*Psicológica (2018), 39,* 142-163
doi: 10.2478/psicolj-2018-0007

# Stimulus response compatibility affects duration judgments, not the rate of subjective time

Alexander Varakin*[1], Amanda Renfro[2], & Jason Hays[2]

[1] *Eastern Kentucky University, USA*

[2] *Florida International University, USA*

The current experiments examined whether non-temporal associations can affect duration judgments without affecting the rate of subjective time. In both experiments, participants performed a temporal bisection task, judging on each trial whether stimulus' duration was closer to pre-learned short or long standards. In each experiment, the spatial compatibility between stimuli and responses was manipulated. In both experiments, stimulus-response compatibility (SRC) affected duration judgments: stimuli that were spatially compatible with the key used for long judgments elicited long responses at shorter objective durations than stimuli that were compatible with the key used for short judgments. The size of SRC's effect did not depend on the magnitude of the standard durations and SRC's effect was magnified even when SRC was introduced after the relevant temporal interval had ended. Thus, these findings are consistent with the idea that duration judgments can be affected without influencing the rate at which subjective time passes.

Casual experience and scientific research both suggest that time perception is malleable. In laboratory experiments, it is common for observers to judge whether a stimulus is short or long relative to pre-defined standard durations. Experiments have demonstrated, for example, that stimulus characteristics such as size (e.g. Ono & Kawahara, 2007), brightness (e.g. Xuan, Zhang, He, & Chen, 2007), and coherence (e.g. Varakin, Klemes, & Porter, 2014) can affect duration judgments. Similarly, manipulating an observer's attentional (Block, Hancock, & Zakay, 2010) or emotional (Droit-Volet & Meck, 2007) state can influence judged duration. Researchers are often content to simply demonstrate that a given manipulation affects responses in a duration judgment task, while other times, researchers make

***Corresponding author:** D. Alexander Varakin. Department of Psychology. 127 Cammack Building, 521 Lancaster Avenue, Richmond, KY, USA, 40475. Tlfn: 1(859) 622-2511. E-mail: donald.varakin@eku.edu

claims about a manipulation affecting the subjective passage of time (see e.g. Hagura, Kanai, Orgs, & Haggard, 2012). It is important to note, however, that finding a manipulation that affects *duration judgments* does not necessarily imply that the *subjective passage of time* changed because duration judgments might be affected by factors other than the rate at which time seems to pass (Ivry & Schlerf, 2008). The purpose of the present experiments is to investigate the idea that duration judgments can be affected even when the subjective rate of the passage of time is not. Theoretical models of time perception allow for such effects (Allman et al., 2014; Ivry & Schlerf, 2008), though very little empirical work has actually tested factors that affect duration judgments without affecting subjective passage of time (Matthews, 2011).

*Pacemaker-accumulator* models of time perception can be useful for testing whether an experimental manipulation affects the rate of subjective time. These models feature prominently in debates about the nature of the mechanisms responsible for the malleability of duration judgments, often being described as *dedicated* models, since the putative pacemaker is presumably dedicated to marking the passage of time (e.g. Ivry & Schlerf, 2008; Spencer, Karmarkar, & Ivry, 2009). However, it is also possible that the pacemaker is not dedicated to marking the passage of time, but is *useful* for doing so. Therefore, in adopting the logic underlying pacemaker-accumulator models, we are not attempting to weigh in on the intrinsic vs. dedicated debate. The logic underlying the pacemaker-model is instead useful as it provides a way to differentiate between effects caused by changes in the subjective rate of time's passage and effects caused by other processes.

Pacemaker-accumulator models of time perception describe a set of information processing stages that together track the passage of time and guide decisions concerning it (Allman et al., 2014). The basic idea is that at the onset of an interval, a pacemaker emits pulses at a given rate, and an accumulator captures the pulses. Pacemaker models also include a gate and a switch that modulate the flow of pulses from the pacemaker to the accumulator (e.g. Zakay & Block, 1995, 1997). The "width" of the gate determines how easily pulses reach the accumulator, while the switch stops the flow at the end of the interval. Thus, time-related decisions and behaviors are determined based on the number of pulses accumulated since the onset of the interval. Other processes believed to contribute to time-related judgments have been identified, however, the above description of the model illustrates that not all variation in time-related judgments can be attributed to pacemaker speed alone.

