namely the ‘unexpected recognition’ component linked to the early dilation amplitude, and the ‘judgment effort’ component linked to the trailing dilation slope.

4.1. The Two Components of the Dilation Response

The current data reveal two functionally dissociable components in the dilation response accompanying individual recognition decisions. The pupil old/new effect identified in the prior literature appears to be linked with the early amplitude response, as quantified on each individual trial as the intercept of a simple linear regression of the dilation. In contrast, the later trajectory of the dilation response was quantified as the slope, and displays the characteristics of voluntary effort as outlined under the Cognitive Load model. The MLM analysis revealed a positive relationship between the slope of the dilation response and slower reaction times (Table 2 columns 1 and 2), erroneous responding (columns 3 and 4), and responses that conflicted with predictive cues (columns 5 and 6; note sign reversal). Collectively, these convergent behavioral findings suggest that the trailing slope dilation component indexes judgment uncertainty in a general fashion, and in this sense is consistent with the voluntary effort component of the Cognitive Load model. Critically, without decomposing the trial-wise dilation response into two components this link with decision effort/uncertainty would have been less clear.

The second, and often overlooked component of the Cognitive Load model is an involuntary attentional phenomenon linked to Sokolov's classic studies of the orienting response (Kahneman, 1973; Sokolov, 1963a, 1963b). Stimuli typically used to study the orienting response in oddball research are perceptual deviants that violate an established sequence or pattern (Friedman et al., 1973; Hillyard et al, 1971), which raises the question of what should occur when episodic content instead serves as the unexpected event. Adhering to the interpretation of involuntary attention as an ‘orientation towards probable sources of future significant information’ (Kahneman, 1973, page 48) that arise unexpectedly in the environment, the appropriate orienting response to unexpectedly recognized memory probes is to engage long term memory search or source monitoring (Johnson, Hashtroudi & Lindsay, 1993), so as to conclusively identify details of the prior episode (Tulving, 1985). The early amplitude dilation response documented here (both in the MLMs and the trial-averaged PD) is hence consistent with an involuntary attentional process, triggered by orienting to unexpected recognition content to facilitate subsequent memory processing. Additionally, because the involuntary component of the Cognitive Load model is not directly linked with subjective judgment uncertainty or effort (Kahneman, 1973), the fact that hits are subjectively easier than correct rejections poses no problem for its application to the old/new pupil effect.

4.2. Specificity of the unexpected recognition response

The current findings also suggest that judgments of newness are incapable of triggering an orienting dilation response regardless of whether expectations are confirmed or disconfirmed (see Figures 2 and 3). The notion that only 'oldness' information is capable of triggering a robust orienting response is consistent with dual process models of recognition that posit the existence of separate familiarity and recollection processes (for review see, Yonelinas, 2002). Under these models, studied materials can yield an acontextual sense of recent encounter accompanied by fluent processing of the recognition probe (viz., familiarity), and/or they can elicit remembrances of supporting contextual information (viz., recollection). The early amplitude 'unexpected recognition' effect could therefore reflect the recollection of contextual information, which engenders high certainty of prior encounter and thus is experienced as unexpected or surprising when one expects new materials. Since recollection is usually not triggered by new materials (Cox & Dobbins, 2011; Dobbins, 2014; Mickes, Hwe, Wais, & Wixted, 2011), this would explain why an analogous pattern doesn't arise when new items are unexpectedly encountered.

Alternatively, it remains possible that the unexpected recognition dilation is driven by familiarity and not recollection - a possibility bolstered by the characterization of familiarity as an automatic or involuntary memory influence (Jacoby 1991), which is sometimes capable of evoking a strong sense of recent encounter. This follows from Mandler's (1980) famous butcher on the bus anecdote, wherein one encounters someone who immediately evokes a sense of high familiarity but for whom there is initially no accompanying episodic content (e.g. when seeing the butcher on the bus, displaced from the context of his butcher shop). Following this, one begins a deliberate search of memory to try to explain the perceived familiarity, which eventually yields successful recollection to help identify the butcher. Critically, it is the unexpected familiarity of the individual that involuntarily triggers orienting in the Mandler account and most individuals have had experiences of orienting during encounters of familiar individuals in unexpected contexts (see also Godden & Baddeley, 1975).

Ultimately, the current data cannot differentiate between a familiarity and/or recollection account of the unexpected recognition dilation response. Nonetheless they establish that the early dilation response is linked to unexpected recognition rather than unexpected novelty experiences, and it is worth re-emphasizing that the response is fully eliminated when recognition experiences are anticipated and hence unsurprising (rendering it distinct from prior 'retrieval success' interpretations).

4.3. Neurobiology of the Dilation Response

The demonstration of two functionally separable components in the trial-wise dilation response raises interesting questions as to its underlying neurobiology, which has primarily been linked to activity in the locus coeruleus (LC) – a brainstem nucleus linked with arousal (Foote, Aston-

Jones & Bloom, 1980) and which provides norepinephrine broadly to cortex and other structures (Aston-Jones & Cohen, 2005). Critically, electrical activity in the monkey LC is closely associated with the moment to moment dilation of the pupil (Rajkowski, Kubiak, & Aston-Jones, 1993), and concurrently measured pupil diameter covaries with LC activation during both rest and the oddball decision task in human fMRI studies (Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014). Given the close association between the LC activity and the PD response in non-human primates, cognitive researchers have typically taken observed dilation responses as tracking LC activity (Aston-Jones & Cohen, 2005; Nassar et al., 2012; Nieuwenhuis, De Geus, & Aston-Jones, 2011). However, recent animal work has also demonstrated that microstimulation of the intermediate layer of the superior colliculus (SC), a midbrain nucleus traditionally linked to eye movement control and bottom-up attention, also results in pupil dilation. However, unlike the LC, which primarily receives cortical input from medial and orbitofrontal PFC, the intermediate layer of the SC primarily receives cortical input from the lateral frontal and parietal areas (for review see, Aston-jones & Cohen, 2005; Corneil & Munoz, 2014). This raises the interesting possibility that different neural systems may drive dissociable patterns of pupillary dilation; a possibility consistent with prior functional neuroimaging studies involving the Explicit Memory Cueing paradigm, which has revealed dissociable parietal and prefrontal activations underlying an analogous unexpected recognition response (for invalid 'old' decisions only), versus a more general effort/uncertainty response (for both invalid 'old' and invalid 'new' decisions; Jaeger et al., 2013). Probing the potential overlap between these separate unexpected recognition and judgment uncertainty fMRI signatures, the early amplitude and trailing slope dilation components identified in the present report, and the emerging distinction between LC and SC neuromodulatory systems therefore provides an intriguing avenue for future research.

5. CONCLUSION

The current data demonstrate two dissociable psychological components of the pupillary dilation response during recognition memory – an early amplitude component tied to unexpected recognition, and a trailing slope component that tracks the effort of recognition judgments. These components are consistent with the Cognitive Load model if the early amplitude response is considered a form of involuntary orienting, and the trailing slope assumed to reflect voluntary effort expended during recognition judgments broadly. Our working hypothesis is that unexpected recognition adaptively prepares the observer for further long term memory processing and under the Cognitive Load model this is ‘effortful’ only in the sense that it reflects a rapid mobilization of cortical resources in service of long term memory search or monitoring. In contrast, regardless of whether a stimulus is unexpectedly recognized, recognition judgments will also individually vary in terms of subjective uncertainty. Voluntary effort should increase for increasingly non-diagnostic memory signals prone to result in slow, erroneous responding, and additionally, when observers must discount a recommendation known to be generally valid. The trailing slope of the PD response demonstrates these characteristics, becoming increasingly positive when behavioral markers reflect heightened uncertainty. Our findings demonstrate that the unfolding pupil dilation response during each recognition decision contains separable orienting and effort components, thereby providing a window onto the dynamics of memory decision-making.

