

Stimulus Intensity and the Perception of Duration

William J. Matthews
University of Essex

Neil Stewart
University of Warwick

John H. Wearden
Keele University

RUNNING HEAD: Intensity and time perception

Address correspondence to:

William Matthews
Department of Psychology,
University of Essex,
Wivenhoe Park
Colchester
C04 3SQ
UK

Tel: +44 (0)1206 873818

Fax: +44 (0)1206 873801

Email: will@essex.ac.uk

Abstract

This article explores the widely-reported finding that the subjective duration of a stimulus is positively related to its magnitude. In Experiments 1 and 2 we show that, for both auditory and visual stimuli, the effect of stimulus magnitude on the perception of duration depends upon the background: Against a high intensity background, weak stimuli are judged to last longer. In Experiment 3 we show that the effect of intensity becomes more pronounced at longer durations, consistent with the idea that increasing stimulus intensity speeds up the pacemaker component of an internal clock, and that it is the difference of a stimulus from the background, rather than its absolute magnitude, which influences the rate of the pacemaker. These results urge a modification to the oft-repeated claim that more intense stimuli seem to last longer, and provide an important constraint on any model of human timing.

KEYWORDS: Time perception; temporal judgments; stimulus intensity; internal clock; scalar expectancy theory.

Stimulus intensity and the perception of duration

Every stimulus has a duration, and the accurate perception of time is essential to successful negotiation of the environment (e.g., Belin et al., 2002). Despite the critical role of time perception, it has long been known that various non-temporal aspects of the stimulus influence its apparent duration (e.g., Brown, 1995; Grondin, 1998; Rammsayer & Lima, 1991; Tse, Intriligator, Rivest, & Cavanagh, 2004; see Eagleman & Pariyadath, 2009, for a review). One particularly robust finding is that perceived duration depends upon the stimulus magnitude. For example, Goldstone, Lhamon, and Sechzer (1978) played participants a standard tone for one second followed by a comparison tone of variable duration, and asked them to indicate whether the latter tone was longer or shorter than the former. On some trials the standard was low intensity (64 dB) and the comparison stimulus was high intensity (76 dB); on other trials, the order of intensities was reversed. Goldstone et al. found that participants were more likely to perceive the comparison stimulus as longer than the standard when it was louder than when it was quieter, a pattern which has been replicated in many other studies (Berglund, Berglund, Ekman, & Frankenhaeuser, 1969; Zelkind, 1973).

Similar results have also been found with luminance (Goldstone & Goldfarb, 1964; Goldstone et al., 1978; Xuan, Zhang, He, & Chen, 2007), electrical stimulation (Ekman, Frankenhaeuser, Levander, & Mellis, 1966), vibrotactile stimulation (Ekman, Frankenhaeuser, Berglund, & Waszak, 1969), sizes of squares (Thomas & Cantor, 1975; 1976; Xuan et al., 2007) and number of dots (Dormal, Seron, & Pesenti, 2006; Xuan et al., 2007). Such effects would be trivial if they were limited to very brief stimulations (c. 50 ms) for which there is integration of duration and intensity (Fraisse, 1978). However, stimulus magnitude is positively related to perceived duration for intervals of several hundred milliseconds or more, leading to the widely-repeated view that “the more intense a stimulation, the longer the duration perceived” (Fraisse, 1978, p. 220) or, more colloquially, that “larger stimuli are judged to last longer” (Xuan et al., 2007, p.1; for other statements of this conclusion, see Allan, 1979; 1998; Fraisse, 1984; Goldstone et al., 1978; Wearden, Edwards, Fakhri, & Percival, 1998).

Over the years many different models of time perception have been developed (see Grondin, 2001, for a review), often emphasizing quite different tasks and timescales (Allan, 1979; Grondin, 2001; Fraisse, 1984). The effects of stimulus intensity on time perception have implications for any model of time perception, and have been central to the development of some accounts. For example, some theorists have posited a common

representational system for temporal and non-temporal magnitudes (e.g., Walsh, 2003; Xuan et al., 2007), presumably implying an intrinsic association between large or intense stimuli and long durations. More recently, Eagleman and Pariyadath (2009) have suggested that subjective duration depends on the efficiency of neural coding, with more intense stimuli seeming to last longer because they evoke a larger neural response. In the current article we shall be particularly concerned with one highly successful account of temporal perception and judgment: scalar expectancy theory (SET) (Gibbon, Church, & Meck, 1984). SET is an internal clock model (Creelman, 1962; Treisman, 1963) developed to account for animal learning and successfully applied to human timing behaviour (Allan, 1992; 1998; Wearden, 2003). Central to SET is the assumption of a pacemaker which emits pulses at some high, stochastic rate. The pacemaker is attached to an accumulator by a switch. When stimulus timing begins, the switch closes and pulses flow into the accumulator; at stimulus offset, the switch opens and the flow of pulses stops. The number of pulses stored in the accumulator provides the basis for judgments of stimulus duration. Within the framework of SET, stimulus intensity may influence time perception by changing the rate of the pacemaker (Zelkind, 1973), the latency to open and close the switch, or the variance of the switch latencies (Wearden et al., 1998).

Although the effect of stimulus intensity on time perception is well-documented in different modalities, and despite the existence of detailed and successful models of timing, the underlying basis for the effect of stimulus magnitude on perceived duration is unclear. The current paper addresses this shortcoming by asking two related questions. Firstly, what is it about more intense stimuli that causes them to have longer apparent duration? In particular, is it the absolute magnitude of the stimulus or its difference from the background context that drives the effect? And secondly, how does stimulus intensity exert its effects? That is, which aspects of the timing process are influenced by stimulus magnitude? Experiments 1 and 2 address the former question; Experiment 3 addresses the latter.

Experiment 1

Despite the large body of evidence demonstrating that high intensity stimuli are judged to last longer than low intensity stimuli, there are two reasons to doubt whether this is always the case.

The first reason comes from the general argument that perception is relational (e.g., Lockhead, 1992). Judgments are influenced by context, and depend upon the immediate

background against which stimuli are presented (e.g., Laming, 1997, pp. 83-86), the item presented on the immediately preceding trial (e.g., Matthews & Stewart, 2009a, 2009b; Ward, 1979) and the distribution of other stimuli presented in the experiment (e.g., Lockhead & Hinson, 1986; Matthews & Stewart, 2008; Parducci, 1965). Many researchers have taken these kinds of results to imply that it is the differences between stimuli (or between stimulus and background) and not absolute magnitudes which determine our sensory judgments and identification of stimuli (e.g., Laming, 1997; Stewart & Matthews, 2009; Stewart, Brown, & Chater, 2005).

The second, more specific reason comes from studies of time perception which have demonstrated the importance of stimulus differences. For example, Xuan et al. (2007) asked people to compare two durations defined by the gaps in a sequence of digits. The perceived duration of a gap was longer when the jump between the flanking digits was larger, indicating that the perception of an interval depends upon the difference between the markers. Similarly, when a stimulus is an “oddball” in an experiment (that is, when it differs from the background established by the other stimuli in the experiment), its apparent duration increases (Tse, Intriligator, Rivest, & Cavanagh, 2004; Ulrich, Nitschke, & Rammsayer, 2006).

