Expectation Expands Subjective Duration for Repeated Stimuli by Altering Perception

Corinna D. McFeaters and Daniel Voyer

University of New Brunswick

Author Note

Corinna D. McFeaters, Department of Psychology, University of New Brunswick; Daniel Voyer, Department of Psychology, University of New Brunswick.

This research was made possible by a research grant awarded by the Natural Sciences and Engineering Research Council of Canada (NSERC) to D. Voyer. We thank Stephen J. Wilkins for assistance with image processing and experimental programming and Susan Voyer for constructive feedback on a draft of the article.

Address correspondence to Corinna D. McFeaters, Department of Psychology, University of New Brunswick, PO Box 4400, Fredericton, NB, Canada, E3B 5A3. E-mail: corinna.mcfeaters@unb.ca

Abstract

Repeated stimuli are generally perceived to be shorter in duration than novel stimuli. Matthews (2015), however, demonstrated that when repetition is predictable, expectations of repetition may expand subjective duration for repeated stimuli. Although this effect is hypothesized to be perceptual, this has yet to be empirically established. The present study therefore examined perceptual and decisional factors in the repetition effect by using psychophysical methods while varying probabilities of repetition, in addition to replicating Matthews' original paradigm. Using faces with neutral expressions, sixty participants completed two judgment tasks, indicating whether a comparison stimulus was longer or shorter in duration than a standard stimulus preceding it. Comparison stimuli were presented for the same duration as the 500-ms standard in the replication task and for one of seven durations (from 200 - 1250 ms) in the crucial extension task, allowing for examination of sensitivity and bias. No evidence of bias was observed, but modulating participants' expectations of repetition affected perception, such that discrimination was more difficult under high than low repetition conditions. Overall, participants were more likely to judge stimuli that met expectations as longer, regardless of whether the expectation was repetition or novelty. Implications for models of repetition, context effects, and time estimation are discussed.

Keywords: time perception, repetition, expectation, context effects

Public Significance Statement

This study demonstrated that the duration of novel and repeated images is experienced differently when repetition is common than when it is uncommon due to changes in the way time is processed. Durations for images that match circumstantial expectations (e.g., a repeated image when repetition is common) are more accurately judged, suggesting that context influences how time is perceived. This work suggests the need to account for context in signal detection situations (e.g., air traffic control) when events are time sensitive but highly repetitive.

Repetition and Subjective Duration: Context-Based Expectations Alter Perception

Does time tick differently for something new than for a stimulus we have come across recently? Recent exposure to a stimulus typically results in a shortening of perceived duration on re-exposure (Matthews, 2011; Matthews & Meck, 2016; Rose & Summers, 1995), but this effect appears to be influenced by stimulus expectations (Matthews, 2015). Matthews proposed a mechanism for this modulation, but it has yet to be evaluated empirically. The present study therefore aimed to serve as a preliminary evaluation of the basis for the interaction observed between stimulus repetition and expectation.

Much of the research in this area has employed the oddball task, in which an oddball stimulus is embedded in train of repeated stimuli (e.g., Tse et al., 2004). In this task, the oddball stimulus seems longer, suggesting that, in comparison, duration has been contracted for the repeated standards. Matthews (2011) demonstrated that this effect is also observable in a two-stimulus paradigm. This contraction of subjective duration has been linked with the attenuation of neurological signals for repeated stimuli (Desimone, 1996). This phenomenon, known as *repetition suppression*, has been alternately hypothesized to reflect both low-level consequences of continuous perceptual information flow (see Grill-Spector, Henson, & Martin, 2006, for a review) and top-down contributions of expectation on information processing (e.g., Auksztulewicz & Friston, 2016; Friston, 2005). Under the latter view, encountering a stimulus in a particular context establishes its existence and likely recurrence in that environment. Subsequent encounters with the same stimulus produce lower levels of initial prediction error (i.e., the difference between observed and expected input), resulting in reduced neural responses to repeated stimuli.

The Role of Expectation

There is evidence, however, that repetition and expectation may actually operate differentially. Summerfield, Trittschuh, Monti, Mesulam, and Egner (2008), for example, found that repetition suppression occurred regardless of expectation, but it was greater when repetition was expected, suggesting that expectation modulated repetition suppression. This result has been well replicated (e.g., Andics, Gál, Vicsi, Rudas, & Vidnyánszky, 2013; Mayrhauser, Bergmann, Crone, & Kronbichler, 2014; Summerfield, Wyart, Mareike Johnen, & De Gardelle, 2011). Additionally, Todorovic and de Lange (2012) observed that expectation and repetition produced suppression under different circumstances and on different time scales. Furthermore, some studies have even documented enhancement of neural signals in response to repeated stimuli, including repeated durations (e.g., Wiener & Thompson, 2015), and a number of non-temporal variables have been shown to produce neural repetition enhancement, including attention, recognition, and explicit memory (see Segaert, Weber, de Lange, Petersson, & Hagoort, 2013, for a review). Taken together, the neural evidence suggests that the effects of repetition and expectation are separable and that task demands may influence whether repetition produces neural suppression or enhancement.

Behaviourally, there are again indications that repetition and expectation may affect duration judgment in different ways. In a comprehensive investigation of the roles of repetition and expectation, Matthews (2015) demonstrated that expectation moderates the degree of duration contraction for repeated stimuli. When the probability of repetition was low (i.e., the expectation was for novelty), Matthews found that repeated stimuli were judged to be shorter in duration than novel stimuli, but when repetition was expected, subjective duration for the repeated stimuli was lengthened, reducing and even reversing the typical repetition effect. Cai, Eagleman, and Ma (2015) also conducted a series of experiments examining the effects of repetition and expectation on duration perception and observed that repetition consistently shortened perceived duration, whereas expectation, if anything, increased it. In the same vein, Birngruber, Schröter, Schütt, and Ulrich (2018) reported that stimuli conforming to participants' self-generated verbal predictions of the upcoming stimulus were judged to be longer in duration than those that differed from these predictions, suggesting that meeting expectations prolongs duration. In contrast, however, some studies have found longer subjective durations for unexpected stimuli (e.g., Ulrich, Nitschke, & Rammsayer, 2006). It thus appears that, similar to the neural findings for repetition, specific task demands may determine the effect of expectation.

Theoretical Accounts of the Interplay between Expectation and Repetition

Matthews (2015) framed the interplay between repetition and expectation as one of opposition. Repetition produces a subjective contraction in duration consistent with low-level repetition suppression secondary to reductions in neural signal strength. In contrast, expectations exert an expansive top-down influence that works against that of repetition. Specifically, expectation was hypothesized to act on the percept, enhancing its representation, which improves information extraction and perceptual decision making. Matthews proposed that subjective duration is a function of the strength of the percept; thus, factors that boost perceptual decision-making will prolong subjective duration (also see Matthews & Gheorghiu, 2016; Matthews & Meck, 2016).

A perceptual origin has been established for the temporal oddball effect (Birngruber, Schröter, & Ulrich, 2014), but it is not clear whether the oddball effect fully shares the same basis as the effect described by Matthews (2015). Specifically, although expectation has been hypothesized to contribute to duration contraction of the repeated standards in the oddball paradigm, in Matthews' paradigm, expectation has expansive properties. In fact, in response to the findings of expectation expansion, the role of expectation in the oddball effect has been called into question (Birngruber et al., 2018). Bang and Rahnev (2017) also recently demonstrated that visual expectations altered decisional criterion, rather than perception, which had previously been hypothesized as a potential mechanism by Summerfield and Egner (2009). It is therefore plausible that the differences in implicit expectations generated through manipulation of repetition probability altered decision-making, rather than perception. Given the specific correspondence between perceptual representation (in this case, a visual representation) and subjective duration that Matthews (2015) proposed, it seems prudent to rule out the possibility that decision-making processes might account for the observed effects.

Perceptual and Decisional Bases

Establishing a perceptual or decisional basis for the interaction between repetition and expectation cannot be accomplished by relying solely on standard psychophysical measures, as the thresholds used to calculate these measures are known to be sensitive to the influence of both nonsensory and decisional factors (Green & Swets, 1966; Macmillan & Creelman, 2005). As a result, we cannot distinguish between a genuine change in the observers' perceptual experience, an effect of nonsensory origin (such as a context effect), or differences in the amount of sensory evidence required to make a decision. Signal detection theory, however, addresses this obstacle by providing separate measures of signal-related and decisional influences, the discrimination sensitivity index (typically, d') and the criterion measure (typically, β), respectively.