ACKNOWLEDGEMENTS

Ravi D. Mill was supported by a PhD studentship from the Scottish Imaging Network: A Platform for Scientific Excellence (SINAPSE), and a period of two months at Washington University in St Louis was funded by a SINAPSE Early Career Researcher Exchange Award. Akira R. O’Connor is supported by a SINAPSE Fellowship.

REFERENCES

Ahern, S.K., & Beatty, J. (1981). Physiological evidence that demand for processing capacity varies with intelligence. In M. Friedman, J. P. Dos & N. O'Connor (Eds.), *Intelligence and Learning*. New York: Plenum Press.

Aston-Jones, G., & Cohen, J. D. (2005). An Integrative Theory of Locus Function : Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience, 28,* 403-450. doi:10.1146/annurev.neuro.28.061604.135709

Baldi, P., & Itti, L. (2010). Of bits and wows: a Bayesian theory of surprise with applications to attention. *Neural Networks, 23*(5), 649-666.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276-292.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289-300.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and recognition memory: an attentional account. *Nature Reviews Neuroscience, 9*, 613-625.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215. doi: 10.1038/nrn755

Corneil, B. D., & Munoz, D. P. (2014). Overt Responses during Covert Orienting. *Neuron*, *82*(6), 1230–1243. doi:10.1016/j.neuron.2014.05.040

Cox, J. C., & Dobbins, I. G. (2011). The striking similarities between standard , distractor-free , and target-free recognition. *Memory & Cognition,* 39(6), 925-940. doi:10.3758/s13421-011-0090-3

Dobbins, I., & Han, S. (2007). What Constitutes a Model of Item-Based Memory Decisions? *Psychology of Learning and Motivation, 48,* 95-144.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience, 3*(3), 277-283.

Fiedler, S., & Glockner, A. (2012). The dynamics of decision making in risky choice: an eye-tracking analysis. *Frontiers in Psychology, 3*, 335.

Friedman, D., Hakerem, G., Sutton, S., & Fleiss, J. L. (1973). Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalography & Clinical Neurophysiology, 34*(5), 475-484.

Foote, S., Aston-Jones, G., & Bloom, F. E. (1980). Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal. *Proceedings of the National Academy of Sciences, 77*(5), 3033-3037.

Gardner, R. M., Mo, S. S., & Borrego, R. (1974). Inhibition of pupillary orienting reflex by novelty in conjunction with recognition memory. *Bulletin of the Psychonomic Society, 3*, 237-238.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: on land and underwater. *British Journal of Psychology, 66*(3), 325-331.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil Dilation Reflects the Creation and Retrieval of Memories. *Current Directions in Psychological Science, 21*(2), 90-95.

Heaver, B., & Hutton, S. B. (2011). Keeping an eye on the truth? Pupil size changes associated with recognition memory. *Memory, 19*(4), 398-405.

Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science, 143*(3611), 1190-1192.

Hillyard, S., Squires, K., Bauer, J., & Lindsay, P. (1971). Evoked potential correlates of auditory signal detection. *Science, 172*(3990), 1357-1360.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. doi:10.1016/0749-596X(91)90025-F

Jaeger, A., Cox, J. C., & Dobbins, I. G. (2012). Recognition confidence under violated and confirmed memory expectations. *Journal of Experimental Psychology. General*, *141*(2), 282–301. doi:10.1037/a0025687

Jaeger, A., Konkel, A., & Dobbins, I. G. (2013). Unexpected novelty and familiarity orienting responses in lateral parietal cortex during recognition judgment. *Neuropsychologia*, *51*(6), 1061–1076. doi:10.1016/j.neuropsychologia.2013.02.018

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological* *bulletin, 114*(1), 3-28.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585.

Kahneman, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrmination task. *Perception & Psychophysics, 2,* 101-105.

Konkel, A., Selmeczy, D., & Dobbins, I. (2015). They can take a hint: Older adults effectively integrate memory cues during recognition. *Psychology and Aging, 30*(4), 781-794.

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary stroop effects. *Cognitive Processing, 12*(1), 13-21.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. doi:10.1037//0033-295X.87.3.252

Mickes, L., Hwe, V., Wais, P. E., & Wixted, J. T. (2011). Strong Memories Are Hard to Scale, *Journal of Experimental Psychology: General, 140*(2), 239 –257. doi:10.1037/a0023007

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. doi:10.1002/hbm.22466

Naber, M., Frassle, S., Rutishauser, U., & Einhauser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of*

Vision, 13(2), 11. doi: 10.1167/13.2.11

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience, 15*(7), 1040-1046. doi: 10.1038/nn.3130

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology, 48*(2), 162-175.

Nowak, W., Hachol, A. & Kasprzak, H. (2008). Time-frequency analysis of spontaneous fluctuation of the pupil size of the human eye. *Optica Applicata, 38*(2), 469-480.

O'Connor, A. R., Han, S., & Dobbins, I. G. (2010). The inferior parietal lobule and recognition memory: expectancy violation or successful retrieval? *The Journal of Neuroscience, 30*(8), 2924-2934. doi: 10.1523/JNEUROSCI.4225-09.2010

Otero, S., Weekes, B., & Hutton, S. (2011). Pupil size changes during recognition memory. *Psychophysiology, 48*(10), 1346-1353.

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology, 83*(1), 56-64.

Posner, M. I., Snyder, C. R. R., Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: general, 109*, 160-174.

Preuschoff, K., 't Hart, B. M., & Einhauser, W. (2011). Pupil dilation signals surprise: evidence for noradrenaline’s role in decision making. *Frontiers in Neuroscience, 5,* 115.

Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society for Neuroscience Abstracts*, *19,* 974.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review, 83*(3), 190-214. doi: 10.1037/0033-295x.83.3.190

Selmeczy, D., & Dobbins, I. G. (2012). Metacognitive Awareness and Adaptive Recognition Biases. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(3): 678-690. doi:10.1037/a0029469

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81*(1), 174-176.

Sokolov, E. (1963a). Higher nervous functions: The orienting reflex. *Annual Review of Physiology, 25,* 545-580.

Sokolov, E. (1963b). *Perception and the conditioned reflex*. Pergamom Press.

Steenbergen, H. Van, & Band, G. (2013). Pupil dilation in the Simon task as a marker of conflict processing. *Frontiers in Human Neuroscience, 7,* 215.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662. doi: 10.1037/0096-3445.121.1.15