There is therefore good reason to think that the relative magnitude rather than the absolute intensity of a stimulus will be the primary determinant of its apparent duration. In the long list of experiments cited above, wherein increasing stimulus magnitude increased apparent duration, the high-intensity stimuli were also the ones which were most different from the background context. That is, absolute stimulus magnitude was confounded with the difference between the stimulus and the background. Our first two experiments ask whether the effect of absolute stimulus magnitude on the perception of duration persists when low intensity stimuli are most different from the background.

Experiment 1 employed a design used in previous studies of time perception (e.g., Goldstone et al., 1978) and used visual stimuli. On each trial the participant was shown two squares, one after the other, and judged whether the duration of the second, comparison stimulus was longer or shorter than that of the first, standard stimulus. On half of the trials the standard stimulus was a light grey square and the comparison stimulus was a dark grey square; on the other half of trials the order was reversed. Participants completed the task in two conditions, one in which the stimuli were shown against a white background and another in which the stimuli were shown against a black background. If absolute intensity determines subjective duration then the bright square should seem longer irrespective of the background;

if it is the difference from the background that matters, the bright square should seem longer against a dark background but shorter against a white background.

Method

Participants. Participants were members of the University of Warwick participant panel. In order to check that participants understood and were motivated to complete the task, we replaced any participants whose performance was not significantly above chance (assessed with a two-tailed binomial test). Across Experiments 1 and 2, a total of 13 participants were replaced. (Their inclusion made no difference to the pattern of significant results from either experiment.) Twenty four above-chance participants completed this experiment. Each was paid £6.

Stimuli. The stimuli were grey squares measuring 150 x 150 pixels (approximately 4.9° visual angle) displayed in the centre of a 19" CRT screen (1024 x 768 pixels) viewed from c. 60 cm. The *dim* square was dark grey (c. 1.8 cd/m²); the *bright* square was light grey (c. 98 cd/m²). On each trial, the two squares were shown in succession for different durations. The first stimulus acted as a standard against which the second, comparison stimulus was to be judged. Two different standard durations were used: 600 ms and 800 ms. For each standard there were four different comparison durations: for the 600-ms standard these were 450, 560, 640 and 750 ms; for the 800-ms standard they were 650, 760, 840 and 950 ms.

Design and procedure. Participants completed two sessions, one with a black background (c. 0.5 cd/m²) and the other with a white background (c. 128 cd/m²). Each session consisted of 16 practice trials followed by three blocks of 80 test trials. The sequence of events on each trial is illustrated in Figure 1, and was as follows: 1000-ms blank interval (during which nothing was shown on the screen), 500-ms fixation cross in the centre of the screen, 500-ms blank interval, the standard stimulus, a variable blank interval of 450-650 ms, and finally the comparison stimulus. The screen then went blank and participants indicated their judgment by pressing one key if the comparison duration was shorter than the standard and another if it was longer. The background (black or white) was constant throughout a session and was visible before, during, and after stimulus presentation (see Figure 1). In each block, each standard duration-comparison duration combination was presented equally often. On half of the trials the standard stimulus was the bright square and the comparison stimulus was the dim square; on the other half the order was reversed. Trial order was randomized.

The order in which participants completed the sessions and the assignment of responses to response keys were counterbalanced. Stimulus presentation was controlled by DMDX (Forster & Forster, 2003) in this and Experiment 3; Experiment 2 used custom-written software.

Results

The dependent measure was the proportion of trials on which the participant judged the duration of the comparison stimulus to be longer than the standard. The results are shown in Figure 2.

As can be seen in the figure, the effects of stimulus luminance on perceived duration were opposite for the black and white background conditions. This impression is supported by a within-subjects ANOVA which indicated a significant four-way interaction between background (black, white), standard duration (600 ms, 800 ms), difference between comparison and standard durations (-150 ms, -40 ms, +40 ms, +150 ms) and luminance of comparison stimulus (dim, bright), $F(2.3, 53.7) = 6.71, p = .002, \eta_p^2 = .23$ (here and at certain points below we applied a Huynh-Feldt correction because of violations of sphericity; in all analyses, α was set to .05). We therefore considered separately the results for each of the two standard durations with each of the two backgrounds. That is, we performed separate within-subject ANOVAs for each of the four panels in Figure 2. [In this experiment and the next, we also estimated the constant error by maximum likelihood fitting of a logistic psychometric function to the data from each participant in each condition (Mattes & Ulrich, 1998). Examination of the constant errors revealed the same pattern as that indicated by the ANOVAs reported below.]

For the black background condition with a 600-ms standard there was a main effect of comparison duration, $F(2.6, 58.8) = 88.37, p < .001, \eta_p^2 = .79$. As one would expect, increasing the comparison duration increased the probability that it would be judged longer than the standard. The effect of stimulus luminance missed significance, $F(1,23) = 2.98, p = .098, \eta_p^2 = .12$, and there was no interaction, $F(3,69) = 2.26, p = .089, \eta_p^2 = .09$. For the 800-ms standard there was again a main effect of comparison duration, $F(2.0, 45.3) = 58.07, p < .001, \eta_p^2 = .72$. Here, there was also a main effect of stimulus luminance, $F(1,23) = 7.45, p = .012, \eta_p^2 = .25$, with a higher proportion of “longer” responses when the comparison stimulus

was brighter than the standard. There was no interaction between comparison duration and luminance, $F(3,69) = 1.10, p = .354, \eta_p^2 = .05$.

For the white background condition with the 600-ms standard there was a main effect of comparison duration, $F(2.4, 54.8) = 121.77, p < .001, \eta_p^2 = .84$. More importantly, there was a main effect of stimulus luminance, $F(1,23) = 14.71, p = .001, \eta_p^2 = .39$; the proportion of “longer” responses increased when the comparison stimulus was darker than the standard. There was also a significant interaction between comparison duration and luminance, $F(3,69) = 4.19, p = .009, \eta_p^2 = .15$. Inspection of Figure 2 suggests that the effect of stimulus luminance increased as the comparison duration lengthened. For the 800-ms standard, there was again a significant effect of comparison duration, $F(2.0, 45.3) = 66.28, p < .001, \eta_p^2 = .74$, and a main effect of luminance, $F(1,23) = 16.25, p = .001, \eta_p^2 = .41$. As for the 600-ms standard, dim stimuli were judged to last longer than bright ones. There was also a significant interaction between comparison duration and luminance, $F(3,69) = 2.80, p = .046, \eta_p^2 = .11$.

Discussion

These results show that the effect of stimulus luminance on judgments of duration depends upon the context. Against a black background, a bright stimulus is judged to last longer than a dim one. This is the result which has been reported in numerous previous studies of time perception (e.g., Goldstone et al., 1978). However, when the background is white, the effect is reversed; using the same stimuli, the darker stimulus is judged to last longer than the bright one.