A criterion shift would indicate that expectation is altering response patterns, rather than expanding duration by strengthening perceptual representations for the expected repeated stimuli, although it would not allow us to distinguish between perceptual and response biases (Witt, Taylor, Sugovic, & Wixted, 2015). In contrast, an alteration in perceptual representation should result in changes in discrimination sensitivity (Dosher & Sperling, 1998). Therefore, in the context of the Matthews' (2015) proposal, because stronger representation is hypothesized to correspond to longer subjective duration and changes in representation are indexed by measures of sensitivity, increases in subjective duration for expected repetitions should also be accompanied by differences in discrimination sensitivity for these items. Indeed, such a relationship between subjective duration and sensitivity has been observed previously. Birngruber et al. (2018), for example, observed matching increases for both duration and sensitivity for expected stimuli, although this effect was not observed consistently across experiments. Ulrich et al. (2006) conversely observed a decrease in sensitivity and duration for expected, relative to unexpected, stimuli, but the magnitude of the sensitivity effect varied across experiments. As such, it is reasonable to expect differences in sensitivity based on expectation's influence, although predicting the direction of the effect might prove challenging.

The current study therefore aimed to investigate whether the effect reported by Matthews (2015) arises due to changes in discrimination sensitivity or because of a criterion shift. Based on Matthews' results, we expected that trial type would interact with repetition condition such that the subjective duration of repeated and novel stimuli would vary as a function of repetition condition. We also expected that this prediction would hold for both a replication task, in which Matthews' approach was reproduced, and an extension task using the wider range of duration necessary to evaluate discrimination and bias. If the expectation of stimulus repetition affects perception by strengthening the perceptual representation of the repeated stimulus, we should also observe a trial type by repetition condition interaction effect for sensitivity (*A'*). If, however, the effect is associated with changes in response criterion, this would suggest another mechanism of action.

Method

Participants

Sixty individuals (13 male, 47 female) between the ages of 17 and 68 (*M* = 26.73, *SD* = 11.37) were recruited from Introductory Psychology classes and the general campus community at a comprehensive eastern Canadian university. Power calculations for the linear mixed effects models fitted in our analyses (see Statistical Analysis, below) rely on simulation, which requires information not readily available from previous reports of the effect under investigation. As such, we computed typical *a priori* power analysis by means of G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for a 2 by 2 repeated measures ANOVA. This analysis suggested that a sample size of 12 would be required to achieve 95% power with the effect size corresponding to the trial type by repeat condition interaction ($\eta^2_p = .25$) reported by Matthews (2015, Experiment 1), whereas 53 participants would be required to achieve 95% power with a medium effect ($f = .25$). Therefore, our selected sample size provided ample statistical power. Participants received either a $15 payment or bonus course credit as compensation for their time. Three participants were excluded from analysis because their error rate for the attention check items was greater than 20% (see Replication, below), and two were removed because at least one of their difference thresholds could not be calculated (see Statistical Analyses, below), leaving a final sample of 55.

Materials

Stimuli for both tasks were drawn from a pool of 1195 color photographs of forward-facing faces with neutral expressions obtained from the following face databases: the Glasgow Unfamiliar Faces Database (Burton, White, & McNeill, 2010), the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015), the Psychological Image Collection at Stirling (PICS; University of Stirling Psychology Department, n.d.), and the Tarrlab Face Database (Stimulus images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, http://www.tarrlab.org/. Funding provided by NSF award 0339122.). The collection of faces depicted individuals of both sexes from a wide range of ages and ethnic backgrounds.

The images from each database differed in terms of size, background, the amount of background space, clothing, and the amount of neck and shoulder displayed. Images with large proportions of background area were cropped to match the amount of background space in other images, and the background was edited in any images not set against a white background. All images were then resized to measure a uniform 300 pixels in width; height varied according to the aspect ratio of the original/cropped images and ranged from approximately 225 pixels to 300 pixels. Stimuli were centrally presented against a white background on 17-in. LED monitors with a 60 Hz refresh rate and a resolution of 1024 x 764 pixels. Participants were positioned approximately 60 cm from the screen. Response options were presented with either a left-*shorter*/right-*longer* or left-*longer*/right-*shorter* mapping in order to control for potential spatial stimulus-response effects (Vallesi, Binns, & Shallice, 2008). Stimulus presentation was controlled using PsychoPy2 v1.84.2 (Peirce, 2007) on computers running Windows 7 Enterprise with on-board Intel graphics.

Procedures

Participants completed two tasks: a replication task designed to verify our ability to obtain the same pattern of results as Matthews (2015) and an extension task designed to assess the psychophysical properties of the observed phenomenon. Order of task completion, stimulus conditions, and spatial mapping of the response options were all counterbalanced across participants. Before beginning their first experimental task, participants completed six practice trials following the method described below with object stimuli consisting of a 500 ms presentation of the standard stimulus and one 250 ms, one 500 ms, and one 1250 ms comparison trial of each trial type (repeat vs nonrepeat).

Replication. The replication task followed the procedures used by Matthews (2015, Experiment 1), with each trial consisting of the serial presentation of a standard stimulus lasting

500 ms, a blank interstimulus interval lasting 300 ms, and a comparison stimulus with a typical duration of 500 ms (but see below). After the comparison stimulus, participants were asked to indicate whether it was shorter or longer in duration than the standard stimulus by clicking the appropriately labelled response option on screen. After response selection, a blank inter-trial interval of random duration (ranging from 880 to 1120 ms) preceded the presentation of the next trial.

Trials were divided evenly over 6 blocks in each task. In half of the blocks, the image presented for the standard and comparison stimuli was the same (repeat trial) for 75% of the trials and different (nonrepeat trials) for 25% of the trials, constituting a high repeat-rate condition. This proportion was reversed in the other half of the blocks (75% nonrepeat trials, 25% repeat trials; low repeat-rate condition). All three blocks for each repeat-rate were presented consecutively (as in Matthews 2015, Experiment 4a).

For this replication task, participants completed 288 experimental trials divided into 6 blocks of 48 trials. Within each block, comparison stimuli were randomly presented for either 250 ms (4 trials), 500 ms (40 trials), or 1250 ms (4 trials). Only the trials in which the comparison stimulus lasted 500 ms were included in the analysis, resulting in 30 repeat trials and 90 nonrepeat trials in the low repeat-rate condition and the reverse in the high repeat-rate condition (90 repeat and 30 nonrepeat trials). The other 8 trials per block were used to check attention to the task. As mentioned earlier, any participant achieving an error rate greater than 20% on these items was excluded from further analysis for both tasks. Stimuli for this task were randomly selected without replacement from a subset of 435 different face images from the pool described above.

Psychophysical Extension. The extension task also was modeled after Matthews (2015, Experiment 1), with the blocked format he used in his fourth experiment. Its procedure and materials were the same as for the replication task, above, with the following changes. It consisted of 504 trials presented over 6 blocks (84 trials per block) and relied on the method of constant stimuli to compute the difference threshold and point of subjective equality (see Statistical Analyses, below, for details). In addition, in this task, the comparison stimulus was randomly presented for one of seven logarithmically-spaced durations: 200 ms, 271 ms, 368 ms, 500 ms, 679 ms, 921 ms, or 1250 ms. A logarithmic scale was chosen on the basis that duration appears to be handled by a general magnitude system that also handles numerosity and distance (see Kadosh, Lammertyn, & Izard, 2008; Walsh, 2003), and this magnitude system, from a neurological perspective, appears to be logarithmically scaled (Dehaene, 2003). Each duration was presented 12 times per block. Thus, for each comparison duration, we had 9 nonrepeat trials and 27 repeat trials in the high repeat-rate condition (totaling 63 nonrepeat and 189 repeat trials) and 27 nonrepeat trials and 9 repeat trials for the low repeat-rate condition (totaling 189 nonrepeat trials and 63 repeat trials). Like in the replication task, stimuli were randomly selected without replacement from a new subset (760) of face images from the pool described above. None of the images used in the replication task were shown during the extension task.

Statistical Analyses. Sensitivity in stimulus discrimination (*A'*) and bias in discrimination probability (B''_D) were calculated using nonparametric signal detection analysis (Donaldson, 1992). The hit rate corresponds to the number of *longer* responses to trials where the comparison stimulus was longer in duration than the standard and the false alarm rate reflects *longer* responses on trials where the comparison stimulus was actually shorter in duration than the standard. Nonparametric indices were used because we expected that variance estimates might differ significantly between conditions as a result of the unbalanced nature of the experimental design. To provide a point of comparison with Matthews' (2015) results and ensure replication, the proportion of *longer* responses were also analyzed for both the replication and extension tasks.

Dixon (2008) recommends using a logistic linear mixed-effect analysis for analyzing accuracy data (such as the proportion of *longer* responses) within a repeated measures design, as using analysis of variance for this type of data can produce distortions in the patterns of significance, especially for interactions. Because the interaction between repeat-rate and trial type is the main effect of interest, the logit approach was chosen to ensure accuracy with respect to significance tests. For consistency, A' and B''_D were also analyzed using the mixed-models approach.