Importantly, pacemaker-accumulator models permit pacemaker effects to be teased out from non-pacemaker effects (Droit-Volet & Wearden, 2002; Matthews, 2011; Wearden et al., 1998). The subjective rate of time passing is effectively set by the pacemaker's rate, so pacemaker-accumulator models can be used to examine whether the subjective rate of time has been altered. If a manipulation affects the rate at which time seems to pass, then its effect should get larger as objective duration increases. To illustrate this point, if the standard rate of a pacemaker is 1 pulse/ms, and some manipulation doubles that rate, then after 500ms, the difference between the standard and manipulated subjective durations would be 500ms. Given a longer duration such as 1000ms, the difference would grow to 1000ms. On the other hand, manipulations that affect other processes, such as how quickly the switch opens or closes, would yield effects that stay constant across a range of objective durations. For example, assume again a 1 pulse/ms pacemaker rate, but a manipulation that adds 5ms to the switch at the end of an interval. After 500ms, the standard subjective duration would be 500ms, and the manipulated subjective duration would be 505ms. After 1000ms, the standard subjective duration would be 1000ms, and the manipulated 1005ms. For both objective durations, the difference between the non-manipulated and manipulated subjective durations is 5ms. In general, the size of pacemaker effects should depend on objective duration, whereas the size of non-pacemaker effects should not.

In duration judgment tasks, participants are often instructed to consider only the duration of some stimulus, and all other features of the stimulus are considered task-irrelevant (see e.g. Gil, Rousset, & Droit-Volet, 2009; Experiment 2 in Ono & Kawahara, 2007; Xuan, et al. 2007). Nevertheless, task-irrelevant features sometimes affect duration judgments. Ivry and Schlerf (2008) raised the possibility that some of these effects might be caused by non-temporal factors, such as the biasing effects of a stimulus's association with the concepts *short* and *long*. Ivry and Schlerf's proposal is interesting because it suggests that variations in performance on tasks designed to measure tracking of duration might not actually be caused by variations in tracking duration. An example would be how size's influence on duration judgments (e.g. Ono & Kawahara, 2007; Xuan, et al. 2007) might be due to the association of *large* with *long,* and *small* with *short* biasing response selection.

While the idea that non-temporal processes might bias duration judgments is theoretically interesting, stimulus attributes such as length, size, and quantity (all of which affect performance in time perception studies; e.g. Casasanto & Boroditsky, 2008; Ono & Kawahara, 2007; Xuan, et al., 2007) might be part of a general magnitude system that is recruited whenever quantities must be processed (Walsh, 2003, but see Young & Cordes, 2013). A portion of the general magnitude system might produce pacemaker-like effects, even in situations where duration is not being judged (Droit-Volet, 2010) or where attributing an effect to changes in the pacemaker would be illogical because the manipulation occurred at the end of an interval (Matthews, 2011). Thus, factors that are related to the general magnitude system may affect duration judgments because time and other quantities are processed similarly. It is therefore interesting to test whether stimulus attributes that would not be processed as quantities can affect duration judgments via associations with *short* and *long*.

One factor that seems unlikely to be processed as a quantity, but might affect duration judgments is *stimulus response compatibility* (SRC). The term SRC applies to tasks that become easier or harder because of the manner in which responses are mapped onto stimuli (Kornblum, Hasbroucp, & Osman, 1990). When stimuli are presented such that required responses are already strongly associated with stimulus features, responding becomes faster and more accurate. However, when the required responses are not associated with stimulus features, or are perhaps associated with a competing response option, responses become slower and less accurate.

The features of stimuli that induce SRC do not have to be relevant to a participant's task. For example, the Simon effect is a well-known phenomenon involving SRC in which an irrelevant stimulus dimension (spatial location) can influence responding (Craft & Simon, 1970; Lu & Proctor, 1995). In a model Simon effect task, participants respond to stimuli that are presented at different spatial locations, but the response rule is based on a non-spatial aspect of the stimuli, such as color (Craft & Simon, 1970). One color might require a left-hand response, and another a right-hand response. Even though the spatial location of the stimulus is not relevant to the task, responses are faster when a stimulus requiring a right-hand response is presented on the right compared to the left, and vice versa for stimuli requiring a left-hand response.

Simon-like effects have been observed in timing tasks (Lalanne, Van Assche, & Giersch, 2012a; Lalanne, Van Assche, Wang, & Giersch, 2012b). In Lalanne et al. (2012a, b), participants performed simultaneity judgments for stimuli that were presented on the left and right side of the screen, with asynchronies ranging from 0ms to 92ms. When the asynchronies were large and easy to notice, a Simon-like effect was observed: participants' responses were biased by the position of the second stimulus. However, when asynchronies were short and harder to notice, patients with schizophrenia showed different patterns of response bias than healthy controls: patients were biased toward the position of the first stimulus, whereas controls were not. These findings are consistent with the idea that altered temporal processing might account for the disturbances of experience reported by patients diagnosed with schizophrenia (e.g. Martin, et al., 2014). For current purposes though, the Lalanne et al., (2012a, b) suggest that Simon-like tasks might be a fruitful way to explore processing of durations.