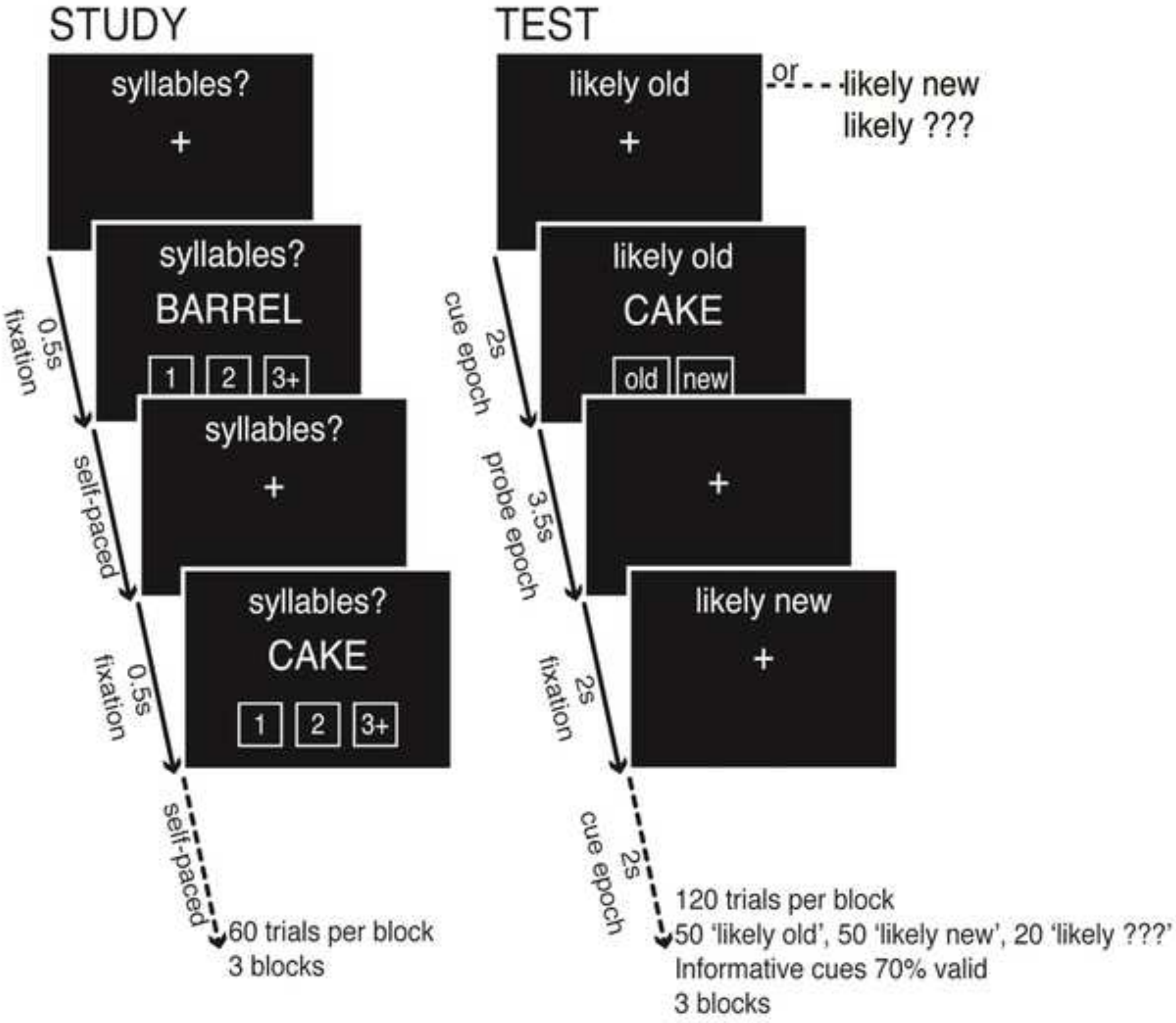
Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*(1), 1–12. doi:10.1037/h0080017

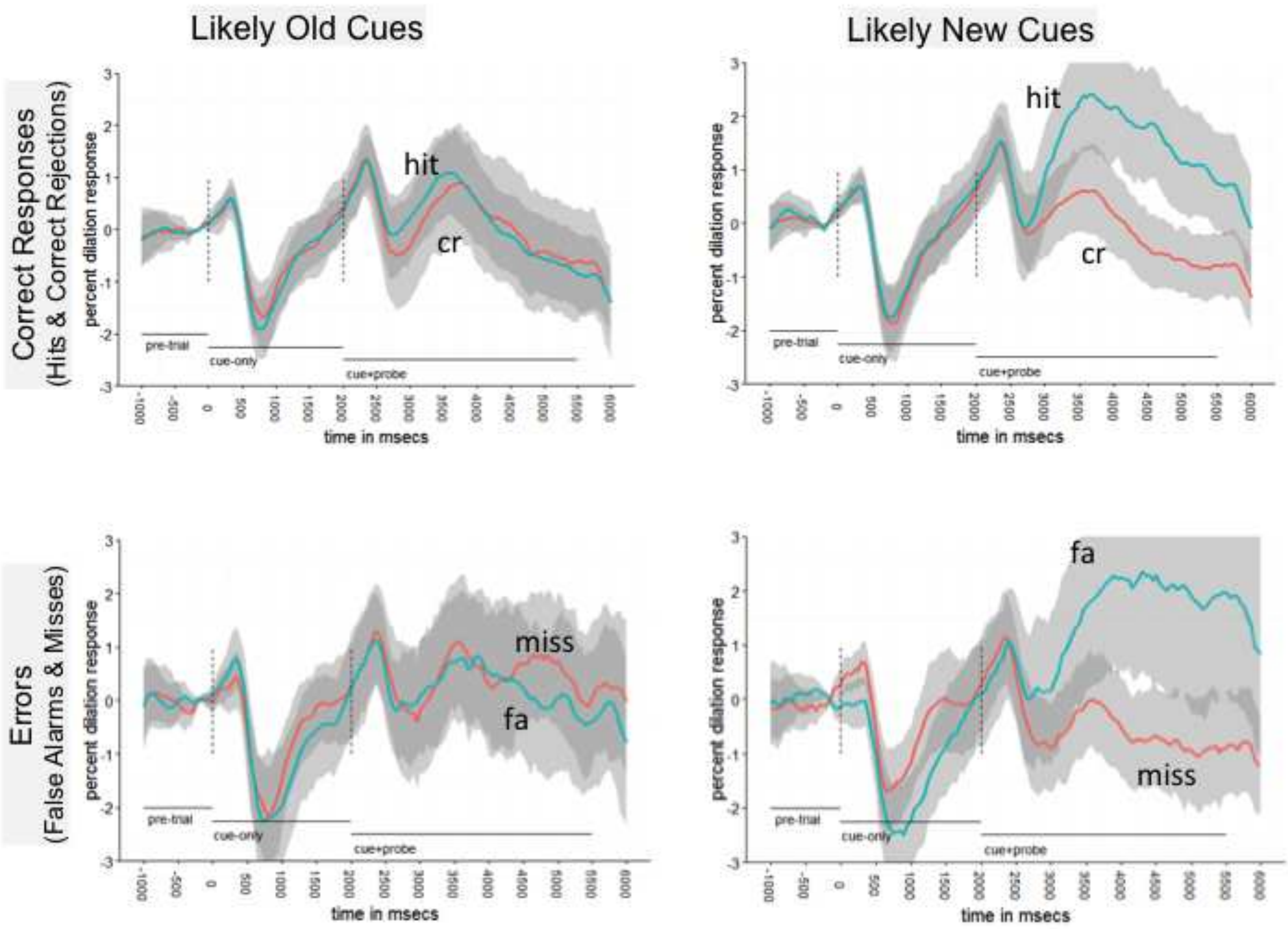
Vo, M. L., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: introducing the pupil old/new effect. *Psychophysiology, 45*(1), 130-140.

Wang, C.A., Boehnke, S.E., Itti, L. & Munoz, D.P. (2014). Transient pupil response is modulated by contrast-based saliency. *The Journal of Neuroscience, 34*(2), 408-417.

Wiese, H., & Daum, I. (2006). Frontal positivity discriminates true from false recognition. *Brain Research, 1075*(1), 183-192.