All model fitting and statistical analyses were conducted in the R environment (R Core Team, 2013) using the *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *phia* (De Rosario-Martinez, 2015) packages. For all mixed effects models, repeat condition (RC: high repeat vs low repeat) and trial type (TT: repeat vs nonrepeat) were entered as independent (fixed) variables, unless otherwise specified. For each dependent variable (Y), models containing random intercepts for each participant

$$Y \sim \text{RC} * \text{TT} + (1 \mid \text{Participant}) \quad (1)$$

were compared with models containing random intercepts for participants plus random slopes for the experimental conditions and their interaction terms

$$Y \sim \text{RC} * \text{TT} + (1 + \text{RC} * \text{TT} \mid \text{Participant}) \quad (2)$$

The duration of the comparison stimulus was also included as an independent variable for the extension task and included in the random slope component. Only results from the best fitting models, selected on the basis of deviance statistics (Bates et al., 2015), are reported below. For simple main effects, a Holm adjustment was used to control for family-wise error. Standard errors and confidence intervals were calculated for within-subjects designs (Morey, 2008) using the *Rmisc* (Hope, 2013) package.

In order to calculate evidence in favor of the perceptual and response criterion explanations, an approximation of the Bayes factor derived from the Bayesian Information Criterion (BIC) was calculated according to the method described by Song, Nathoo, and Masson (2017) for the psychophysical measures. The Bayes factor is a statistical value that quantifies the weight of evidence in favor of one hypothesis over another (Wagenmakers, 2007). The Bayes factor can be approximated using the formula $BF_{12} \approx \exp\{(BIC_{M2} - BIC_{M1}) / 2\}$ where M1 and M2 represent the models being compared. For additional information, see, among others, Jeffreys (1961), Kass and Raftery (1995), and Lee and Wagenmakers (2013).

Finally, information on difference thresholds and the point of subjective equality (PSE) may be of interest to some readers. Therefore, although this was not the primary purpose of the present study, the difference threshold and PSE were calculated and analyzed. Further information about the calculation of these variables and their analyses is included in an Appendix. It should be noted that two participants were excluded from all statistical analysis because a lower threshold for their psychophysical function could not be calculated (see the Appendix for more information).

Results

A' and *B''*$_D$

No effects (main or interaction) were found for B''_D (repeat condition: $LR(1) = 0.09$, $p = 0.77$; trial type: $LR(1) = 1.44$, $p = 0.23$; interaction: $LR(1) = 0.15$, $p = .70$). See Figure 1 for relevant means. One might be tempted to argue that the lack of significant effects for B''_D could reflect low statistical power. However, the estimated Bayes factor suggests that the data were 7.17 times more likely to occur under the null model than a model including an effect for trial type, which was the effect closest to statistical significance on B''_D.

For *A'*, a main effect of repeat condition was observed, $LR(1) = 4.13$, $p = .042$. Sensitivity to stimulus difference was significantly higher in the low repeat condition (M = 0.959, SE = 0.002) than the high repeat condition (M = 0.952, SE = .003). No effect of trial type, $LR(1) = 0.72$, $p = .40$, was found, nor was the interaction term significant, $LR(1) = 1.82$, $p = .18$.

The dissociation in significant results that we found for *A'* and B''_D is crucial to our examination of whether the influence of repetition on time estimation reflects changes in perceptual sensitivity or a criterion shift, but direct comparison across these dependent variables is difficult. With this in mind, we also computed estimated Bayes factors relevant to the significant effects reported above for *A'* as a complement of information. For *A'*, the estimated Bayes factor suggested that the null model was 1.92 times more likely than an alternative model that included an effect of repeat condition, although this does not provide more than weak evidence in favor of one model over the other (Kass & Raftery, 1995). In contrast, the Bayes factor of 7.17 obtained in favor of the null model for B''_D provides positive evidence in favor of a null effect.

Proportion of *Longer* Responses

For the replication task, a repeat condition by trial type interaction, $LR(1) = 31.23$, $p < .001$ (Figure 2) was observed. Analysis of the simple main effects revealed that in the low repeat condition, mean proportion of *longer* responses was significantly higher for nonrepeat trials than for repeat trials, $LR(1) = 4.82$, $p = .041$. In the high repeat condition, repeat trials had a significantly higher mean proportion of *longer* responses than nonrepeat trials, $LR(1) = 5.36$, $p = .041$. There was no main effect of repeat condition, $LR(1) = 0.86$, $p = .35$, or trial type, $LR(1) = 0.005$, $p = .95$.

Analysis of mean proportion of *longer* responses in the extension task also resulted in a significant repeat condition by trial type interaction, *LR*(1) = 11.17, $p < .001$ (Figure 3). Analysis of simple main effects revealed that mean proportion of *longer* responses was smaller for repeat trials in the low repeat condition than in the high repeat condition, *LR*(1) = 12.21, $p < .001$, whereas no difference was observed for nonrepeat trials, overall, *LR*(1) = 0.10, $p = .75$. A main effect of stimulus duration, *LR*(6) = 7793.92, $p < .001$, was qualified by a repeat condition by stimulus duration interaction, *LR*(6) = 17.62, $p < .001$. Mean proportion of *longer* responses for the 200 ms duration was significantly higher in the high repeat condition than in the low repeat condition (Figure 4), but no other differences were found (200 ms: *LR*(1) = 14.00, $p = .001$; for all other durations, $p > .26$). There was no effect of repeat condition, *LR*(1) = 1.71, $p = .19$, or trial type, *LR*(1) = 0.09, $p = .76$, and no interaction between trial type and stimulus duration, *LR*(6) = 5.36, $p = .50$, or trial type, repeat condition, and stimulus duration, *LR*(6) = 3.06, $p = .80$.

Supplemental analyses. The data for the replication were also compared to the 500-ms subset of the data from the extension task to assess for findings at this duration that may be attributable to the addition of the other comparison durations in the extension task. A repeat condition by trial type interaction was again observed, *LR*(1) = 30.89, $p < .001$. No three-way interaction was observed, however, suggesting that the interaction between trial type and repeat condition did not differ between the replication and the 500-ms subset of the extension task, *LR*(1) = 0.30, $p = .58$. Additionally, there was an interaction between task and repeat condition, *LR*(1) = 9.39, $p = .002$. Mean proportions of *longer* responses were significantly larger in the high repeat condition of the extension compared to the replication, *LR*(1) = 10.92, $p = .002$, but no differences were observed between tasks for the low repeat condition, *LR*(1) = 1.07, $p = .30$ (Figure 5). No effect was observed for the interaction between task and trial type, *LR*(1) = 0.50, *p* = .48 or for the main effects of task, *LR*(1) = 2.36, *p* = .12, repeat condition, *LR*(1) = 0.39, *p* = .53, or trial type, *LR*(1) = 0.81, *p* = .37.

Effect of previous trials. In order to refine the interpretation of his results, Matthews (2015) conducted a supplemental analysis to determine whether the previous trial type influenced subjective duration under any of the observed conditions. Accordingly, we also analyzed the effect of previous trial type, repeat condition, and current trial type on the proportion of *longer* responses with our data. In our replication task, a main effect of previous trial type, *LR*(1) = 9.66, *p* = .002, and a repeat condition by trial type interaction, *LR*(1) = 11.76, *p* < .001, were qualified by repeat condition by trial type by previous trial type interaction, *LR*(1) = 11.13, *p* < .001 (Figure 6). No main effect of trial type, *LR*(1) = 0.0004, *p* = .99, or repeat condition, *LR*(1) = 1.01, *p* = .32, were observed, nor were the interactions of trial type and previous trial type, *LR*(1) = 0.69, *p* = .41, or repeat condition and previous trial type, *LR*(1) = 0.13, *p* = .71, significant.

Analysis of simple main effects suggested that within the low-repeat condition, trial type interacted with previous trial type, *LR*(1) = 8.17, *p* = 0.004. Mean proportion of longer responses was significantly higher for repeat trials following repeat trials than for repeat trials following nonrepeat trials, *LR*(1) = 11.89, *p* = .001, but no effect was observed for nonrepeat trials, *LR*(1) = 11.88, *p* = .65. Within the high repeat condition, a main effect of previous trial, *LR*(1) = 5.00, *p* = .025, was observed. Trials preceded by the same trial type (*M* = .47) were judged to be longer in duration than those preceded by a different trial type (*M* = .42). There was no difference in duration judgment between repeat and novel trials, *LR*(1) = 2.44, *p* = .12, and the interaction between current trial type and previous trial type was not significant, *LR*(1) = 2.88, *p* = .09.

No effect of repeat condition, *LR*(1) = 0.004, *p* = .95, trial type, *LR*(1) = 0.03, *p* = .86, or previous trial type, *LR*(1) = 0.17, *p* = .68, was found for our extension task, nor were any of the interaction terms significant (trial type by repeat condition: *LR*(1) = 1.84, *p* = .17; trial type by previous trial type: *LR*(1) = 1.83, *p* = .18; repeat condition by previous trial type: *LR*(1) = 0.04, *p* = .85; trial type by repeat condition by previous trial type: *LR*(1) = 0.80, *p* = .37).