As an additional observation, we note that the comparison stimuli in the 600-ms standard conditions tended to be underestimated but that comparison stimuli in the 800-ms conditions tended to be overestimated. It seems likely that these results are instances of the time order error (TOE, Hellström, 1985). When a standard stimulus is followed by a comparison item, the latter is routinely found to be systematically overestimated (a negative TOE) or underestimated (a positive TOE). The magnitude and direction of the TOE depends upon a number of factors, including the interstimulus interval and the duration of the stimuli. Previous studies (see, e.g., Hellström & Rammsayer, 2004) have shown that when, as in the current experiments, standards of varying lengths are intermixed in an experimental session,

shorter standards elicit a positive TOE whilst longer standards elicit a negative TOE. The current results therefore match this body of research, and presumably result from the same perceptual mechanisms (see Hellström, 1985; 2003, for a discussion).

Although the observed time order errors are interesting, they in no way compromise the primary conclusion of Experiment 1. The TOE effects comprise an overall tendency to over- or under-estimate the duration of the comparison stimulus. Irrespective of these tendencies, presenting the stimuli against a black background meant that the comparison stimulus was more likely to be judged longer than the standard when it was brighter, and more likely to be judged shorter when it was dimmer. Against a white background, the opposite was true.

Experiment 2

Experiment 2 replicated the visual effect in Experiment 1 with auditory stimuli. High intensity and low intensity sounds were presented which interrupted a background context of noise that was either loud or quiet. If perceived duration depends upon absolute magnitude then the loud stimulus ought to appear to last longer irrespective of the context; if it is the magnitude of the change in intensity that matters then loud sounds will seem to last longer when presented in a quiet context, but quiet sounds will seem longer when presented in a loud context.

Method

Participants. Twenty four above-chance participants took part. Each was paid £7.

Stimuli. The stimuli were bursts of white noise, a quiet stimulus (c. 46 dBA) and a loud stimulus (c. 68 dBA). There was a single standard duration of 600 ms and four comparison durations of 450, 570, 630 and 750 ms. Stimuli were presented diotically over Sennheiser HD 265 headphones.

Design and procedure. Participants were tested individually in a quiet testing cubicle. Each participant performed in two conditions, a quiet background condition and a loud background condition, completed in two sessions conducted back-to-back. Each session comprised 16 practice trials followed by four blocks of 48 test trials. Each trial began with a

1000-ms blank (silent) interval followed by the word “Ready” appearing on the screen for 500 ms. White noise was then played for 1000 ms to establish context: In the quiet background condition this noise was at c. 42 dBA; in the loud background condition it was at c. 72 dBA. The standard stimulus was then presented for 600 ms, followed by a variable ISI of context noise for 700-800 ms after which the comparison stimulus was presented. Context noise then played until the participant made a response. Each of the four comparison durations were presented equally often. On half of the trials in each block the standard stimulus was loud and the comparison stimulus quiet, for the other half the order of loudnesses was reversed. Trial order was randomized within each block. Participants were told that they could take as long as they liked over each response, and to respond as quickly and accurately as possible once the comparison stimulus had ended. The order in which the two conditions were completed and the assignment of keys to responses were counterbalanced. As in Experiment 1, stimuli were well above threshold and we made sure that participants understood the task before they began.

Results

The results are shown in Figure 3. A within-subjects ANOVA revealed a significant three-way interaction between context (quiet background, loud background), comparison stimulus duration (450, 570, 630, 750 ms) and loudness of standard and comparison stimuli (quiet standard and loud comparison vs. loud standard and quiet comparison), $F(3,69) = 30.03, p < .001, \eta_p^2 = .57$. We therefore conducted separate 2 x 2 within subject ANOVAs for the quiet background and loud background conditions.

In the quiet background condition there was a significant main effect of comparison duration, $F(1.8, 42.5) = 204.86, p < .001, \eta_p^2 = .90$. There was also a significant main effect of stimulus loudness, $F(1,23) = 78.20, p < .001, \eta_p^2 = .77$, and a significant interaction, $F(2.7, 62.6) = 26.95, p < .001, \eta_p^2 = .54$. When the comparison stimulus was louder than the standard, the probability that it would be judged longer increased; this effect was greatest for the two most difficult discriminations.

In the loud background condition there was again a main effect of comparison duration, $F(1.3, 30.1) = 38.19, p < .001, \eta_p^2 = .62$, a main effect of stimulus loudness, $F(1,23) = 4.72, p = .040, \eta_p^2 = .17$, and a significant interaction, $F(2.6, 59.1) = 6.00, p = .002, \eta_p^2 =$

.21. In this condition the effect of stimulus loudness was the opposite of that in the quiet background condition. When the comparison stimulus was louder than the standard, it was more likely to be perceived as having shorter duration.

As in Experiment 1, the effect of stimulus magnitude on perceived duration depends upon the experimental context. When the background intensity is low, loud stimuli are judged to last longer, as found in previous studies (e.g., Berglund et al., 1969). However, when the background intensity is high, quiet stimuli are perceived to have longer duration. Inspection of Figure 3 suggests that the effect of stimulus intensity may be somewhat weaker for the loud background condition. One plausible explanation for this is that the context was less well established in this condition: There was a silent interval of 1.5 s between trials, which is much more similar to the quiet background than the loud background.

Discussion

Many studies have found that increasing the non-temporal magnitude of a stimulus increases its apparent duration, (e.g., Berglund et al., 1969; Dormal et al., 2006; Ekman et al., 1966; 1969; Goldstone & Goldfarb, 1964; Goldstone et al., 1978; Xuan et al., 2007; Zelkind, 1973) leading to the empirical generalization that, across different sensory modalities, more intense or larger stimuli are judged to last longer than less intense or smaller ones (e.g., Allan, 1979; 1998; Fraisse, 1978; 1984; Goldstone et al., 1978; Wearden et al., 1998). Experiments 1 and 2 support this conclusion when the background intensity is low, but show that the reverse pattern obtains when the background intensity is high; against a high-intensity background, weak stimuli are judged to last longer. This is true for both visual and auditory stimuli. Our experiments therefore suggest that the empirical generalization that larger stimuli last longer should be modified: The perceived duration of a stimulus depends positively upon how different it is from the background.

In the remaining experiments, we investigate how intensity (or intensity difference) influences temporal perception. Within the framework of scalar expectancy theory, stimulus intensity may influence the switch component, the rate of the pacemaker, or both. Suppose a stimulus is presented for a certain duration t . At stimulus onset, the switch closes with latency l_c , allowing the flow of pulses into the accumulator; at offset, the switch opens with latency l_o . If the pacemaker produces pulses with rate r , the total number of pulses accumulated during stimulus presentation will be:

$$N = r(t - l_c + l_o) \quad (1)$$

Subjective differences in duration might result from a difference in relative switch latencies. If one type of stimulus causes the switch to close more quickly, or to open more slowly, then it will have longer subjective duration. (Note that it is the difference in switch opening/closing times that is crucial; if both latencies are increased by a constant amount, there will be no change in apparent duration.) Alternatively, differences in subjective duration might result from a change in the rate of the pacemaker, with a faster rate yielding longer subjective duration. Finally, changes in switch variability will increase or decrease the variability of judgments without affecting their average.