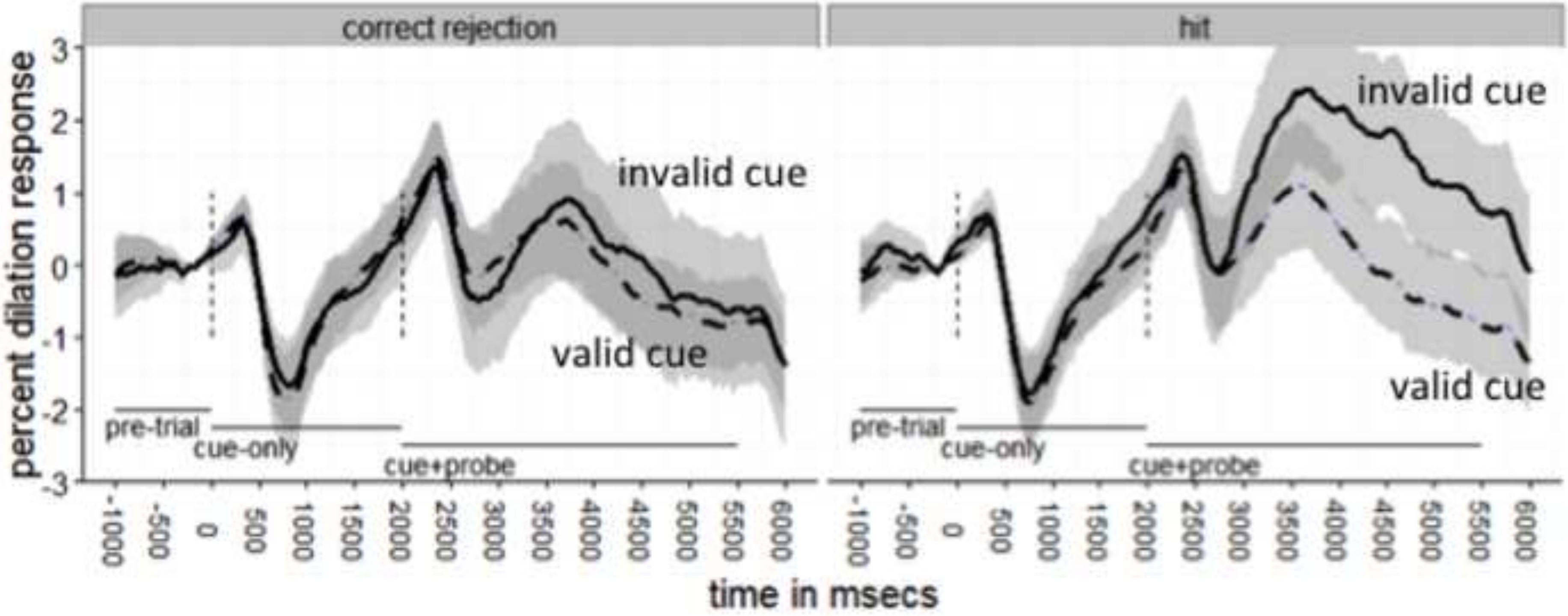
Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity : A Review of 30 Years of Research. *Journal of Memory and Language*, *517*, 441–517. doi:10.1006/jmla.2002.2864


STUDY
syllables?
+
syllables?
BARREL
1
2
3+
syllables?
+
syllables?
CAKE
1
2
3+
0.5s
fixation
self-paced
0.5s
fixation
self-paced
60 trials per block
3 blocks
TEST
likely old
+
or
likely new
likely ???
likely old
CAKE
old
new
+
likely new
+
2s
cue epoch
3.5s
probe epoch
2s
fixation
2s
cue epoch
120 trials per block
50 'likely old', 50 'likely new', 20 'likely ???'
Informative cues 70% valid
3 blocks


Likely Old Cues
Likely New Cues
Correct Responses
(Hits & Correct Rejection)
Errors
(False Alarms & Misses)
percent dilation response
time in msecs
hit
cr
miss
fa
pre-trial
cue-only
cue+probe
3
2
1
0
-1
-2
-3
-1000
-500
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000

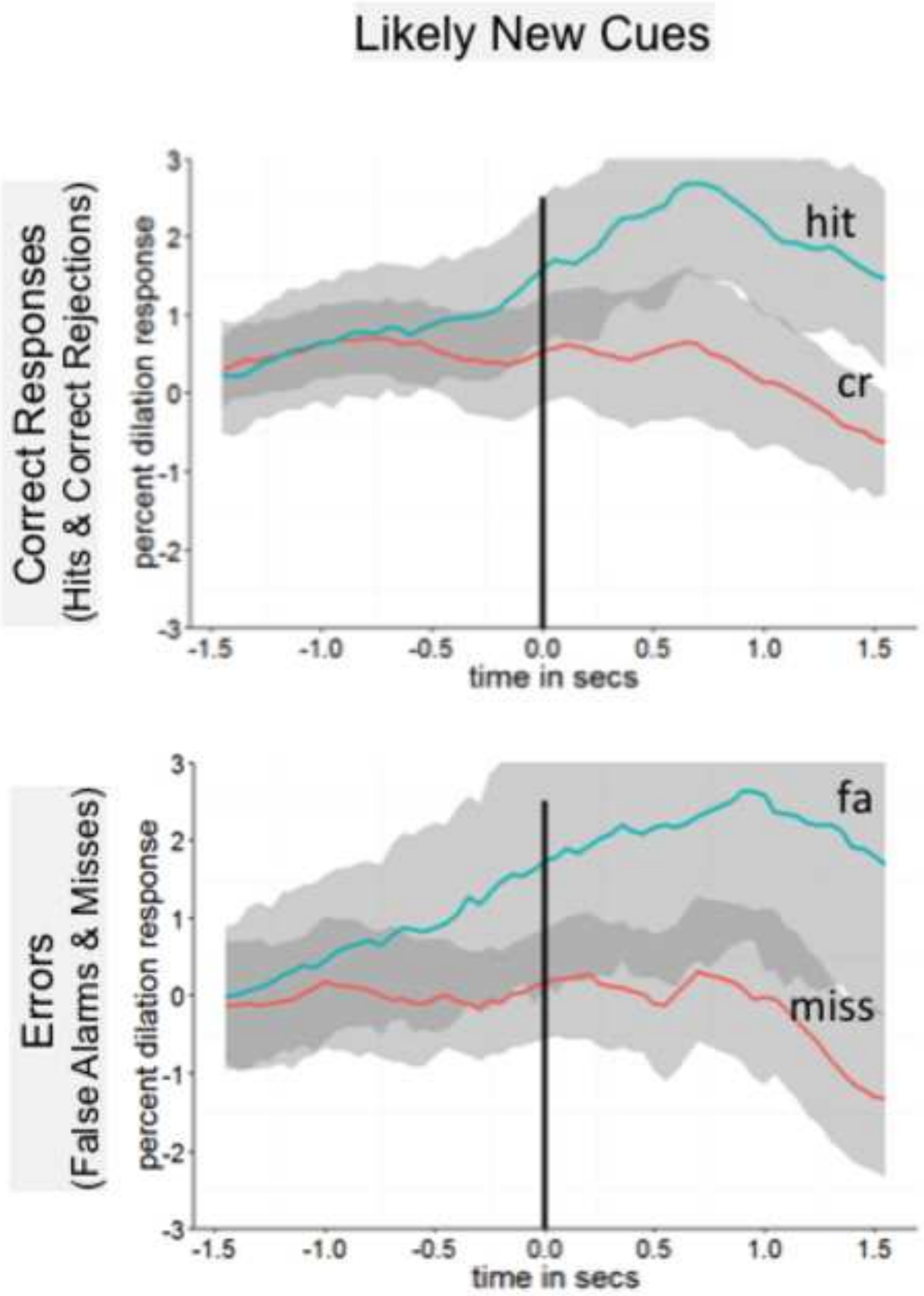
Figure(s)