Discussion

The present experiment aimed to replicate and further explore an effect reported by Matthews (2015) that the typically observed subjective shortening of duration for repeated stimuli can be modulated by participants' expectations of repetition. Using the comparison task employed by Matthews in his Experiments 1 and 4a and extending it to include a range of durations, we were able to observe differential effects of stimulus repetition based on global repetition contingencies. In both our replication and psychophysical extension tasks, participants' duration judgments for repeated stimuli depended on whether the probability of repetition was high or low.

Matthews (2015) hypothesized that this effect is perceptual, resulting from a mechanism whereby enhancements in perceptual processing correspond to increases in subjective duration. The expectation of repetition is proposed to strengthen the percept for a repeated stimulus, lengthening subjective duration (also see Matthews & Gheorghiu, 2016; Matthews & Meck, 2016). Bang and Rahnev (2017), however, recently demonstrated that expectation can shift decisional criteria without affecting perception, as Summerfield and Egner (2009) had previously proposed. The application of psychophysical methods in our extension tasks allowed us to conduct a preliminary assessment of whether expectation mitigates the effect of repetition by changing perceptual sensitivity or whether it induces a bias.

Primary Findings

Crucially, we found no evidence of a criterion shift, demonstrated by the lack of significant effects on B''_D and corroborated by a Bayes factor of 7.17 in favor of the null hypothesis. Thus, it does not appear that the observed effect has a decisional basis. Conversely, there was evidence that perception was affected by the manipulation of repetition probability, in that sensitivity (*A'*) was significantly lower in the high repetition condition than in the low repetition condition. Taken together, our psychophysical findings do provide support for a perceptual origin for the expansion of subjective duration that was observed for expected stimuli, in agreement with Matthews' (2015) hypothesis.

We note that, although we did observe that sensitivity differed based on repetition condition, the difference was small, and the Bayes Factor estimate was weakly in favor of the null model. Even so, a review of the relevant literature suggests that expectation's effects on sensitivity have been demonstrated previously, although the direction of the effect is inconsistent. Ulrich et al. (2006) reported that sensitivity was poorer for expected stimuli, in agreement with current findings. Birngruber et al. (2018) observed a difference in discrimination sensitivity (measured by the difference threshold) for expected stimuli relative to unexpected stimuli in one of two experiments, although they found better discrimination for expected stimuli, which contrasts with what we observed in the present investigation. Skylark and Gheorghiu (2017), however, observed that discrimination sensitivity was not affected by stimulus type, repetition probability, or the combination of the two when replicating Matthews' (2015) paradigm, although caution in directly interpreting these results may be in order. The difference of 33 ms between standard and comparison stimuli employed by Skylark and Gheorghiu was far smaller than the difference thresholds obtained in the current experiment (which ranged from 85 – 92 ms, see Appendix), suggesting that the difference may have been undetectable in the Skylark and Georghiu experiment. In fact, their reported hit rates (H: 38.9% - 48.9%) were much smaller and their false alarm rates (FA: 33.9% - 43.6%) much larger than what were obtained in the current experiment (H: 90% - 92%; FA: 7% - 8%), further suggesting that participants were unable to discriminate between stimulus durations in the Skylark and Gheorghiu experiment.

A perceptual origin for expectation's influence in the current paradigm corresponds with findings for the oddball effect (Birngruber et al., 2014). Expectation has been theorized to play a part in the oddball effect by suppressing duration for repeated stimuli through predictive coding processes (e.g., Friston, 2005), although Birngruber et al. (2018) have recently questioned expectation's role in this regard. The expansive properties of expectation observed in recent studies would rule out its involvement in the repetition contraction observed in the oddball paradigm. Nevertheless, a perceptual basis for both the current paradigm and the oddball effect, from which the current research stems to some extent, points to a mutual mechanism of action. It would, however, be hard to reconcile Matthews' (2015) proposal with the well-established oddball findings. This difficulty, coupled with our failure to observe the interaction effect between repetition condition and stimulus type for sensitivity that would be predicted by Matthews' hypothesis, suggests that the effect of expectation may operate through other perceptual means. Although expectation's role in the oddball effect is under question, we propose that predictive coding may still be able to offer a common perceptual mechanism for both oddball effects and our current findings.

Under predictive coding, context is a crucial factor in establishing the statistical regularities that allow for construction of prior expectations (Auksztulewicz & Friston, 2016; Clark, 2013; Friston, 2005). Manipulations affecting context should result in suppressed prediction error for stimuli conforming to contextual expectations, regardless of whether they are novel or repeats (Auksztulewicz & Friston, 2016). This phenomenon has been demonstrated by Richter et al. (2018) for expected objects that were not repetitions of the previous stimulus. Thus, the oddball effect may reflect anticipation of the oddball, with expectation expanding duration for that stimulus. The standards, on the other hand, seem shorter due to repetition.

Predictive coding also provides a convenient framework to explain how context (i.e., repetition condition) reversed the effect of repetition on the proportion of *longer* responses in the current experiment, in contrast to the reduction in effect observed by Matthews (2015). Consistent with predictive coding's theorized hierarchies of prediction, recent findings suggest that expectation operates on at least two levels neurally: a global level corresponding to the overall probability of encountering a stimulus given the context and a local level that makes predictions based on recent stimulus encounters (e.g., Todorovic & de Lange, 2012). Although these two levels of prediction are independent of one another, they do interact.

In the current paradigm, the local level tracks the most recent stimulus presentations, and the global level represents the overall probability of repetition or novelty within a given block. In our replication task, we observed that trials matching the global expectation were judged similarly, regardless of local context (Figure 5). From this perspective, our analysis considering the nature of the previous trial revealed evidence of such a hierarchical effect in our data. Duration judgments for the unexpected trial type were affected by local context. Initial presentations of a globally unexpected trial were judged to be shorter than that of an expected trial, but consecutive presentation expanded duration judgments for the subsequent trials. This demonstrates an updating of local expectations, in agreement with a hierarchical predictive coding account. Furthermore, the increase in subjective duration for consecutive unexpected trials resulted in mean duration judgments that were quite similar to those of the globally expected trials, suggesting that expectation exerted the same influence on subjective duration in the current task regardless of whether it was a local or global expectation that was met.

By what perceptual mechanism, however, might meeting expectations translate to the changes in subjective duration observed here and in the oddball effect under a predictive coding framework? One possibility is that the degree of distortion in temporal processing observed may be related to the magnitude of the top-down prediction error signals that arise from context violation. Lateral frontal regions and the anterior cingulate cortex have been implicated both in time perception (Schirmer, Meck, & Penney, 2016) and in the establishment of contextual learning, behaviour, and cognitive control (Alexander & Brown, 2015, 2018; Miller & Cohen, 2001). Error signals from one source (e.g., contextual learning) could affect other processes relying on the same networks. As such, violation of perceptual inferences, rather than suppression due to perceptual learning, may be responsible for the duration distortion observed for repeated stimuli in the low repetition blocks. This mechanism could also explain the reversal found when repetition was expected and the effect of updating local expectations described above. It could also potentially account for the oddball effect if it is the oddball that is anticipated. It is less clear whether Matthews' (2015) results are consistent with a relationship between error signals and subjective duration, although it is also not apparent that Matthews' suggestion of an enhancement of the percept for repeated stimuli can explain our findings.

Overall, our results support a perceptual basis for the observed increases in subjective duration for expected repeated stimuli. At the same time, it is obvious that additional research is needed. Given that we observed subjective duration to be longer for any stimulus meeting local or global expectations (whether it was repeated or novel), our findings suggest that the perceptual enhancement for the features of a recent stimulus that Matthews (2015) proposed might not be broad enough to account for our results. Our failure to observe an interaction effect for perceptual sensitivity that would support Matthews' hypothesized mechanism of action also suggests the need for continued investigation of the representational explanation. Further replication and examination of the extent of contextual influences is also warranted. Can we, for example, observe the same effect for other sensory modalities or for stimuli that share semantic or categorical relationships? It will also be important to determine with more certainty the means by which participants establish context, as the analysis on the effect of the previous trial reported both by Matthews (2015) and here suggested that it may be updated when as few as two trials have similar structure. Factors demonstrated to produce repetition enhancement, such as memory, attention, and recognition, should also be investigated for their potential influence in this paradigm.

Additional Results Requiring Explanation and Exploration

Although the discrepancy between our results and those of Matthews for the proportion of *longer* responses could be related to differences in the stimuli used, we believe it is unlikely, as Matthews' results held across stimulus types, task demands, and samples. Our results also evidence that the effect reported by Matthews may be reduced in the presence of genuine temporal differences in the stimulus intervals. It is possible, though, that the differences between duration intervals in our extension task were too large, and therefore too apparent to participants, suggesting that expectations do not influence time perception significantly enough to distort subjective duration for obviously different intervals. Rather, its impact may be limited to situations in which ambiguity is present. Furthermore, in the presence of uncertainty, we observed a tendency for participants to respond that the comparison was *shorter* in duration than the standard, regardless of stimulus type or expectation. This could indicate a global response bias for this type of task, as the tendency was observable in both our replication and in the 500-ms subset of the extension task, as well as being common in Matthews' (2015) experiments.