We can decompose the effects of switch latencies and pacemaker rate by examining how subjective duration varies with objective stimulus time. The expression giving the total number of accumulated pulses can be written as two additive components, $rt + r(l_o - l_c)$. If two stimuli differ in their switch latencies then this will be manifested as a difference in the intercept of the function relating subjective duration to objective time. If two stimuli differ in the rate of the pacemaker then there will be differences in the slope of the line, with the difference between the stimuli becoming more pronounced at longer times. There may also be a difference in intercept, provided that the difference between switch opening/closing latencies is not zero.

This logic has been used to determine the basis for several factors that affect subjective duration. For example, Wearden et al. (1998) examined the long-established finding that tones are judged to last longer than lights of the same duration. Participants made verbal estimates of the durations of tones and lights lasting 77-1183 ms. Wearden et al. found that when mean judgments were plotted against objective durations, the line for auditory stimuli was steeper than that for visual stimuli, indicating a higher pacemaker rate; there was no difference in intercept. Similarly, Wearden et al. found that playing a series of clicks prior to stimulus presentation produced an increase in slope, consistent with a speeding up of the internal pacemaker (Penton-Voak et al., 1996). Similar results have been found in an investigation of the difference between filled and empty intervals (Wearden et al., 2007).

Experiment 3 addresses two questions. Firstly, we ask which component(s) of the internal clock are affected by stimulus intensity. Previous studies have shown that stimulus modality, click trains, and whether the interval is filled or empty all seem to influence the rate

of the pacemaker (Wearden et al., 1998; 2006; 2007), and it has been suggested that stimulus intensity also influences pacemaker rate (Zelkind, 1973). However, it also seems plausible that switch latencies will be particularly affected by stimulus intensity; intensity effects may be solely due to differences in onset/offset detection time. Secondly, we ask whether it is absolute stimulus magnitude or difference from the background which influences the clock. Experiments 1 and 2 suggest that it is the difference of a stimulus from the background, rather than the absolute intensity, which affects subjective duration. One possibility is therefore that the rate of the pacemaker is higher for stimuli which are more different from the background. Alternatively, it may be that the pacemaker runs faster for large stimulus magnitudes even when they are presented against a high-intensity background, but that there is a difference in the relative switch opening/closing latencies which offsets this effect at the durations studied in Experiments 1 and 2. In Experiment 3, we clarified the effects of absolute and relative stimulus magnitude by employing a version of the verbal estimation task in which the background context was varied between participants. We choose to focus on visual stimuli because it is easier to manipulate the background context for this modality.

Experiment 3

Method

Participants. Sixty four participants completed the experiment for a payment of £3.

Stimuli. The stimuli were grey squares measuring 150 x 150 pixels displayed on a 19" CRT monitor with a 1280 x 1024 pixel resolution and an 85 Hz refresh rate, viewed from a distance of approximately 60 cm. The stimuli subtended a visual angle of approximately 3.9 degrees. One square was dim (c. 1.2 cd/m²) and the other was bright (c. 104.8 cd/m²). Stimuli were either shown against a white background (c. 117.5 cd/m²) or against a black background (c. 0.1 cd/m²). Stimuli were presented for six different durations similar to those used by Wearden et al. (2007): 82, 353, 576, 765, 953, and 1176 ms.

Design and procedure. Half of the participants were in the black background condition; half were in the white background condition. There were four blocks of 24 trials. Each block comprised two presentations of each stimulus at each duration, in random order. Participants were tested individually in quiet testing cubicles. On each trial, the participant

saw one of the squares and judged its duration, in milliseconds. They were told that all durations were in the range 50 to 1500 ms, but that this did not necessarily mean that the shortest duration was 50 ms and the longest 1500 ms. Participants typed their judgments on the computer keyboard and were able to clear their entry and re-type their judgment if they wished. They then pressed a key to register their response; after a pause of 1.0 - 2.2 s, the next stimulus was presented. Participants were instructed to attend to the screen at the point when they pressed the button to register their judgment, to ensure that they would not miss the next stimulus. After the participant entered his or her response there was a pause of 1.4 - 2.0 s before the next item. Trials on which the participant failed to enter a response, or entered a response less than 50 ms or greater than 1500 ms, were assumed to indicate typing errors and were discarded (Wearden et al., 1998; 2007). Two participants (one from each condition) were discarded because fewer than 90% of trials produced valid data. (The proportion for the remaining 62 participants was 99%.)

Results

The mean judgments are shown in the top panels of Figure 4. An initial 2 x 6 x 6 mixed ANOVA revealed a significant three-way interaction between background, stimulus brightness, and stimulus duration, $F(4.2, 250.7) = 10.33, p < .001, \eta_p^2 = .15$. We therefore conducted separate analyses for the black and white background conditions.

In the black background condition, judgments increased with stimulus duration, $F(2.3, 69.3) = 376.75, p < .001, \eta_p^2 = .93$. Bright stimuli were judged to last longer than dim ones, $F(1,30) = 25.47, p < .001, \eta_p^2 = .46$, and this effect was more pronounced for longer durations, $F(3.9, 116.4) = 7.69, p < .001, \eta_p^2 = .20$.

In the white background condition, judgments again increased with stimulus duration, $F(2.0, 60.3) = 217.85, p < .001, \eta_p^2 = .88$, but dim stimuli were judged to last longer than bright ones, $F(1,30) = 18.59, p < .001, \eta_p^2 = .38$, and this effect became more pronounced at longer durations, $F(4.1, 122.7) = 4.44, p = .002, \eta_p^2 = .13$.

Following the example of previous studies (e.g., Wearden et al., 1998), we regressed judgment on stimulus duration for each participant, calculating separate intercept and slope coefficients for dim and bright stimuli. For the slope coefficients, there was a significant

interaction between background and stimulus, $F(1,60) = 27.30, p < .001, \eta_p^2 = .31$, so we used paired-sample t tests to compare the slope coefficients for the bright and dim stimuli separately for each background. Against a black background the slope was greater for bright stimuli ($M = 0.79, SD = 0.18$) than for dim stimuli, ($M = 0.70, SD = 0.19$), $t(30) = 4.26, p < .001, d = 0.77$. Against a white background, the reverse was true: The slope for the dim condition ($M = 0.75, SD = 0.24$) was greater than that for the bright condition ($M = 0.67, SD = 0.22$), $t(30) = 3.20, p = .003, d = 0.58$. (Note that some caution is needed when applying the regression analysis to the white background condition because there is a rather abrupt jump in judgments at c. 800 ms.)