The current experiments utilize spatial SRC much like the Simon Task (Craft & Simon, 1970). Participants use left and right hand responses, and stimuli appear on either the left or right side of a computer screen. The relevant stimulus dimension is duration: participants performed a temporal bisection task, judging on each trial whether the stimulus' duration is closer to a pre-learned short or long standard duration. Experiment 1 tested whether SRC has effects on temporal bisection performance. If so, then screen side of the stimulus should increase the probability that participants respond using the key that matches the screen side. For example, the screen side corresponding to the key mapped to "long judgment" would reliably elicit long responses at a shorter duration than the screen side corresponding to "short".

Although designed with SRC in mind, the current experiment also allows testing of the *mental time-line* (MTL) hypothesis (Bonato, Zorzi, & Umiltà, 2012). This hypothesis is derived from the fact that in many cultures, time is often represented using an arrow pointing from left to right, such that *short* is associated with *left* and *long* is associated with *right*. Indeed, Vicario, Pecoraro, Turriziani, et al., (2008) found that spatial position of a stimulus affected duration judgments: stimuli on the left were judged as shorter than stimuli on the right (see also Vicario, Rappo, Pepi, & Oliveri, 2009). The current experiments utilize stimuli that appear either on the left or the right side of a computer monitor, thus allowing us to test the MTL hypothesis in addition to testing for an effect of SRC.

## EXPERIMENT 1

The purpose of the first experiment was to test for an effect of spatial SRC on duration judgments using a Simon-like version of the temporal bisection task. As discussed in the introduction, there is reason to predict that SRC should affect duration judgments. However, SRC has been manipulated in experiments using different methods to measure the processing of durations, and the results suggest that participants might be able to discount the effect of SRC. Rakitin (2005) investigated the effects of spatial SRC on choice time production. In the experiments, on each trial the participants were tasked with reproducing the duration of an imperative stimulus via a button press. The position of the stimulus signaled which of 4 possible buttons was to be used for the reproduction. In a compatible group, the spatial location of the stimuli and the buttons aligned, and in an incompatible group the spatial locations did not align. Choice time production was affected, but not in a straightforward manner. The coefficient of variation (*CV*, which is the standard deviation divided by mean) was larger in the incompatible than compatible conditions, especially at the shortest objective durations. This effect suggests a source of variability (i.e. SRC's influence on response selection) other than the pacemaker was contributing to performance. However, SRC did not affect the mean choice time productions, which suggests that participants were able to effectively discount the effect of SRC. Rakitin described this temporal discounting mechanism as error correction. If temporal discounting occurs in a temporal bisection task in which participants simply judge durations rather than reproduce them, then SRC might not affect the duration at which long responses are elicited.

## METHOD

**Participants**. A total of 53 undergraduate students from Eastern Kentucky University provided informed consent and participated in the study to earn course credit. One participant was dropped because of poor performance (<50% short-responses for longest durations), leaving a final n = 52 [29 females, 23 males; $M_{age}$ = 20.62 (*SD* = 3.46)]. All participants reported normal or corrected vision.

**Materials**. Visual stimuli were presented to participants on the 21.5" wide-screen iMac computer (1680 x 1050 resolution, LED-backlight). SuperLab 4.0 (Cedrus Corporation, San Pedro, CA, USA) was used to manage the presentation of the stimuli and to record participant response data. The input device was a Cedrus RB-740 (Cedrus Corporation, San Pedro, CA, USA) response box, which has 7 horizontally arranged buttons. Only the two outermost buttons were used for this experiment. The imperative stimulus was a black square (50 x 50 pixels; 1.2cm x 1.2 cm) presented on a white background. A black cross (.2 cm x .2 cm) was used as a fixation point.

**Procedure and design.** The temporal bisection task consisted of a training phase and a testing phase. Each training phase trial began with the presentation of the fixation cross for 500ms, followed by a black box 6.4 cm (~6.1 degrees visual angle assuming 60cm viewing distance, which was not controlled) to the left or right side of the fixation. The duration of the black box was 400ms (the short standard duration), or 1600ms (the long standard).

At stimulus offset, a response prompt appeared reminding participants of the response mapping. Participants were randomly assigned to one of two response mapping conditions: right-most key = long judgment, left-most key = short judgment; or right-most key = short judgment, left-most key = long judgment. There were 20 total training trials: 10 short and 10 long standards. Instructions stressed accuracy: if a response was incorrect, participants received feedback in the form of a beep played through headphones.

Trials in the testing phase were identical to the practice phase, except the stimulus was presented on the left or right side of the screen for either 400, 600, 800, 1000, 1200, 1400, or 1600 ms (randomly selected with replacement on each trial). The task was to indicate whether stimulus duration was closer to the short or long standard learned during training. As in training, instructions stressed accuracy and a response prompt reminded participants about the response mapping after stimulus offset on each trial. In the test phases, however, no feedback about the accuracy of response was provided. There were 280 test trials with a 500 ms inter-trial interval.