We also observed that duration judgments differed significantly between repetition conditions for the 200 ms interval but not for any other interval (Figure 3). In the low repeat condition, the difference between the comparison and standard may have been more obvious, making participants more likely to respond that the comparison stimulus was *shorter* in duration, and this response would have been further enhanced for repeat trials because of the contractive effects of repetition. When repetition was common, however, expectation expansion would have made the comparison appear to be more similar in duration to the standard for the (majority) repeat trials, resulting in a greater overall proportion of *longer* responses. Although this effect also occurs for other durations (e.g., see Figures 3 and 4), 200 ms might represent a “sweet spot” where the observed differences were maximal.

Conclusions

The present experiment tested Matthews’ (2015) perceptual hypothesis of the effect of expectation on stimulus repetition for time estimation by including an extension of Matthews’ original approach in which we could examine bias (criterion) and sensitivity. The possibility that the observed effects are perceptual was supported by an effect of our manipulations on sensitivity but not on criterion. The absence of such a criterion shift provides preliminary validation for research that assumes that the expansion of duration is perceptual rather than decisional. Accordingly, our study provides a meaningful contribution in supporting the validity of both past and future work. Additionally, our findings suggest the need for more research on the perceptual mechanisms of action responsible for this effect.

References

Alexander, W. H., & Brown, J. W. (2018). Frontal cortex function as derived from hierarchical predictive coding. *Scientific Reports, 8*, 3843. doi:10.1038/s41598-018-21407-9

Alexander, W. H., & Brown, J. W. (2015). Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation*, *27*, 2354-2410. doi:10.1162/NECO_a_00779.

Andics, A., Gál, V., Vicsi, K., Rudas, G., & Vidnyánszky, Z. (2013). FMRI repetition suppression for voices is modulated by stimulus expectations. *Neuroimage, 69*, 277–283. doi:j.neuroimage.2012.12.033

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, *80*, 125–140. doi:10.1016/j.cortex.2015.11.024

Bang, J. W, & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports, 7*, 17072. doi:10.1038/s41598-017-16885-2

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

Birngruber, T., Schröter, H., Schütt, E., & Ulrich, R. (2018). Stimulus expectation prolongs rather than shortens perceived duration: Evidence from self-generated expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 117. doi:10.1037/xhp0000433

Birngruber, T., Schröter, H., & Ulrich, R. (2014). Duration perception of visual and auditory oddball stimuli: Does judgment task modulate the temporal oddball effect?. *Attention, Perception, & Psychophysics*, *76*, 814-828. doi: 10.3758/s13414-013-0602-2

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, *42*, 286–291. doi:10.3758/BRM.42.1.286

Cai, M. B., Eagleman, D. M., & Ma, W. J. (2015). Perceived duration is reduced by repetition but not by high-level expectation. *Journal of Vision, 15*(13), 19:1–17. doi:10.1167/15.13.19

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181–204. doi:10.1017S0140525x12000477

De Rosario-Martinez, H. (2015). phia: Post-hoc interaction analysis [Computer software manual]. Retrieved November 2, 2017, from https://CRAN.R-project.org/package=phia (R package version 0.2-1)

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*, *7*, 145–147. doi:10.1016/S1364-6613(03)00055-X

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences*, *93*, 13494–13499.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*, 447–456. doi:10.1016/j.jml.2007.11.004

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, *121*, 275. doi:10.1037/0096-3445.121.3.275

Dosher, B. A., & Sperling, G. (1998). A century of human information-processing theory: Vision, attention, and memory. In J. E. Hochberg, (Ed.), *Perception and cognition at century's end* (pp. 199–252). San Diego: Academic Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. doi:10.3758/BF03193146

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*, 815–836. doi:10.1098/rstb.2005.1622

Green, D. M. & Swets, J. A. (1974/1966). *Signal detection theory and psychophysics*. Huntington, NY: Robert E. Krieger Publishing Company.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23. doi:10.1016/ j.tics.2005.11.006

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*, 561–582. doi:10.3758/APP.72.3.561

Hope, R. M. (2013). Rmisc: Ryan miscellaneous [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=Rmisc (R package version 1.5)

Jeffreys, H. (1961). *Theory of probability* (3rd Ed.). Oxford, UK: Oxford University Press.

Kadosh, R. C., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*, 132–147. doi:10.1016/ j.pneurobio.2007.11.001

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773-795.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian modeling for cognitive science: A practical course*. Cambridge, UK: Cambridge University Press.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122–1135. doi:10.3758/ s13428-014-0532-5

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Matthews, W. J. (2011). Stimulus repetition and the perception of time: The effects of prior exposure on temporal discrimination, judgment, and production. *PLoS one*, *6*(5), e19815. doi: 10.1371/journal.pone.0019815

Matthews, W. J. (2015). Time perception: The surprising effects of surprising stimuli. *Journal of Experimental Psychology: General*, *144*, 172. doi:10.1037/xge0000041

Matthews, W. J., & Gheorghiu, A. I. (2016). Repetition, expectation, and the perception of time. *Current Opinion in Behavioral Sciences*, *8*, 110–116. doi:10.1177/1948550616673877

Matthews, W. J., & Meck, W. H. (2016). Temporal cognition: Connecting subjective time to perception, attention, and memory. *Psychological Bulletin*, *142*, 865. doi:10.1037/ bul0000045

Mayrhauser, L., Bergmann, J., Crone, J., & Kronbichler, M. (2014). Neural repetition suppression: Evidence for perceptual expectation in object-selective regions. *Frontiers in Human Neuroscience*, *8*, 225. doi:10.3389/fnhum.2014.00225

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61-64.

Peirce, J. W. (2007). PsychoPy − psychophysics software in python. *Journal of Neuroscience Methods*, *162*, 8–13. doi:10.1016/j.jneumeth.2006.11.017

R Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved October 9, 2017, from http://www.R-project.org/

Richter, D., Ekman, M., & de Lange, F. P. (2018). Suppressed sensory response to predictable object stimuli throughout the ventral visual stream. *The Journal of Neuroscience*, *38,* 7452–7461. doi:10.1523/JNEUROSCI.3421-17.2018

Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli. *Perception, 24,* 1177–1187. doi:10.1068/p241177

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: a review of fMRI studies. *Neuropsychologia*, *51*, 59–66. doi:10.1016/j.neuropsychologia.2012.11.006

Schirmer, A., Meck, W. H., & Penney, T. B. (2016). The socio-temporal brain: Connecting people in time. *Trends in Cognitive Sciences, 20*, 760-772. doi: 10.1016/j.tics.2016.08.002

Skylark, W., & Gheorghiu, A. (2017). Further evidence that the effects of repetition on subjective time depend on repetition probability. *Frontiers in Psychology, 8*, 1915, 1–6. doi:10.3389/fpsyg.2017.01915

Song, Y., Nathoo, F. S., & Masson, M. E. (2017). A Bayesian approach to the mixed-effects analysis of accuracy data in repeated-measures designs. *Journal of Memory and Language*, *96*, 78-92. doi: 10.1016/j.jml.2017.05.002

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences, 13*, 403–409. doi:10. 1016/j.tics.2009.06.003

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*, 1004–1006. doi:10.1038/nn.2163

Summerfield, C., Wyart, V., Mareike Johnen, V., & De Gardelle, V. (2011). Human scalp electroencephalography reveals that repetition suppression varies with expectation. *Frontiers in Human Neuroscience, 5,* 67. doi:10.3389/fnhum.2011.00067

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *Journal of Neuroscience*, *32*, 13389–13395. doi:10.1523/JNEUROSCI.2227-12.2012

Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics*, *66*, 1171-1189. doi:10.3758/BF03196844

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychological research*, *70*, 77-87. doi:10.1007/s00426-004-0195-4

University of Stirling Psychology Department. (n.d.). *The Psychological Image Collection at Stirling (PICS)*. Retrieved June 30, 2017, from https://pics.stir.ac.uk

Vallesi, A., Binns, M. A., & Shallice, T. (2008). An effect of spatial-temporal association of response codes: Understanding the cognitive representations of time. *Cognition*, *107*, 501–527. doi:10.1016/j.cognition.2007.10.011

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*, 779-804. doi:10.3758/BF03194105

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488. doi:10.1016/j.tics.2003.09.002

Wiener, M., & Thompson, J. C. (2015). Repetition enhancement and memory effects for duration. *NeuroImage*, *113*, 268–278. doi:10.1016/j.neuroimage.2015.03.054

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception*, *44*, 289-300. doi: 10.1068/p7908

Appendix

Analysis for the point of subjective equality (PSE) and just noticeable difference (JND)

Logistic curves were fit for all stimulus conditions for each participant (average R^2 = .97; range = .59 - .99), from which PSE and JND values were extracted. The PSE is the duration at which participants respond *longer* 50% of the time, signifying that they perceive the duration of the comparison stimulus to be equal to that of the standard stimulus. The difference threshold is calculated as the difference between the durations at which *longer* responses are obtained 75% and 25% of the time, respectively, divided by two. The lower threshold value was not available for two participants (i.e., their lowest proportion of *longer* responses for any interval was larger than 25%). Accordingly, these participants were excluded from all analyses for both tasks. Psychometric functions for each of the four stimulus conditions are presented in Figure A1.