The intercepts did not differ between the four conditions (dim stimuli on black background, $M = 79.33$ ms, $SD = 154.86$ ms; bright stimuli on black background, $M = 58.79$, $SD = 154.86$; dim stimuli on white background, $M = 129.80$ ms, $SD = 152.39$; bright stimuli on white background, $M = 136.96$ ms, $SD = 156.22$ ms). A 2 x 2 mixed ANOVA revealed no effect of stimulus, $F(1,60) < 0.65, p = .423, \eta_p^2 = .01$, no effect of background, $F(1,60) = 2.98, p = .090, \eta_p^2 = .05$, and no interaction, $F(1,60) = 2.97, p = .100, \eta_p^2 = .04$.

We also examined the absolute and relative variability of judgments. In pacemaker-accumulator models, the variance in the number of counts accumulated is positively related to the objective duration and is expected to be higher when the pacemaker runs at a faster rate (see Rammsayer & Ulrich, 2001, for a general treatment), although in the case of verbal estimation tasks the effect on judgment variability is complicated by the need to convert raw counts into numerical estimates. (See Wearden et al., 2007, for a large body of data and a comprehensive discussion of judgment variability in the paradigm used here). The middle panels of Figure 4 show the standard deviations of the judgments. An initial ANOVA indicated a significant interaction between background condition and stimulus brightness, $F(1,60) = 5.27, p = .025, \eta_p^2 = .08$, so we analyzed the results for the two conditions separately. In the black background condition, judgments became more variable at longer durations, $F(4.0, 120.8) = 54.27, p < .001, \eta_p^2 = .64$. However, there was no effect of stimulus brightness, $F(1,30) = 3.11, p = .088, \eta_p^2 = .09$, and no interaction between brightness and duration, $F(4.2, 125.0) = 1.03, p = .399, \eta_p^2 = .03$. The same pattern was found in the white background condition [for effect of duration, $F(3.3, 98.7) = 31.49, p < .001, \eta_p^2 = .51$; for brightness, $F(1,30) = 2.21, p = .148, \eta_p^2 = .07$; for interaction, $F(5,150) =$

0.63, $p = .675$, $\eta_p^2 = .02$]. Thus, in both background conditions, judgment standard deviations increase with increasing stimulus duration. Although the main ANOVA hints that brightness differentially affects the variability of judgments for black and white background conditions, individual analysis of these conditions reveals no significant effect of brightness in either.

Finally, we looked at the coefficients of variation (standard deviation divided by mean judgment). The results are shown at the bottom of Figure 4. Only stimulus duration had a significant effect, $F(1.8, 110.7) = 15.09$, $p < .001$, $\eta_p^2 = .20$ (all other $ps > .115$).

Discussion

The results of Experiment 3 support the conclusion from Experiments 1 and 2 that it is stimulus differences, rather than absolute intensities, which influence subjective duration. More importantly, they suggest that this effect stems from changes in the rate of the pacemaker: Against a low intensity background, the pacemaker runs faster for bright stimuli; when the background has high intensity, the pacemaker runs faster for dim stimuli.

In the supplementary materials for this paper, we report two additional verbal estimation experiments (conducted prior to Experiment 3) which also examine the influence of intensity on the internal clock. In the first of these, bright and dim lights were presented against a dark background; in the second, quiet and loud tones were presented in silence. The supplementary materials also report a re-analysis of data from Ekman et al. (1966) on the effects of intensity on the perceived duration of electrical stimulation. The results of these studies suggest that, across three sensory modalities, stimulus intensity influences the rate of the pacemaker: In every case, the effects of stimulus intensity become more pronounced as duration increased. However, the effects of intensity upon the intercept and coefficients of variation was more complex, and in some cases the results differed from Experiment 3; more work is required to clarify the effects of intensity on the switch component of the internal clock.

General Discussion

The current research has produced two core findings. Firstly, our results challenge the oft-repeated claim that stimulus magnitude is positively related to perceived duration (e.g., Allan, 1979; Fraisse, 1984; Goldstone et al., 1978; Xuan et al., 2007). It is the difference

between the stimulus and background, not its absolute magnitude, which is important. Secondly, these effects become more pronounced for longer durations, consistent with a change in the rate of the pacemaker component of an internal clock.

Our finding that the difference between stimulus and background influences the perception of duration adds to a growing body of work demonstrating the importance of stimulus differences to subjective time (Gomez and Robertson, 1979; Ono & Kawahara, 2007; Tse et al., 2004; Ulrich et al., 2006; Xuan et al., 2007). Perhaps most relevant to the current paper is a study by Hirsh, Bilger, and Deatherage (1956). Hirsh et al. found that the background loudness could influence the reproduced duration of both tones and lights, but only when the context changed between presentation and reproduction stages. In contrast, the level of background illumination made no difference to the reproduced durations of either auditory or visual stimuli. The data of Hirsh et al. are somewhat noisy, and they used much longer durations than the current experiments (up to 16 s), so that counting strategies were possible. The current data provide a clear demonstration that the effect of intensity on the apparent duration of both visual and auditory stimuli is influenced by context; the results of Hirsh et al. (1956) provide an intriguing hint that these context effects may be at least partially cross-modal.

The fact that stimulus magnitude influences the perception of duration has been taken as support for the idea that there is a common system for representing temporal and non-temporal magnitudes (e.g., Conson, Cinque, Barbarulo, & Trojano, 2008; Walsh, 2003; Xuan et al., 2007). Just as size, number and luminance interfere with judgments of duration, so these dimensions interfere with one another (e.g., Dormal & Pesenti, 2007; Hurewitz, Gelman, & Schnitzer, 2006; Pinel, Piazza, Le Bihan, & Dehaene, 2004) and there seems to be a spatial component to the representation of both quantity (Dehaene, Bossini, & Giraux, 1993) and time (Vallesi, Binns, & Shallice, 2008). Our results suggest that, if there is a common system for temporal and non-temporal magnitudes, it cannot be based upon an intrinsic association of high magnitudes on one dimension with high magnitudes on another; there must be some flexibility in the representation system and sensitivity to the background against which stimuli occur.

In Experiment 3, the difference between stimulus and background influenced the slope of the line relating subjective duration and stimulus duration but had no effect on intercept. Within the framework of scalar expectancy theory, these results suggest that the more different a stimulus is from its background context, the higher the pulse rate of the internal pacemaker. Thus, our data support the assertion that stimulus “energy level” is an

important determinant of the speed of the pacemaker (Allan, 1998, p. 111), although they demonstrate that it is relative rather than absolute energy levels that matter. Our results do not preclude the possibility of an effect of intensity on switch latencies; they only show that intensity effects are not solely due to a change in the difference between switch opening and closing times.

The form of our results matches that from other investigations of slope and intercept effects in human timing. However, the effects are weaker than those reported in some other studies. In the black background condition of Experiment 3, the slope for bright stimuli was 1.13 times that for dim stimuli; in the white background condition, dim:bright slope ratio was 1.12. In contrast, Wearden et al. (2007, Experiment 2A) found a much bigger effect (a slope ratio of 1.55) when comparing filled and empty intervals. Our intensity manipulations are closer to the effects obtained when comparing tones and lights (e.g., Wearden et al., 1998, Experiment 2, found that tones produce a slope 1.23 times that for lights) or the increase in pacemaker rate that occurs when stimuli are preceded by a series of clicks (e.g., Wearden et al., 1998, Experiment 3, found that auditory stimuli preceded by clicks had a slope that was 1.17 times that for stimuli without clicks). Although the effects of intensity on the pacemaker are relatively modest, data from Ekman et al. (1966) demonstrate that more extreme effects result from more extreme intensity differences.