**Results and Discussion.** Data were analyzed by fitting two logit models (using the `pandas` and `statsmodels.formula.api` libraries in Python 2.7) to each participant's individual data, one model for each screen side (left and right). The predictor for each model was stimulus duration and the outcome variable was short response vs. long response. The models were used to calculate what would become the dependent variables for a 2x2 ANOVA (performed with the ez package in R). The primary dependent variable (DV) was a bisection point (BP, sometimes called a point of subjective equality or PSE), which is the duration at which 50% of responses are predicted to be long-judgments (see Table 1). As subjective duration increases, bisection points decrease, because long responses will be elicited at shorter objective durations. Figure 1 plots the average proportion of long responses as a function of stimulus duration, both from the actual data, and from the average of the predictions of the logit models, and Figure 2 plots mean bisection points for each screen side for each response mapping condition. In addition to examining BPs, which measures the threshold between short and long durations, we also analyzed two additional DVs that measure temporal acuity. These analyses are reported in the appendix.

**Table 1. Descriptive statistics from Experiment 1. Bisection Point (BP) means (M) and standard deviations (SD) reported in milliseconds (ms).**

| Statistic | Response Mapping | Screen Left *M* (*SD*) | Screen Right *M* (*SD*) | *N* |
|---|---|---|---|---|
| BP | Left-Key: Short<br>Right-Key: Long | 924<br>(161) | 863<br>(147) | 26 |
| | Left-Key: Long<br>Right-Key: Short | 866<br>(158) | 883<br>(157) | 26 |

An ANOVA with response mapping as a between-subject variable, and screen side presentation as a within-subject variable was conducted with BP as the DV. Neither main effect was significant: screen side, $F(1, 50) = 3.88$, $p = .054$, $\eta^2_{generalized} = .005$; response mapping, $F(1, 50) = 0.20$, $p > .05$, $\eta^2_{generalized} = .003$. However, the interaction of screen side and response mapping was significant, $F(1, 50) = 11.89$, $p < .05$, $\eta^2_{generalized} = .02$. The source of the interaction can be seen in Figures 1 and 2. When the right-most key was used to indicate long judgment, the psychometric function (and hence the BP) for screen-right was shifted leftward relative to the psychometric function for screen left (Figure 1). In other words, when the right-most key served as the long response, stimuli presented on the right side elicited long judgments at a shorter objective duration compared to stimuli presented on the left of the screen (Figure 2). Likewise, when the left key was used for long judgments, the screen-left function shifted leftward relative to screen-right.

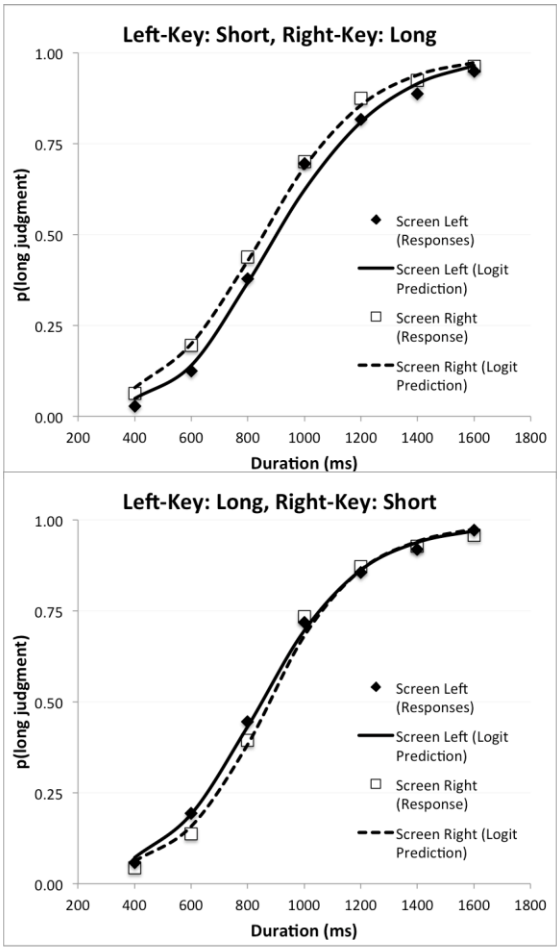


**Figure 1: Mean proportion of long judgments for each duration for Experiment 1. Responses are averaged over participants' actual responses; Logit predictions are averaged over the predictions of the logit models that were fit to individual participants' data.**