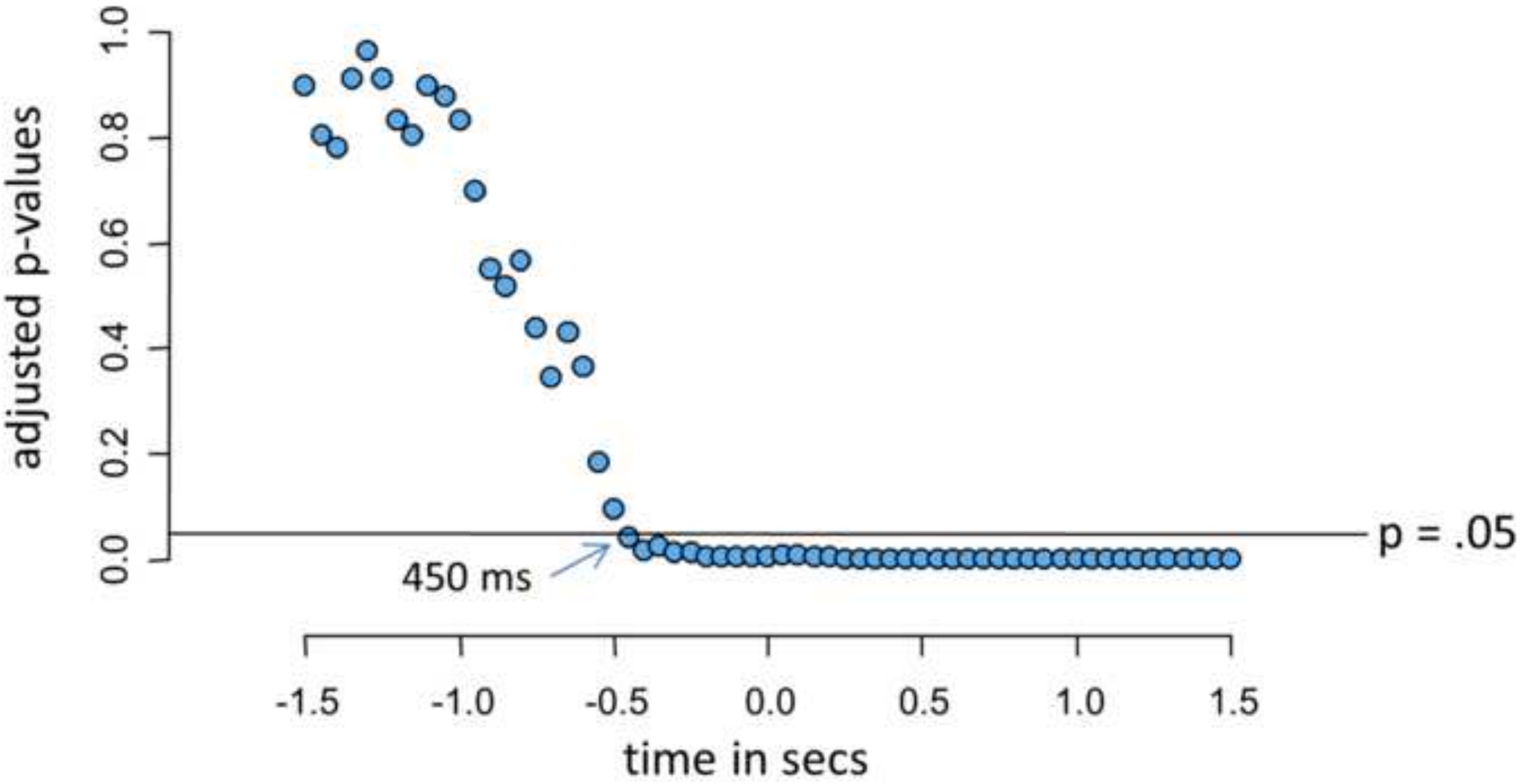
Click here to download high resolution image