We have primarily related our results to SET, as this provides a well-developed framework for human timing behaviour (Allan, 1992; 1998; Wearden, 2003). There are, of course, other interpretations for our results. As noted above, our data do not rule out a role for effects of onset/offset detection, and low-level effects such as retinal persistence (Long & Beaton, 1980; 1981) may be partially responsible for our results. While such effects may play some role, Fraisse (1984) has pointed out that they are limited to very brief stimuli (where Bloch's law holds – that is, where there is temporal summation), and are not applicable to auditory stimuli; the current experiments used durations of several hundred milliseconds and employed both visual and auditory modalities. Beyond the possibility of low-level effects, more general frameworks for interpreting human timing have emphasized the importance of attention (e.g., Brown, 1997; see Lejeune, 1998, for a discussion). For example, Zakay and Block (1998) have proposed that the flow of pulses into the accumulator depends on the amount of attention paid to the stimulus. We can readily imagine that stimuli which are more different from the background context demand more attention and, thus, a greater flow of counts per unit time. Similarly, Penney, Gibbon, and Meck (2000) have suggested that the switch between pacemaker and accumulator only remains open whilst attention is paid to a

stimulus; stimuli which are closer to the background may cause more “flickering” of the switch and, ergo, less accumulation of counts. Irrespective of the theoretical framework one adopts, our results provide important empirical constraints on any model of time perception.

References

- Allan, L. G. (1979). The perception of time. *Perception & Psychophysics*, *26*, 340-354.
- Allan, L. G. (1992). The internal clock revisited. In F. Macar, V. Pouthas, & W. Freidman (Eds.), *Time, action and cognition: towards bridging the gap* (pp. 191-202). London: Kluwer Academic.
- Allan, L. G. (1998). The influence of the scalar timing model on human timing research. *Behavioural Processes*, *44*, 101-117.
- Belin, P., McAdams, S., Thivard, L., Smith, B., Savel, S., Zilbovicius, M., Samson, S., & Samson, Y. (2002). The neuroanatomical substrate of sound duration discrimination. *Neuropsychologia*, *40*, 1956-1964.
- Berglund, B., Berglund, U., Ekman, G., & Frankenhaeuser, M. (1969). The influence of auditory stimulus intensity on apparent duration. *Scandinavian Journal of Psychology*, *10*, 21-26.
- Brown, S. W. (1995). Time, change, and motion: The effects of stimulus movement on temporal perception. *Perception & Psychophysics*, *57*, 105-116.
- Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal processing. *Perception & Psychophysics*, *59*, 1118-1140.
- Conson, M., Cinque, F., Barbarulo, A. M., & Trojano, L. (2008). A common processing system for duration, order and spatial information: evidence from a time estimation task. *Experimental Brain Research*, *187*, 267-274.
- Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, *34*, 582-593.
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and

- number magnitude. *Journal of Experimental Psychology: General*, 122, 371-396.
- Dormal, V., & Pesenti, M. (2007). Numerosity-length interference: A stroop experiment. *Experimental Psychology*, 54, 289-297.
- Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity-duration interference: A stroop experiment. *Acta Psychologica*, 121, 109-124.
- Eagleman, D. M., & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B*, 364, 1841-1851.
- Ekman, G., Frankenhaeuser, M., Berglund, B., & Waszak, M. (1969). Apparent duration as a function of intensity of vibrotactile stimulation. *Perceptual and Motor Skills*, 28, 151-156.
- Ekman, G., Frankenhaeuser, M., Levander, S., & Mellis, I. (1966). The influence of intensity and duration of electrical stimulation on subjective variables. *Scandinavian Journal of Psychology*, 7, 58-64.
- Forster, K. I., and Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116-124.
- Fraisse, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of Perception, Vol. VIII: Perceptual Coding* (pp 203-247. London: Academic Press.
- Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, 35, 1-36.
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Ed.s), *Annals of the New York Academy of Sciences, Vol. 423: Timing and time perception* (pp. 52-77). New York: New York Academy of Sciences.

- Goldstone, S., & Goldfarb, J. L. (1964). Auditory and visual time judgment. *The Journal of General Psychology*, *70*, 369-387.
- Goldstone, S., Lhamon, W. T., & Sechzer, J. (1978). Light intensity and judged duration. *Bulletin of the Psychonomic Society*, *12*, 83-84.
- Gomez, L. M., & Robertson, L. C. (1979). The filled-duration illusion: The function of temporal and nontemporal set. *Perception & Psychophysics*, *25*, 432-438.
- Grondin, S. (1998). Judgments of duration of visually marked empty time intervals: Linking perceived duration and sensitivity. *Perception & Psychophysics*, *60*, 319-330.
- Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin*, *127*, 22-44.
- Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processing in comparing. *Psychological Bulletin*, *97*, 35-61.
- Hellström, Å. (2003). Comparison is not just subtraction: Effects of time- and space-order on subjective stimulus difference. *Perception & Psychophysics*, *65*, 1161-1177.
- Hellström, Å., & Rammsayer, T. (2004). Effects of time-order, interstimulus interval, and feedback in duration discrimination of noise bursts in the 50- and 1000-ms ranges. *Acta Psychologica*, *116*, 1-20.
- Hirsh, I. J., Bilger, R. C., & Deatherage, B. H. (1956). The effect of auditory and visual background on apparent duration. *American Journal of Psychology*, *69*, 561-574.
- Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*, 19599-19604.
- Laming, D. R. J. (1997). *The measurement of sensation*. London: Oxford University Press.

- Lejeune, H. (1998). Switching or gating? The attentional challenge in cognitive models of psychological time. *Behavioural Processes*, *44*, 127-145.
- Lockhead, G. R. (1992). Psychophysical scaling: judgments of attributes or objects? *Behavioral and Brain Sciences*, *15*, 543-601.
- Lockhead, G. R., & Hinson, J. (1986). Range and sequence effects in judgment. *Perception & Psychophysics*, *40*, 53-61.
- Long, G. M., & Beaton, R. J. (1980). The effects of spatial frequency and target type on perceived duration. *Perception & Psychophysics*, *28*, 413-421.
- Long, G. M., & Beaton, R. J. (1981). The effects of stimulus numerosity, retinal location, and rod contrast on perceived duration of brief visual stimuli. *Perception & Psychophysics*, *29*, 389-394.
- Masson, M. E. J., & Loftus, G. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203-220.
- Mattes, S., & Ulrich, R. (1998). Directed attention prolongs the perceived duration of a brief stimulus. *Perception & Psychophysics*, *60*, 1305-1317.
- Matthews, W. J., & Stewart, N. (2008). The effect of stimulus range on two-interval frequency discrimination. *Journal of the Acoustical Society of America*, *123*, EL45-EL51.
- Matthews, W. J., & Stewart, N. (2009a). Psychophysics and the judgment of price: Judging complex objects on a non-physical dimension elicits sequential effects like those in perceptual tasks. *Judgment and Decision Making*, *4*, 64-81.
- Matthews, W. J., & Stewart, N. (2009b). The effect of interstimulus interval on sequential effects in absolute identification. *Quarterly Journal of Experimental Psychology*, *62*, 2014-2029.