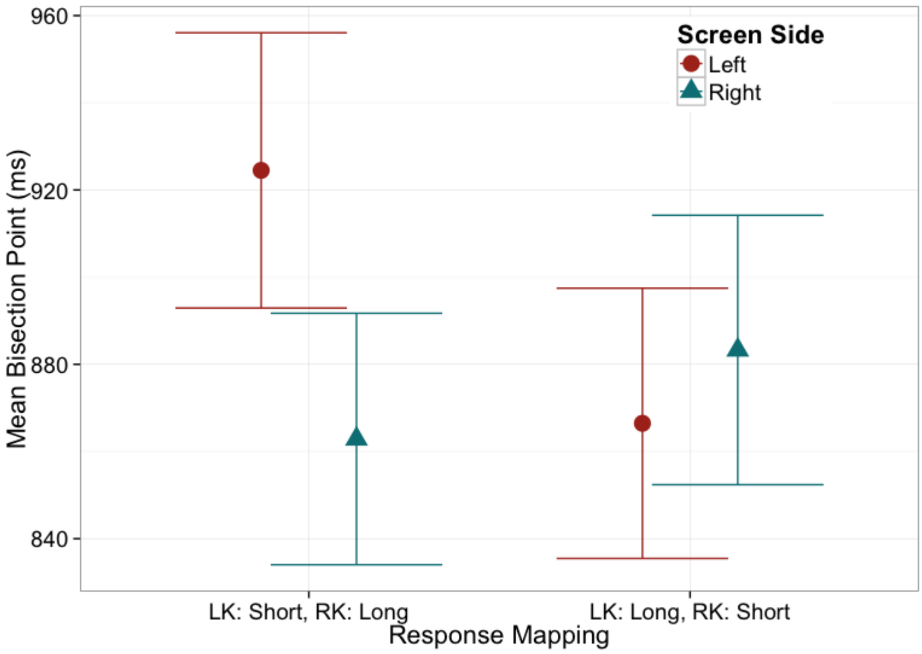


**Figure 2: Mean bisection points from Experiment 1. Error bars represent 1 standard error (not pooled). (LK = Left Key; RK = Right Key).**

The significant screen side by response mapping interaction with BP as the DV supports the hypothesis that SRC affects duration judgments. However, the results do not support the prediction of the MTL hypothesis that screen side should affect temporal-bisection performance: the effect of screen side was not statistically significant.

Before moving onto Experiment 2, it is important to note that the results of Experiment 1 should not be taken as evidence against Rakitin's (2005) temporal discounting model due to the many procedural differences between the choice time production task used in Rakitin's work and the temporal bisection task. The temporal discounting model suggested a reason why SRC might not affect duration judgments, however, the current experiments were not designed as a strong test of the model. Importantly, the results of Experiment 1 do not provide evidence that SRC influences the subjective passage of time, as only one set of standard durations was utilized. As such, it is possible that SRC simply biased responses without affecting the rate of subjective time. Experiment 2 was designed to test whether SRC influences the subjective rate of time.

# EXPERIMENT 2

In order to test whether a manipulation affected the rate at which subjective time passes, it is necessary to examine whether the size of the effect is dependent on duration (e.g. Matthews, 2011; Wearden et al., 1998). Thus, in Experiment 2, two sets of short/long standard durations were used: 200ms/800ms and 400ms/1600ms. The latter effectively contains a replication of Experiment 1.

In addition to manipulating the magnitude of the standard durations, Experiment 2 also included a manipulation of when SRC was introduced. In a *during interval* condition, the imperative stimulus appeared on the left or right side of the screen, as in Experiment 1. In a *post interval* condition, the imperative stimulus always appeared in the center of the screen, and the response-mapping reminder, which had previously been presented in the center during Experiment 1, now appeared on either the left or right side of the screen. Since SRC was not manipulated until after the relevant interval had ended, any effect on duration judgments cannot possibly be due to changes in the rate of subjective time (see Matthews, 2011).

## METHOD

The methods for Experiment 2 were identical to Experiment 1 except as noted below.

**Participants**. A total of 223 undergraduate students from Eastern Kentucky University provided informed consent and participated in the study to earn course credit. Eight participants were dropped because of poor performance (more than 50% short-responses for longest durations or more than 50% long-responses for the shortest durations), leaving a final n = 215 [138 females, 77 males; $M_{age}$ = 20.80 (*SD* = 4.08)]. All participants reported normal or corrected vision.

**Procedure and design.** Standard duration magnitude and SRC timing were both manipulated between subjects. The training phase was the same as in Experiment 1, except that the standard durations in the 200ms/800ms conditions were changed.

In the test phase, the positions of the imperative stimulus and the response prompt for the *during interval* conditions were identical to Experiment 1. In the *post interval* conditions, the imperative stimulus always appeared in the center of the screen, and the response prompt appeared on either the left or right side of the screen, at the same position as the imperative stimulus in the *during interval* conditions. The sample size per condition is given in Table 2.