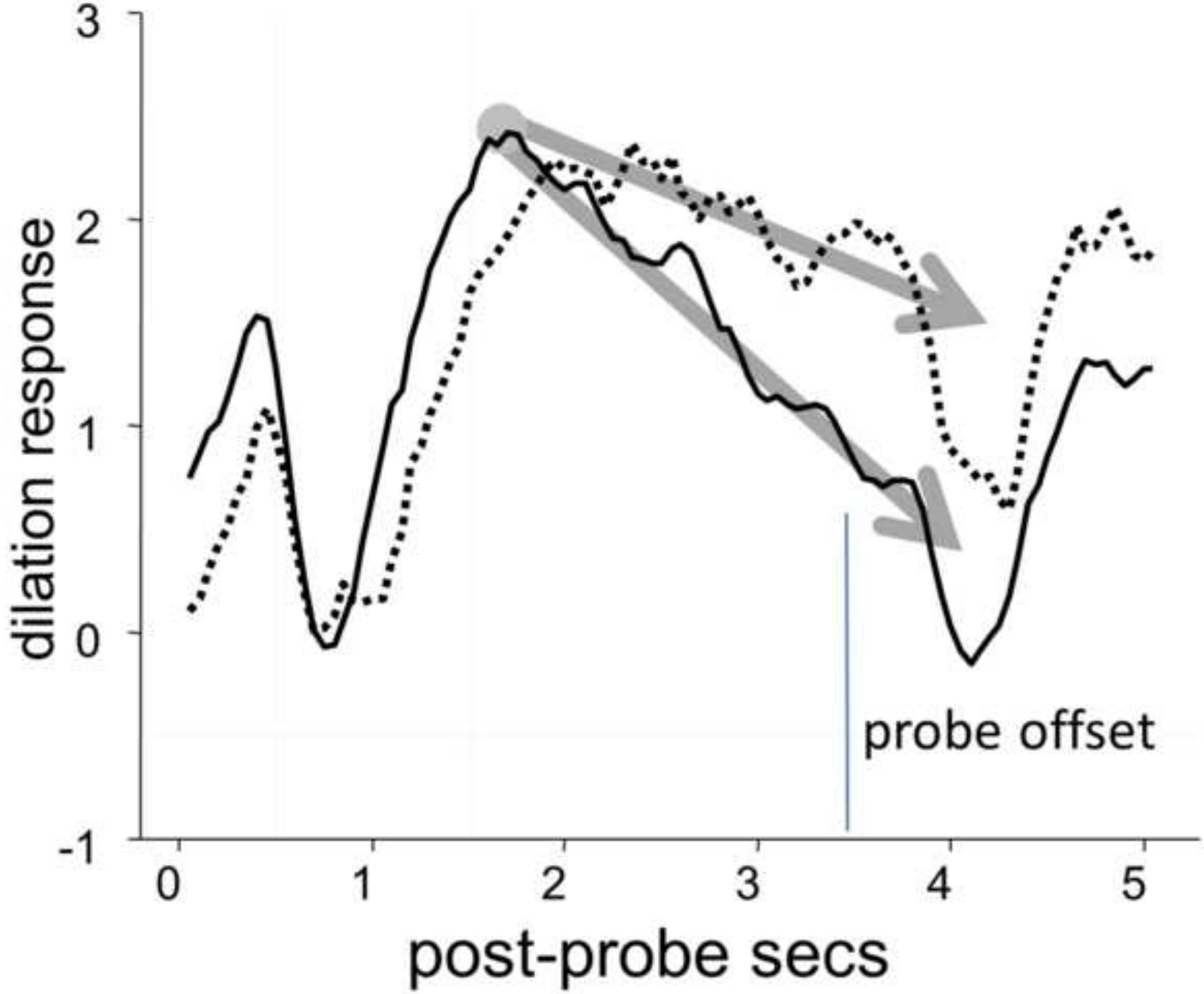


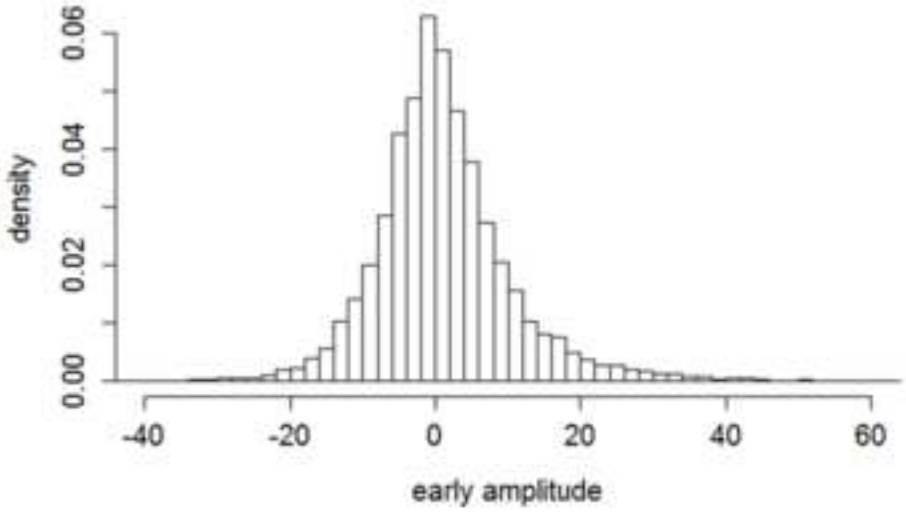
Figure(s)

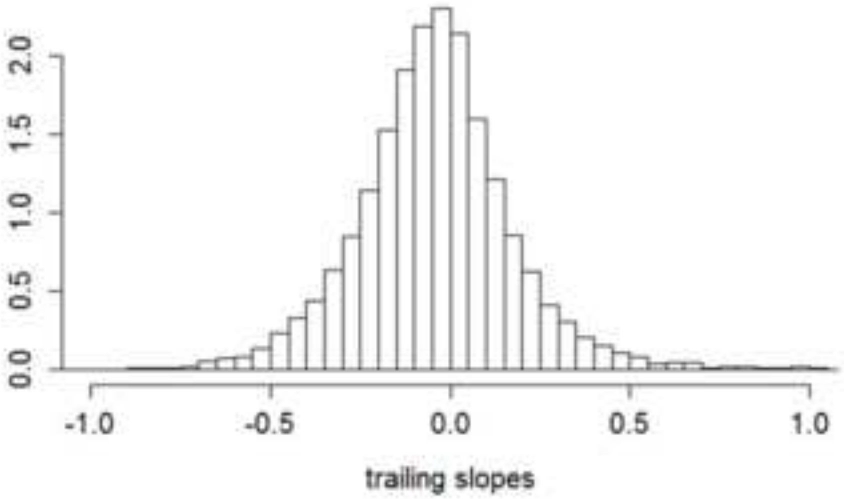
Click here to download high resolution image




adjusted p-values
1.0
0.8
0.6
0.4
0.2
0.0
p = .05
450 ms
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
time in secs