- Noguchi, Y., & Kakigi, R. (2006). Time representations can be made from nontemporal information in the brain: An MEG study. *Cerebral Cortex*, *16*, 1797-1808.
- Ono, F., & Kawahara, J.-I. (2007). The subjective size of visual stimuli affects the perceived duration of their presentation. *Perception & Psychophysics*, *69*, 952-957.
- Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, *72*, 407-418.
- Penney, T. P., Gibbon, J., & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and memory processes. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1770-1787.
- Penton-Voak, I. S., Edwards, H., Percival, A., and Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*, 307-320.
- Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*, 983-993.
- Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics*, *50*, 563-574.
- Rammsayer, T., & Ulrich, R. (2001). Counting models of temporal discrimination. *Psychonomic Bulletin & Review*, *8*, 270-277.
- Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, *112*, 881-911.
- Stewart, N., & Matthews, W. J. (2009). Relative judgment and knowledge of the category structure. *Psychonomic Bulletin & Review*, *16*, 594-599.

- Thomas, E. A. C., & Cantor, N. (1975). On the duality of simultaneous time and size perception. *Perception & Psychophysics*, *18*, 44-48.
- Thomas, E. A. C., & Cantor, N. E. (1976). Simultaneous time and size perception. *Perception & Psychophysics*, *19*, 353-360.
- Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the 'internal clock'. *Psychological Monographs*, *77*, whole no. 576, 1-31.
- Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics*, *66*, 1171-1189.
- Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychological Research*, *70*, 77-87.
- Vallesi, A., Binns, M., & Shallice, T. (2008). An effect of spatial-temporal association of response codes: Understanding the cognitive representations of time. *Cognition*, *107*, 501-527.
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483-488.
- Ward, L. M. (1979). Stimulus information and sequential dependencies in magnitude estimation and cross-modality matching. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 444-459.
- Wearden, J. H. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and mind II: Information processing perspectives* (pp. 21-29). Göttingen, Germany: Hogrefe & Huber.
- Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged

longer than lights”: Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology*, 51B, 97-120.

Wearden, J. H., Norton, R., Martin, S., & Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 716-729.

Wearden, J. H., Todd, N. P. M., & Jones, L. A. (2006). When do auditory/visual differences in duration judgments occur? *Quarterly Journal of Experimental Psychology*, 59, 1709-1724.

Xuan, B., Zhang, D., He, S., and Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision*, 7, 1-5.

Zakay, D., & Block, R. A. (1998). New perspectives on prospective time estimation. In V. DeKeyser, G. D’Ydewalle, & A. Vandierendonck (Eds.), *Time and dynamic control of behaviour* (pp. 129-141). Göttingen, Germany: Hogrefe and Huber.

Zelkind, I. (1973). Factors in time estimation and a case for the internal clock. *The Journal of General Psychology*, 88, 295-301.

Figure Legends

Figure 1. Trial structure in Experiment 1. The figure illustrates the black background condition. In this example, the standard stimulus was bright grey and the comparison stimulus was dim; on half of the trials, these luminances were reversed. The trial structure was identical in the white background condition except that the background (including the blank screens) was white, and the fixation cross was black. (Note: Figure not drawn to scale.)

Figure 2. Results of Experiment 1. The plots show the proportion of trials on which the comparison stimulus was judged longer than the standard stimulus. The solid lines show the results when the comparison stimulus was a dim grey square and the standard was a bright grey square; the dashed lines show the results when the comparison stimulus was a bright square and the standard stimulus was a dim square. Error bars show plus or minus one standard error of the mean. Note that, since ours was a within-subject design, these error bars provide no indication of the significance of differences between means (Masson & Loftus, 2003; violations of sphericity meant it was not possible to use a pooled within-subject error term to calculate standard errors). This is true of the error bars shown on all plots in this paper.

Figure 3. Results of Experiment 2. The plots show the proportion of trials on which the comparison stimulus was judged to last longer than the 600-ms standard. The solid line shows the results when the comparison stimulus was loud and the standard stimulus was quiet; the dashed line shows the results when the comparison stimulus was quiet and the standard stimulus was loud.

Figure 4. Results of Experiment 3.

Figure 1.

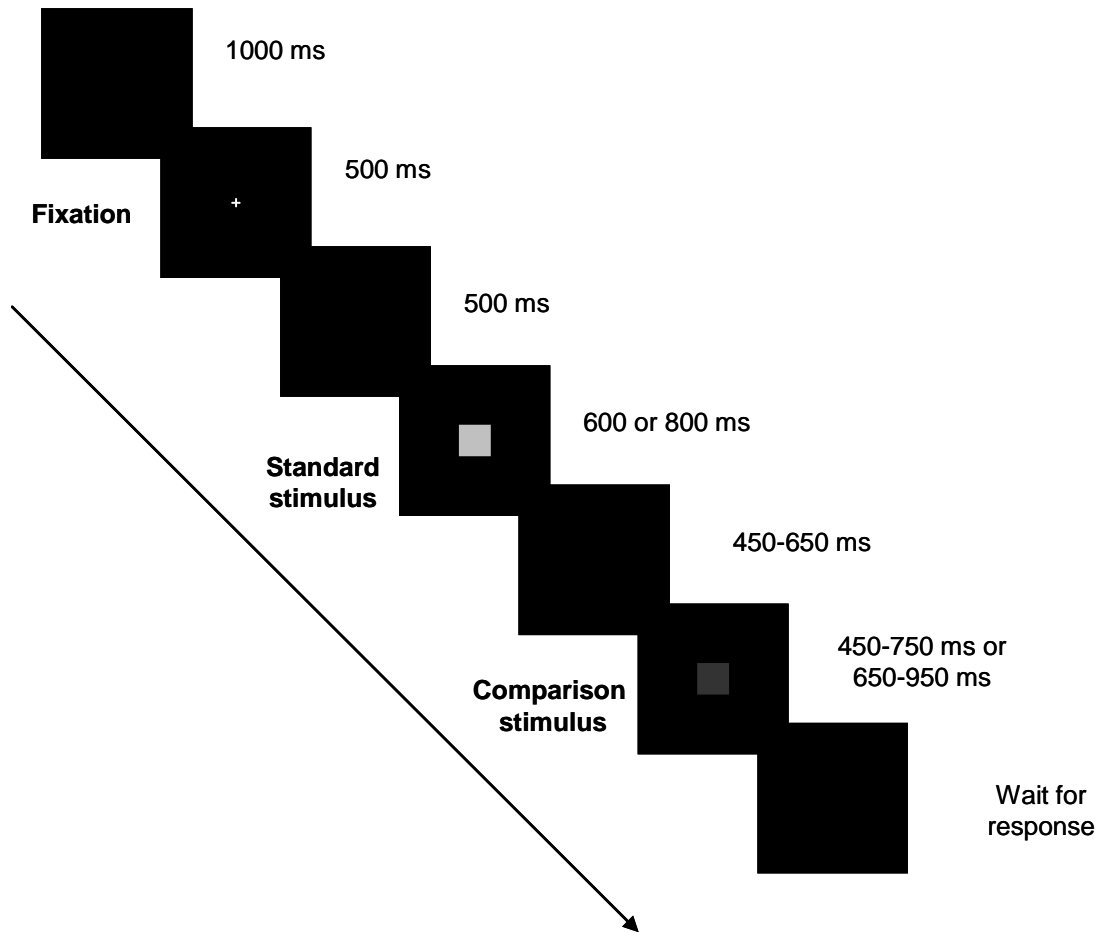


Figure 2.