**Table 2. Bisection Point (BP) descriptive statistics from Experiment 2. Means (M) and standard deviations (SD) reported in milliseconds (ms).**

| SRC Timing | Standard Duration Magnitude | Response Mapping | Screen Left *M* (*SD*) | Screen Right *M*(*SD*) | N |
|---|---|---|---|---|---|
| During Interval | 200ms/800ms | Left-Key: Short<br>Right-Key: Long | 532<br>(98) | 477<br>(99) | 26 |
| | | Left-Key: Long<br>Right-Key: Short | 432<br>(85) | 502<br>(85) | 28 |
| | 400ms/1600ms | Left-Key: Short<br>Right-Key: Long | 896<br>(118) | 838<br>(130) | 23 |
| | | Left-Key: Long<br>Right-Key: Short | 919<br>(162) | 983<br>(216) | 22 |
| Post Interval | 200ms/800ms | Left-Key: Short<br>Right-Key: Long | 535<br>(80) | 454<br>(102) | 23 |
| | | Left-Key: Long<br>Right-Key: Short | 491<br>(81) | 564<br>(84) | 33 |
| | 400ms/1600ms | Left-Key: Short<br>Right-Key: Long | 934<br>(188) | 825<br>(162) | 25 |
| | | Left-Key: Long<br>Right-Key: Short | 854<br>(139) | 990<br>(194) | 35 |

**Results and Discussion.** As in Experiment 1, three DVs were derived from logit models fitted to each participant's individual data for each screen side: bisection point (BP) as a measure of threshold between short and long durations, and two measures of temporal acuity. Descriptive statistics for BP are given in Table 2. Figure 3 plots the average proportion of long responses as a function of stimulus duration, both from the actual data, and from the average of the predictions of the logit models, and Figure 4 plots mean bisection points for each screen side for each condition. The analyses of temporal acuity are in the appendix.

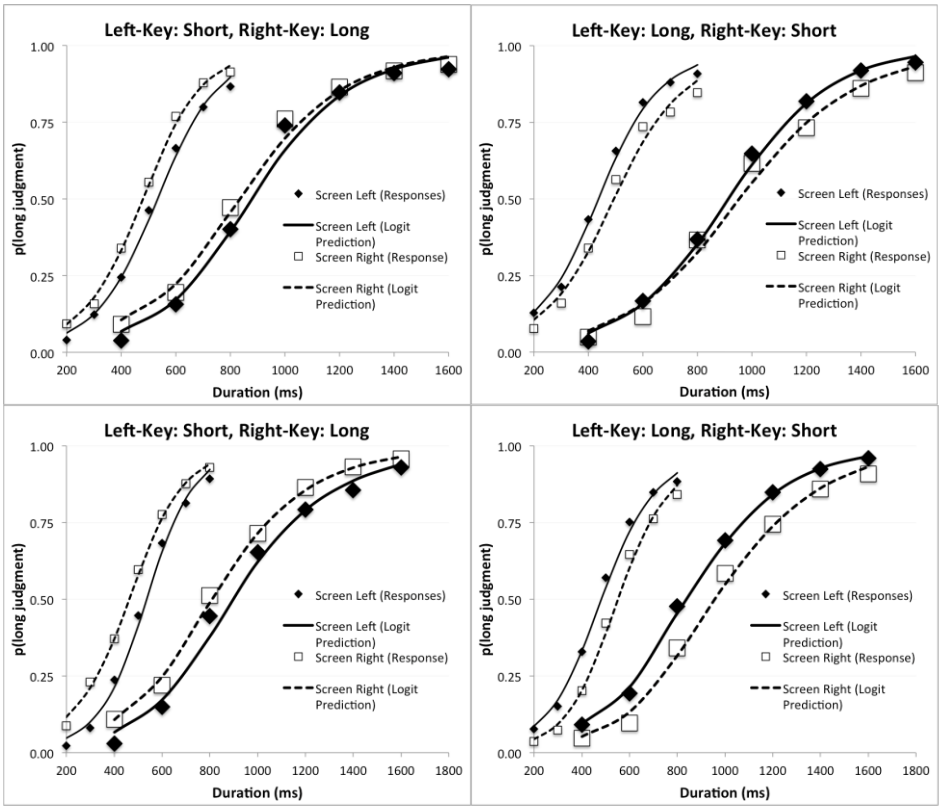


**Figure 3: Mean proportion of long judgments for each duration for Experiment 2. Responses are averaged over participants' actual responses; Logit predictions are averaged over the predictions of the logit models that were fit to individual participants' data. The top row is from the during interval conditions; the bottom row is from the post interval conditions. The smaller markers/thinner lines on the left within each graph represent the 200ms/800ms standard magnitude conditions; the larger markers/thicker lines on the right within each graph represent the 400ms/1600ms standard magnitude conditions.**