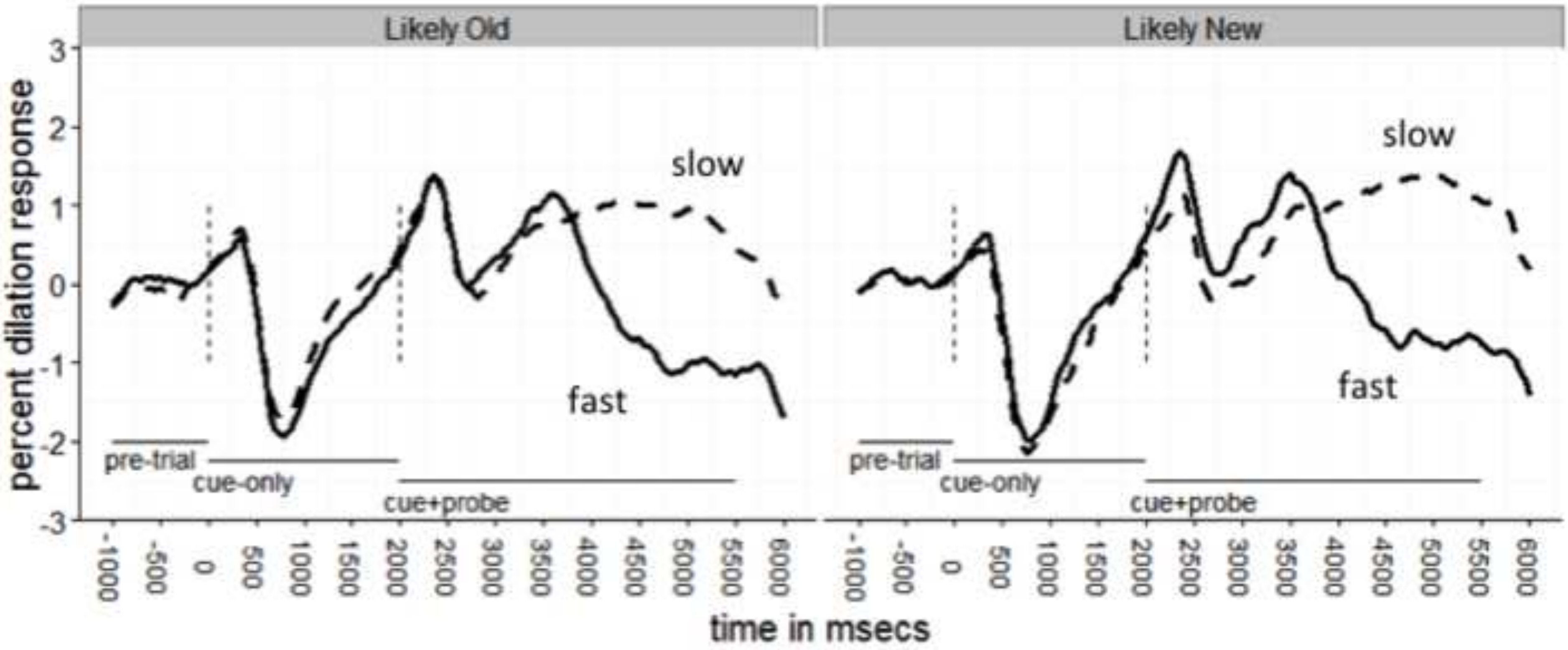

3
2
1
0
-1
dilation response
0
1
2
3
4
5
post-probe secs
probe offset

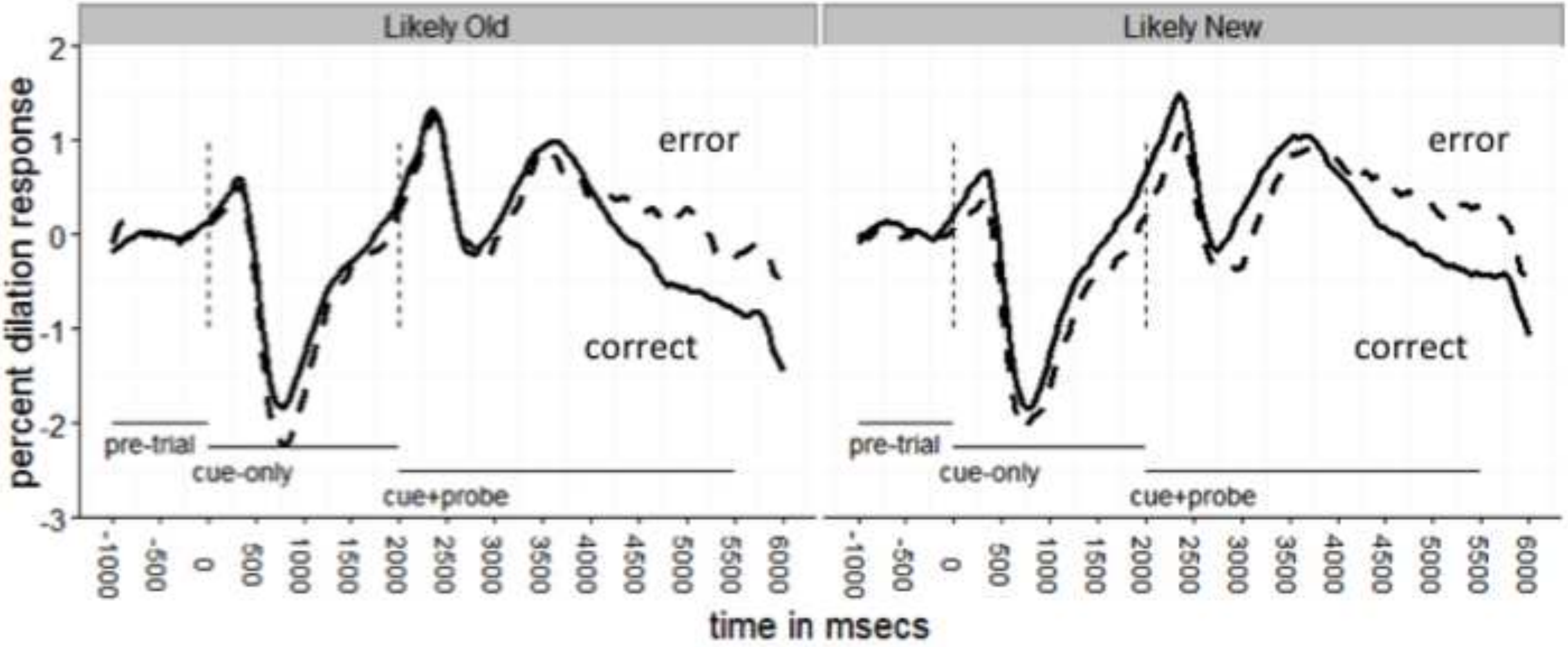

density
0.00
0.02
0.04
0.06
-40
-20
0
20
40
60
early amplitude

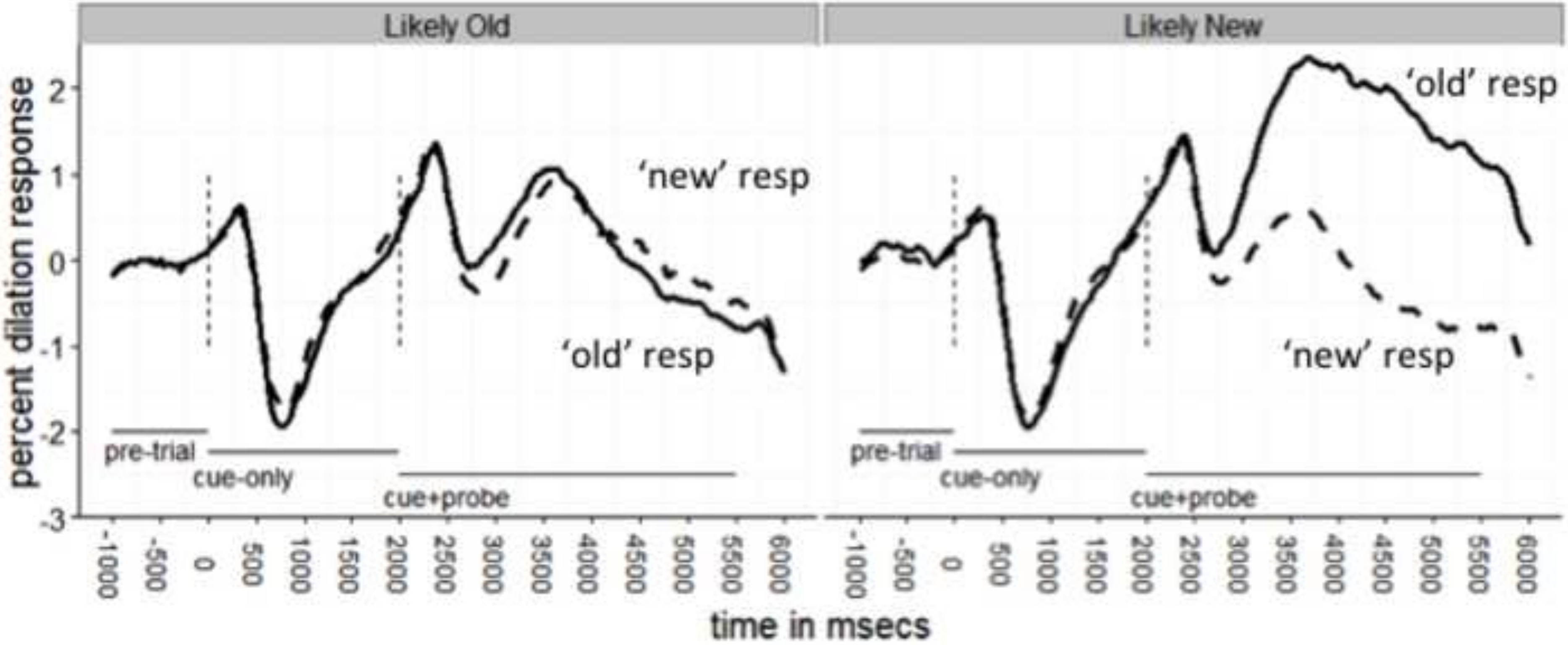

0.0
0.5
1.0
1.5
2.0
-1.0
-0.5
0.0
0.5
1.0
trailing slopes

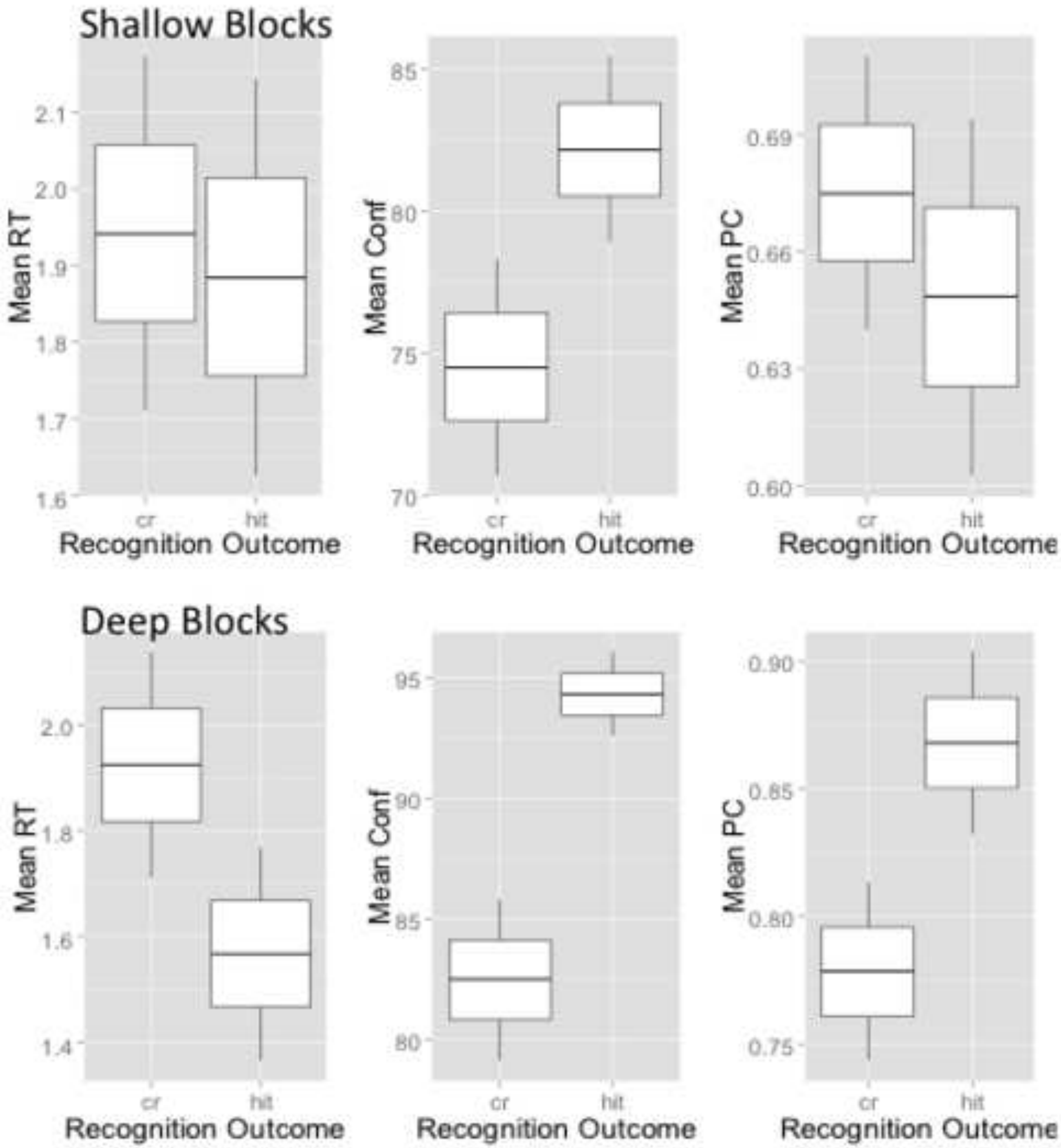
Figure(s)