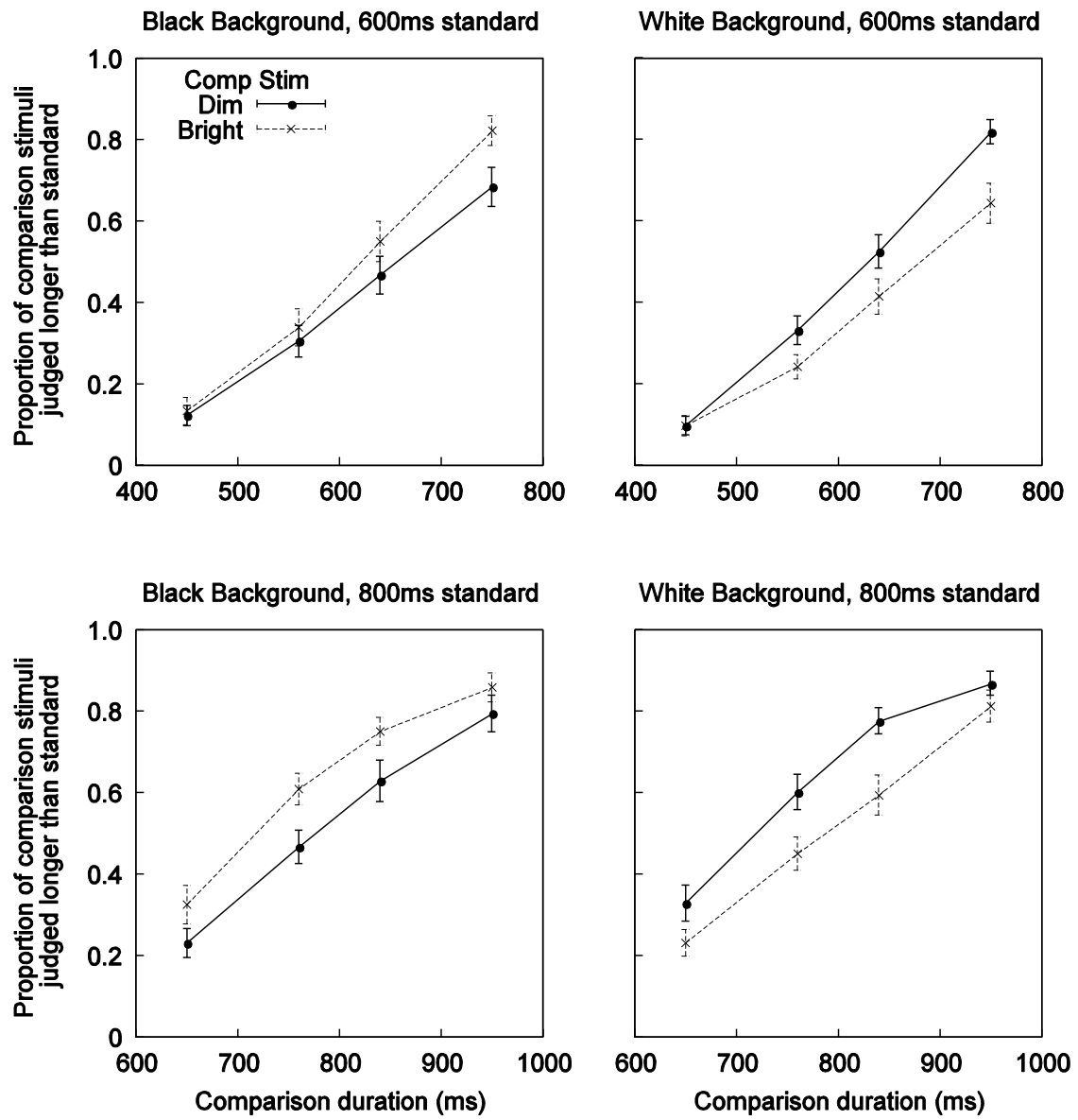


Figure 3.

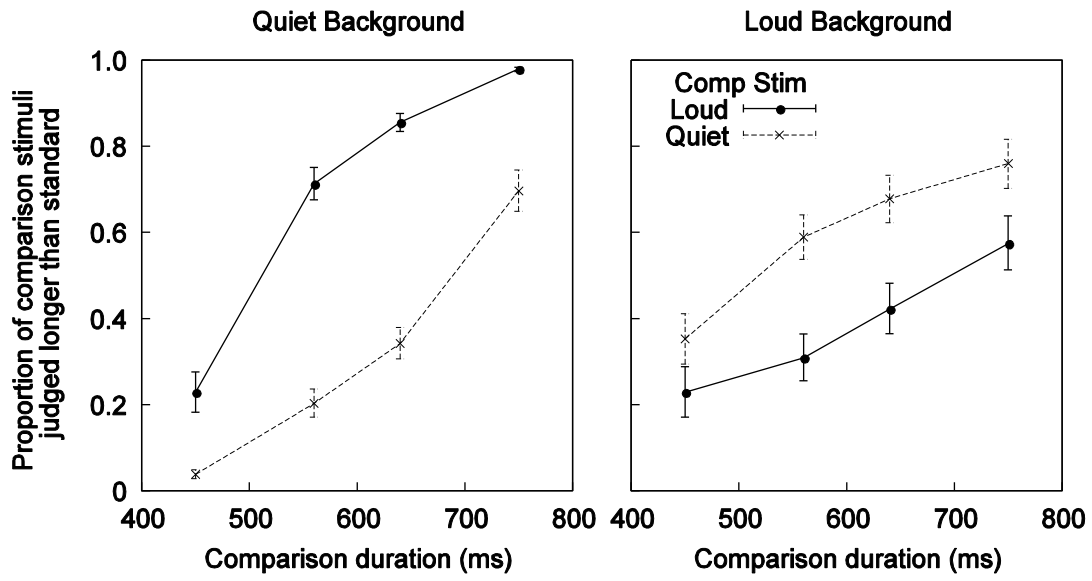


Figure 4.