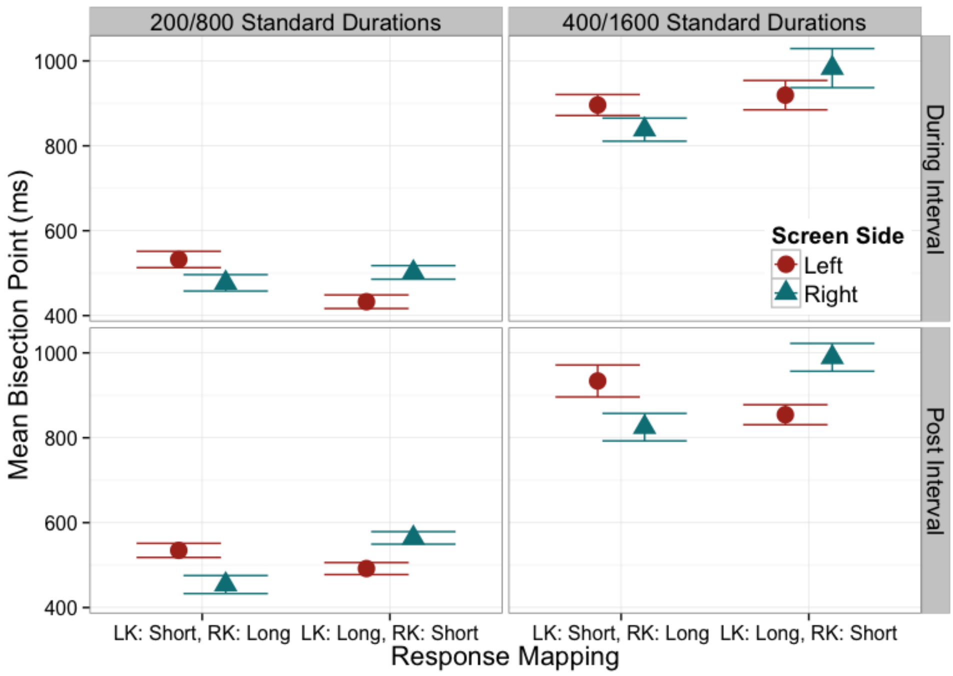


**Figure 4: Mean bisection points from Experiment 2. Error bars represent 1 standard error (not pooled). (LK = Left Key; RK = Right Key).**

An ANOVA with response mapping, standard duration magnitude (200ms/800ms vs. 400ms/1600ms; between-subjects), SRC timing (during vs. post interval; between-subjects), and screen side (left vs. right; within-subjects) as factors was performed with BP as the DV. The only significant main effect was standard duration magnitude ($F(1, 207) = 603.31$, $p < .001$, $\eta^2_{generalized} = .70$). It is to be expected that the BP for the small magnitude standards would be smaller than the BP for the large magnitude standard, and that is indeed what the data reveal. The main effect for screen side was not significant, $F(1, 207) =. 354$, $p = .55$, $\eta^2_{generalized} < .001$, which again fails to support MTL's prediction that screen side should affect duration judgments.

There were several significant interactions. Response mapping interacted with standard duration magnitude, $F(1, 207) = 3.90$, $p = .05$, $\eta^2_{generalized} = .02$. The difference between the BPs in the right/long and left/long response mapping conditions was larger in the 400ms/1600ms (873ms vs. 933ms respectively) than in the 200ms/800ms conditions (499ms vs. 500ms). In Experiment 1 (which also used 400ms and 1600ms as standard durations) the mean right/long BP was 894ms, which was larger than the mean left/long BP of 875ms. Since this interaction was unexpected, is driven by differences of means that go in the opposite direction of the previous experiment, and barely meets the criteria for statistical significance, it should be interpreted with caution.

More important for current purposes are the significant response mapping x screen side ($F(1, 207) = 101.81$, $p < .001$, $\eta^2_{generalized} = .08$) and response mapping x screen side x SRC timing ($F(1, 207) = 5.61$, $p < .05$, , $\eta^2_{generalized} = .01$) interactions. The response mapping x screen side interaction (as seen in Figures 3 and 4) occurred because the BP for the screen side that corresponded to the location of the key used to provide long judgments was smaller than the BP for the screen side that corresponded to the key used to provide short judgments. This result replicates Experiment 1, demonstrating again that SRC influences duration judgments. The three way interaction occurred because the screen side x response mapping interaction was larger in the *post interval* condition than the *during interval* condition. The difference between compatible and incompatible screen side BPs was about 60ms in the *during interval* condition, and 100ms in the *post interval* condition.

Importantly, there was no three way interaction between response mapping, screen-side, and standard duration magnitude, $F(1,207) = 1.93$, $p > .05$. Thus, with BPs as the DV, there is no evidence that the SRC effect (i.e. the interaction between response mapping and screen side) depends on standard duration magnitude.