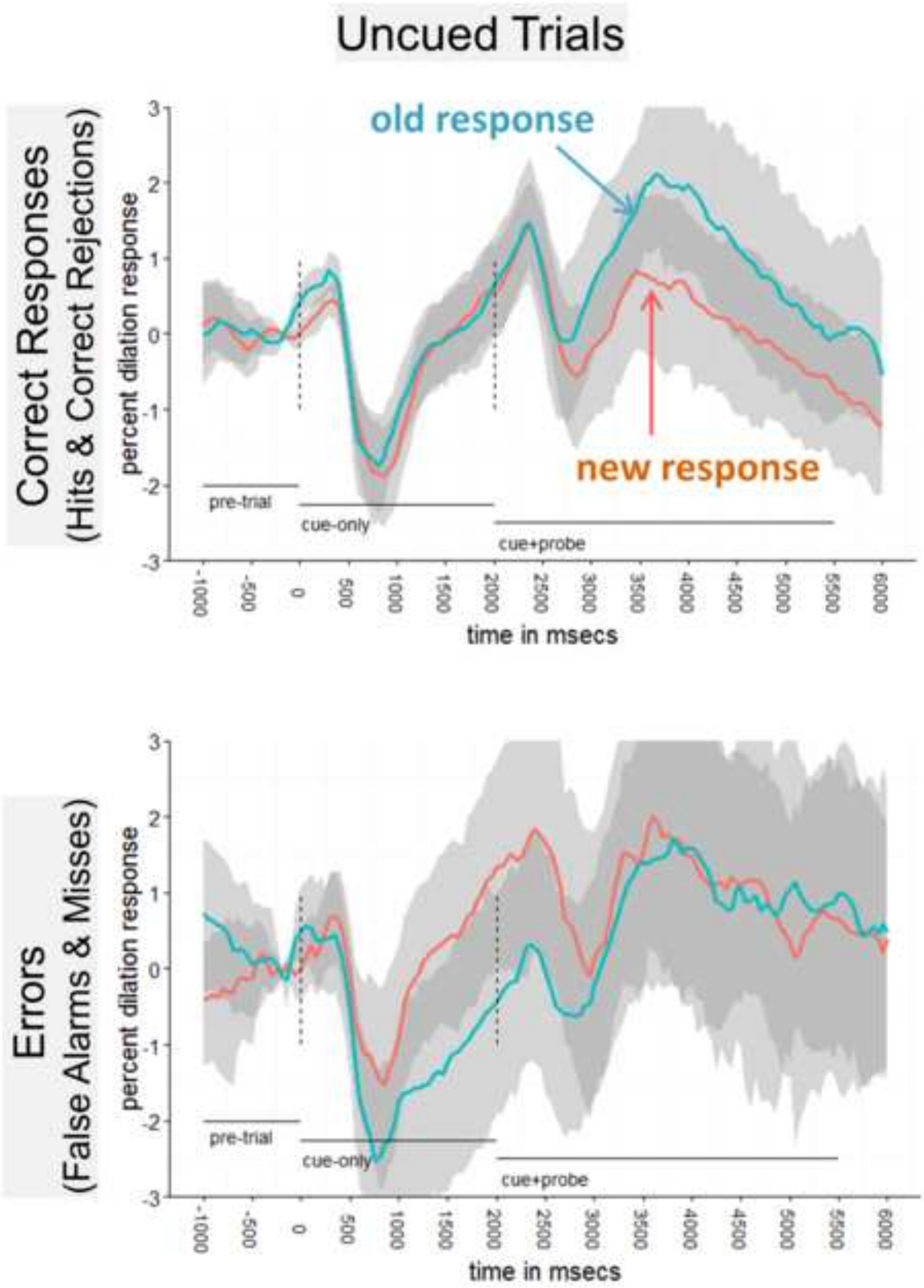
Click here to download high resolution image




Likely Old
Likely New
percent dilation response
2
1
0
-1
-2
-3
error
correct
pre-trial
cue-only
cue+probe
-1000
-500
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
time in msecs


Likely Old
Likely New
percent dilation response
2
1
0
-1
-2
-3
'new' resp
'old' resp
'old' resp
'new' resp
pre-trial
cue-only
cue+probe
pre-trial
cue-only
cue+probe
-1000
-500
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
time in msecs


Shallow Blocks
Mean RT
2.1
2.0
1.9
1.8
1.7
1.6
cr
hit
Recognition Outcome
Mean Conf
85
80
75
70
cr
hit
Recognition Outcome
Mean PC
0.69
0.66
0.63
0.60
cr
hit
Recognition Outcome
Deep Blocks
Mean RT
2.0
1.8
1.6
1.4
cr
hit
Recognition Outcome
Mean Conf
95
90
85
80
cr
hit
Recognition Outcome
Mean PC
0.90
0.85
0.80
0.75
cr
hit
Recognition Outcome


Uncued Trials
Correct Responses
(Hits & Correct Rejection)
old response
new response
percent dilation response
3
2
1
0
-1
-2
-3
pre-trial
cue-only
cue+probe
-1000
-500
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
time in msecs
Errors
(False Alarms & Misses)
percent dilation response
3
2
1
0
-1
-2
-3
pre-trial
cue-only
cue+probe
-1000
-500
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
time in msecs

Highlights for submission to *Cognition*: ‘Pupil dilation during recognition memory: Isolating unexpected recognition from judgment uncertainty’

- Orienting and uncertainty separately drive pupil size during memory decisions.
- ‘Early’ pupil dilation during judgment reflects unexpected recognition.
- Trajectory of pupil dilation during judgment tracks general uncertainty regardless of recognition or expectation.
- This dissociation emerged from novel multi-level modeling of trial-wise dilation responses.
- The same task-evoked dilation timecourse can be split into distinct cognitive components.