Results

Analysis of the point of subjective equality revealed a significant repeat condition by trial type interaction (Figure A2), *LR*(1) = 7.04, *p* = .008. The repeat trials had a higher mean PSE than the nonrepeat trials in the low repeat condition, and the nonrepeat trials had higher a mean PSE than the repeat trials in the high repeat condition, though these differences were not statistically significant (low repeat condition: *LR*(1) = 2.04, *p* = .17; high repeat condition: *LR*(1) = 2.97, *p* = .17). Analysis of the difference threshold produced no significant effects (repeat condition, *LR*(1) = 2.99, *p* = .08; trial type, *LR*(1) = 0.04, *p* = .84; repeat condition by trial type interaction, *LR*(1) = 1.19, *p* = .28).

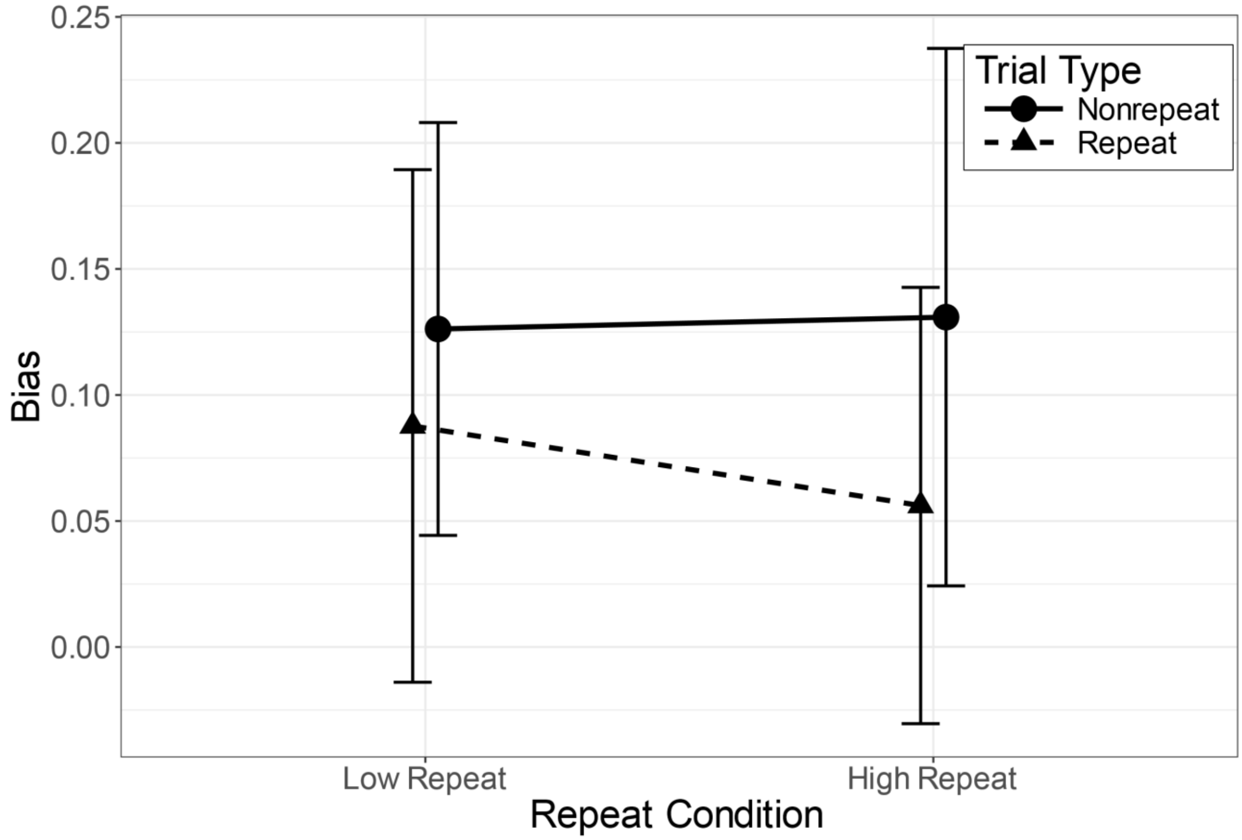

Figure 1. Mean bias estimates as a function of repeat condition and trial type. Bias ranges from -1 to +1, with 0 reflecting no bias. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).

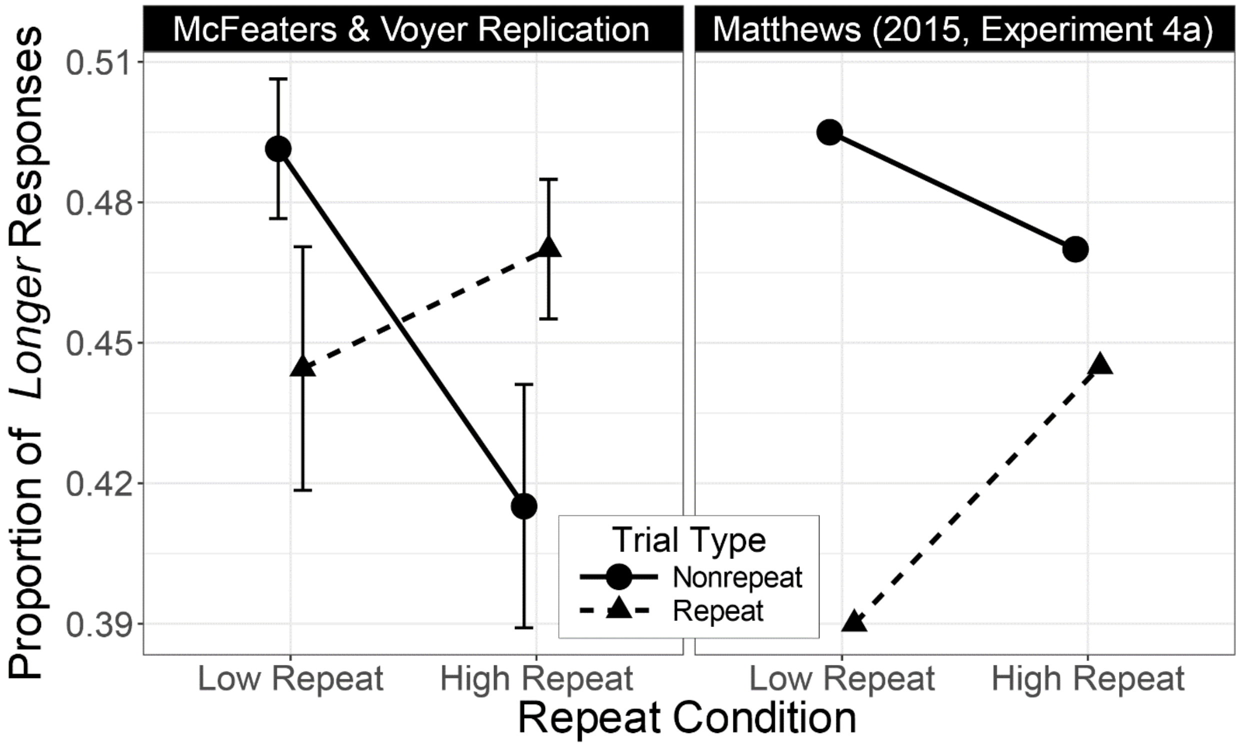

Figure 2. Mean proportion of *longer* responses as a function of repeat condition and trial type in the replication task. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008). For comparison purposes, Matthews' (2015) results from Experiment 4a are shown on the right. Adapted with permission from "Time perception: The surprising effects of surprising stimuli," by W. J. Matthews, 2015, *Journal of Experimental Psychology: General*, *144*(1), pp. 172 – 197. Copyright 2015 by The American Psychological Association.

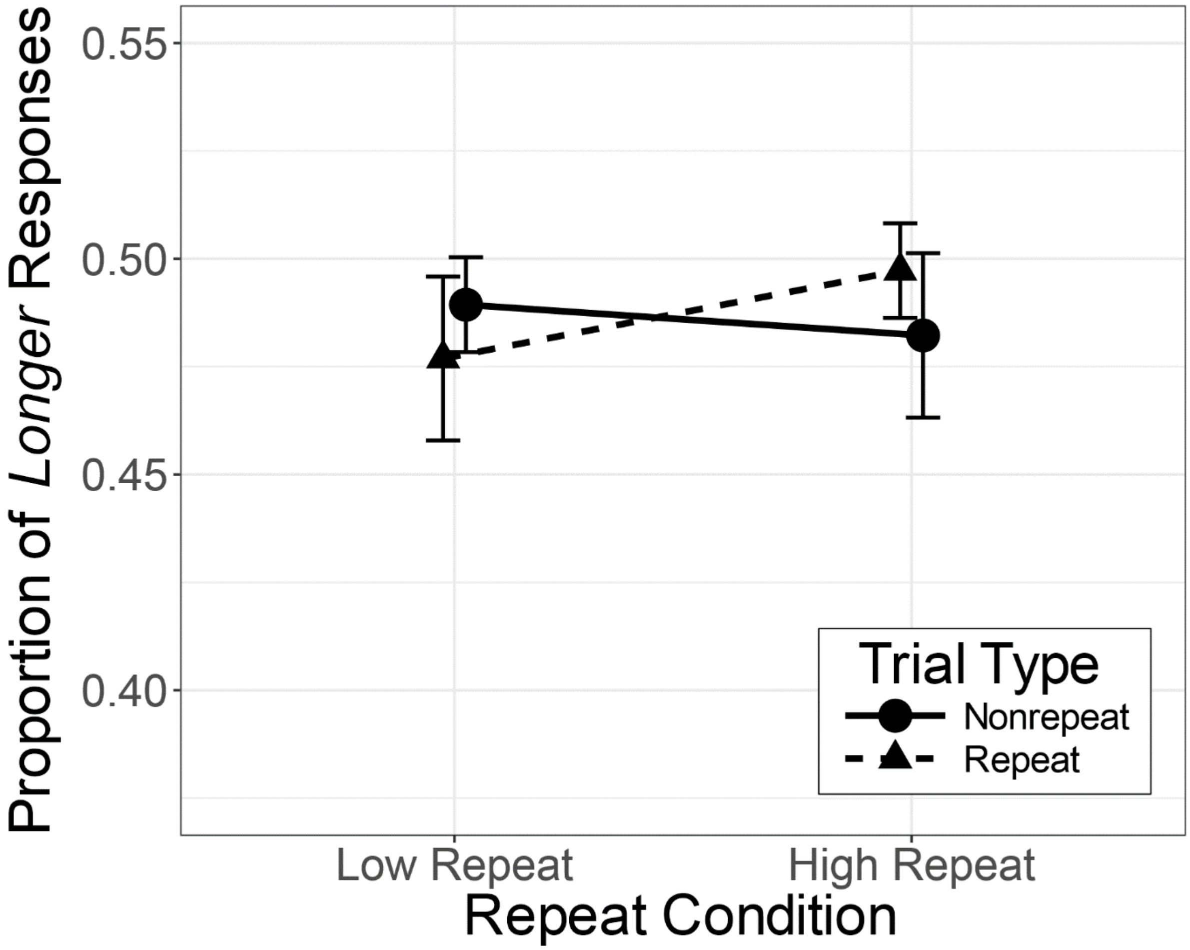

Figure 3. Mean proportion of *longer* responses as a function of repeat condition and trial type in the extension task. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).

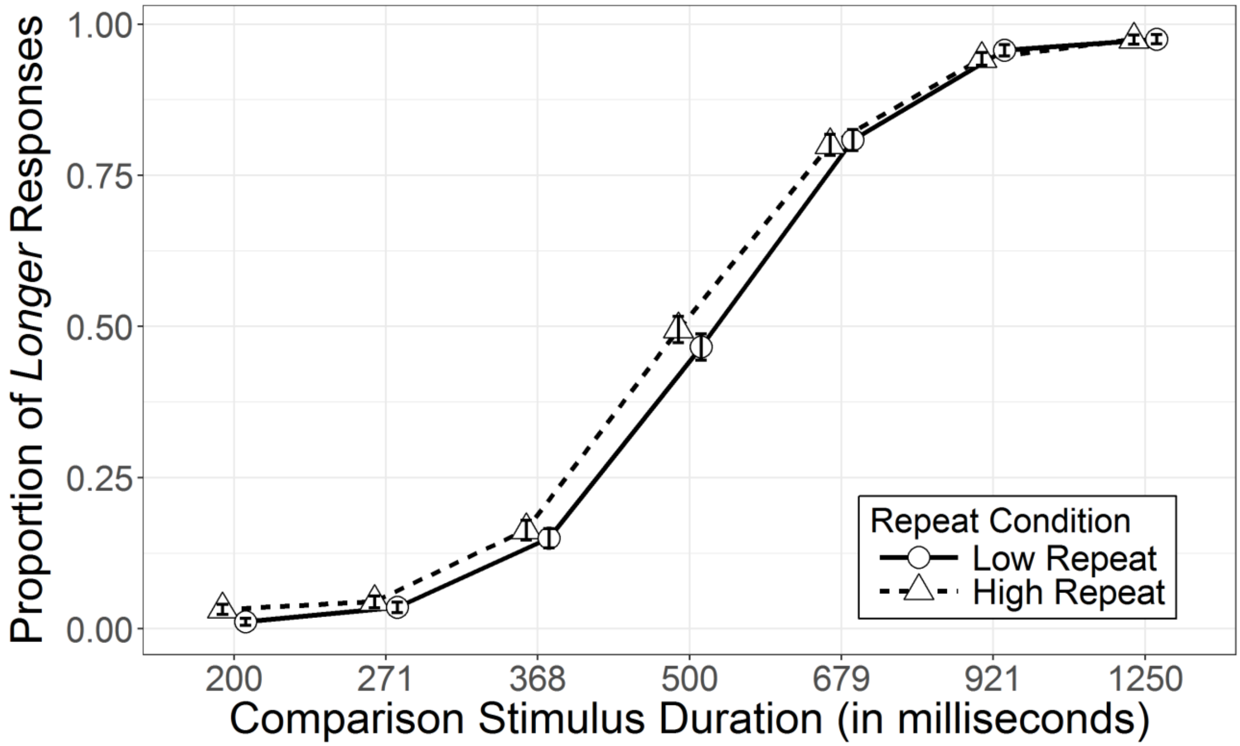

Figure 4. Mean proportion of *longer* responses as a function of repeat condition and stimulus duration in the extension task. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).

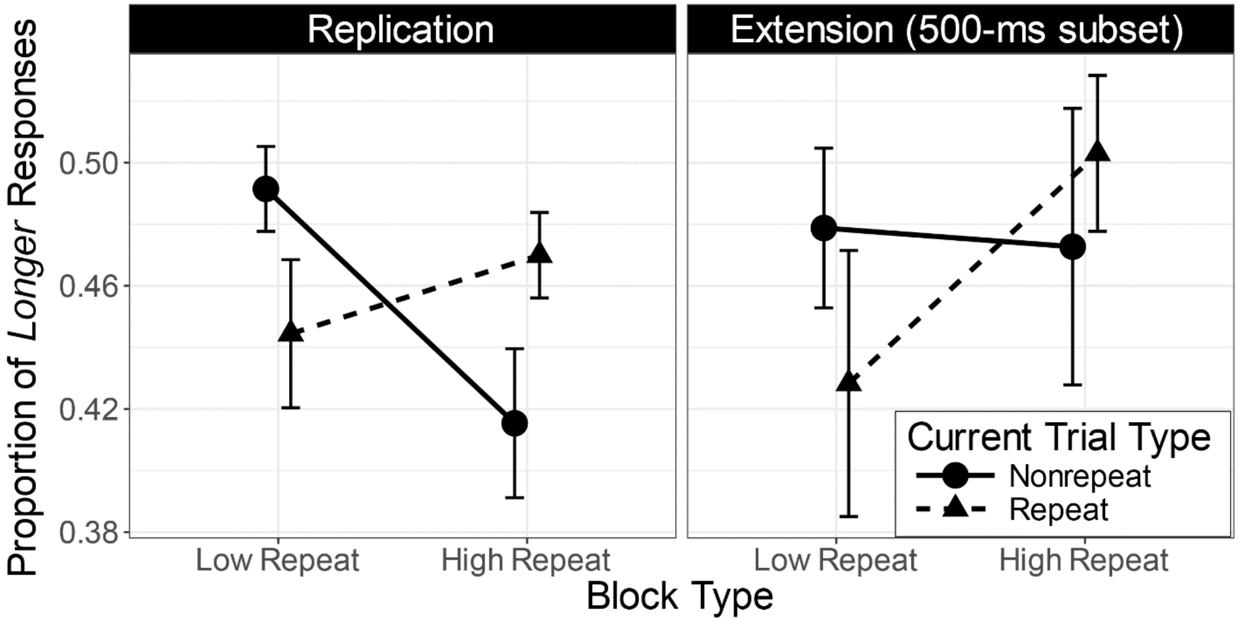

Figure 5. Mean proportion of *longer* responses for the 500-ms comparison duration as a function of repeat condition, trial type, and task. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).

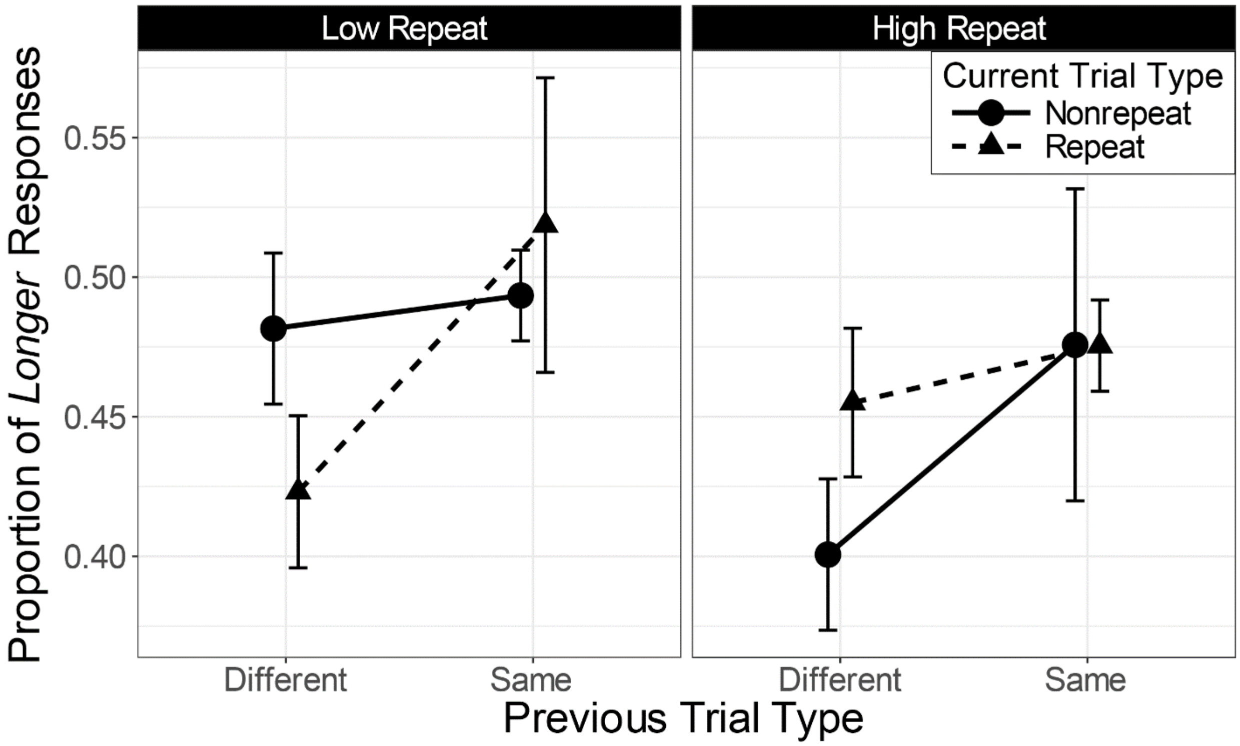

Figure 6. Mean proportion of *longer* responses by repeat condition, current trial type, and previous trial type for the replication task. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).

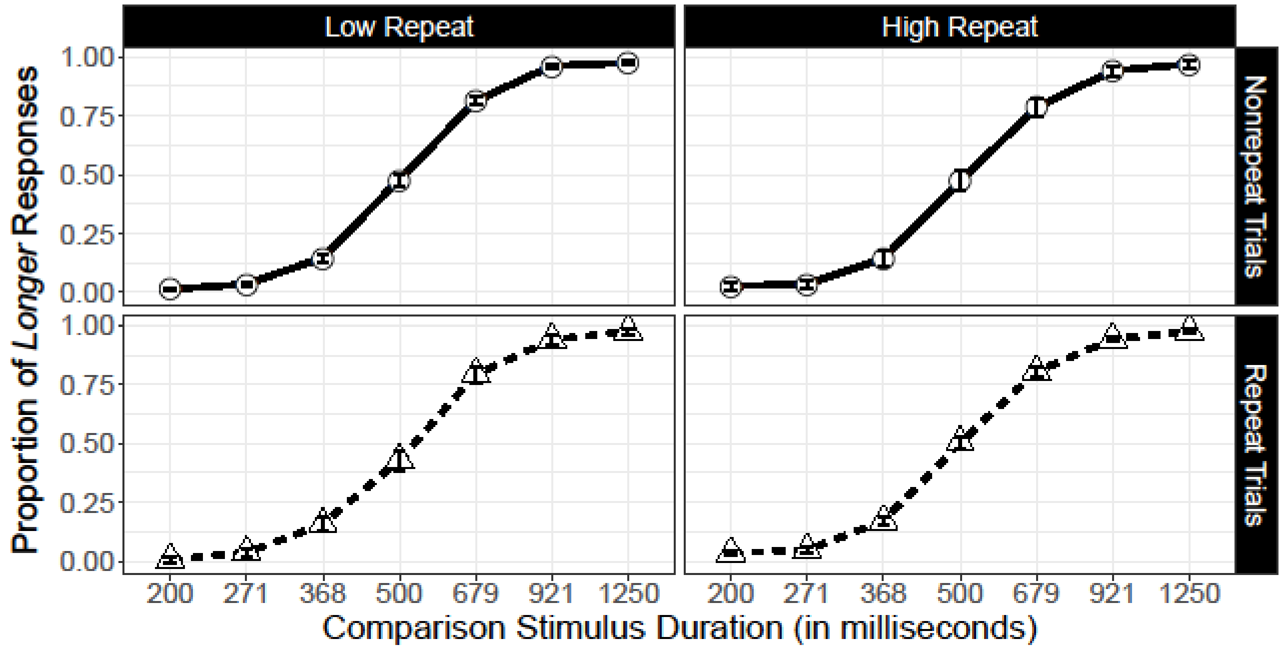

Figure A1. Mean proportion of *longer* responses as a function of repeat condition, trial type, and stimulus duration in the extension task. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).

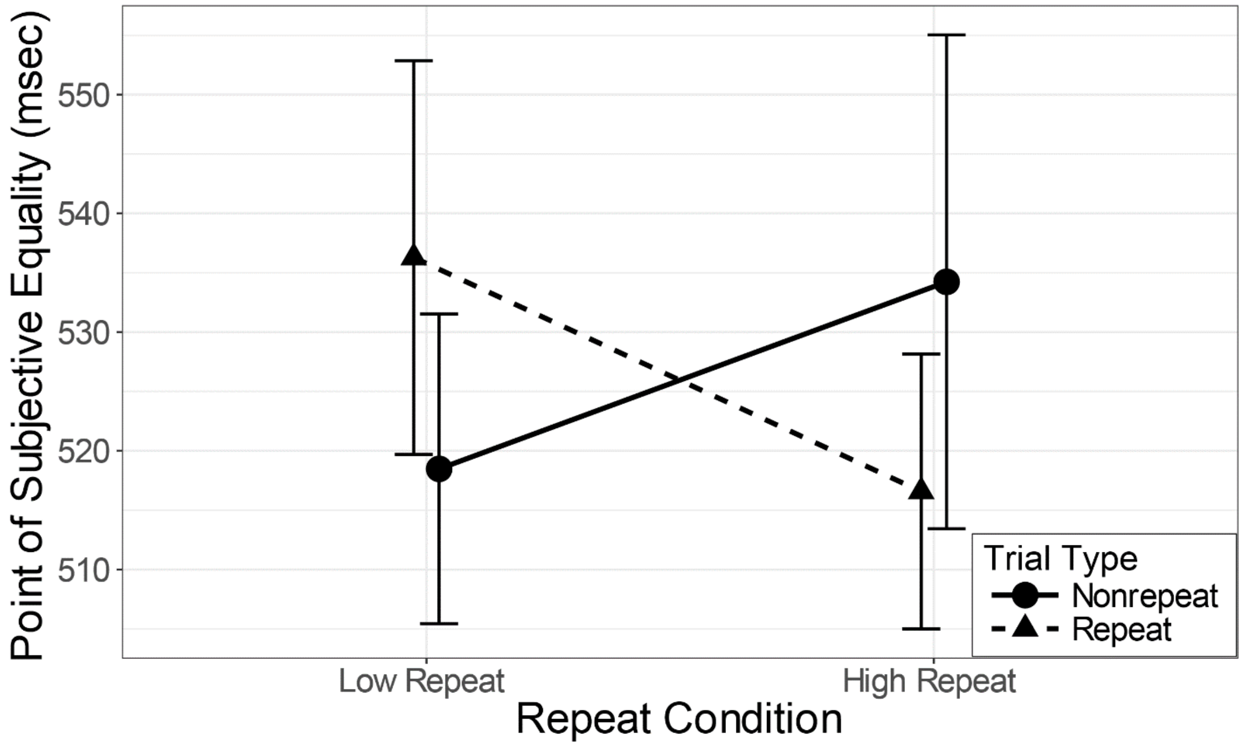

Figure A2. Mean PSE as a function of repeat condition and trial type. Error bars are 95% confidence intervals, calculated for within subjects designs, per Morey (2008).