## GENERAL DISCUSSION

In both experiments, participants performed a temporal bisection task, judging on each trial whether the objective duration of a visual stimulus was closer to pre-learned short or long standard durations. In Experiment 1, stimuli that appeared on the screen side that was spatially compatible with the button used to indicate long judgments reliably elicited long judgments at a shorter duration than stimuli on the screen side that was spatially compatible with the short judgment response. Thus, Experiment 1 demonstrated that SRC influences duration judgments. Experiment 2 further demonstrated that the size of the SRC effect did not depend on the magnitude of the standard durations: the effect was approximately the same size for short/long standards of 200ms/800ms and 400ms/1600ms. Furthermore, Experiment 2 demonstrated that SRC affected duration judgments even more when SRC was introduced *after* the relevant temporal interval had ended than when it occurred during the relevant temporal interval.

This pattern of results suggests that the rate at which subjective time passed was not affected by SRC, even though SRC did affect duration judgments. If SRC influenced the passage of subjective time, then the SRC effect should have been smaller for the shorter standard durations than the longer standard durations. Moreover, an effect that was driven by changes in the rate of subjective time should not have occurred when the manipulation was introduced after the relevant temporal interval. However, the size of SRC effect did not depend on the duration of the standard durations, and the SRC effect was actually bigger when SRC was introduced after the interval had ended. It is important to note that the larger SRC effect in the post interval condition does not, by itself, imply anything about what caused the effect in the during interval condition (or in Experiment 1). However, the post interval SRC effect does demonstrate that changes in temporal bisection performance may not always be due to changes in the rate of subjective time – in the post interval condition, the shift in bisection point could not have been caused by changes in the rate of subjective time. Moreover, the post interval SRC effect is just one piece of evidence: the lack of an effect of standard duration magnitude on the size of the SRC effect also suggests that the rate of subjective time was not changed. Still, it is logically possible that the rate of an internal timer was affected in the during interval condition, although this explanation seems rather implausible given the overall pattern of results.

The current experiments lend little support to the MTL hypothesis. According to the MTL hypothesis, screen side should have affected duration judgments, with screen-left stimuli being judged as shorter than screen-right stimuli, but this effect was not found. However, these results should not be taken as strong evidence against the MTL hypothesis because the current procedure differs from past work demonstrating an effect of screen side on duration judgments (e.g. Vicario et al, 2008; Vicario et al., 2009) in a number of potentially important ways. First, the magnitude of comparison durations was different. In Vicario et al.'s work, participants judged stimuli that ranged from 200 to 400ms in duration. The current experiment used durations ranging from 200 to 800ms in some conditions, and 400 to 1600ms in the rest. It's widely believed that the mechanisms responsible for processing very short durations are different than the mechanisms for longer durations, and some value close to 300ms might be the upper limit of "very short" (e.g. Wearden & Bray, 2001). Second, participants in Vicario et al.'s experiments only used one hand to respond using adjacent buttons, whereas in the current experiments participants used two hands and buttons were several centimeters apart. It is possible that the two-handed response in the current experiment lead to a stronger SRC effect, masking the effect of screen side. However, past work suggests that single-handed tasks still lead to robust SRC effects (e.g. Heister, Ehrenstein, & Schroder-Heister, 1986), which argues against the proposal that the key factor is single-handed vs. two-handed responses. Future work will be required to tease apart the relationship between spatial SRC and the MTL hypotheses.

In conclusion, the current results are consistent with Ivry and Schlerf's (2008) idea that non-temporal associations between the stimuli and short vs. long judgments can significantly affect performance, biasing responses in a way that does not involve changing the rate of an internal clock. However, the current experiments do not support the idea that a mental timeline, from left to right, influences duration judgments.

## REFERENCES


Allman, M.J., Teki, S., Griffiths, T.D., & Meck, W.H. (2014). Properties of the internal clock: First- and second-order principles of subjective time. *Annual Review of Psychology, 65*, 743-771.

Block, R.A., Hancock, P.A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica, 134*, 330-343.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience and Biobehavioral Reviews, 36,* 2257-2273.

Casasanto, D., and Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition, 106*, 579-593.

Craft, J.L. & Simon, J.R. (1970). Processing symbolic information from a visual display: Interference from an irrelevant directional cue. *Journal of Experimental Psychology, 83*(3), 415-420.

Dehaene, S. (2003). The neural basis of the Weber-Fechner Law: a logarithmic mental number line. *Trends on Cognitive Sciences, 7*(4), 145-147.

Droit-Volet, S. (2010). Speeding up a master clock common to time, number and length? *Behavioral Processes, 85,* 126-134.

Droit-Volet, S. & Meck, W.H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513.

Droit-Volet, S., & Wearden, J. (2002). Speeding up the internal clock in children? Effects of visual flicker on subjective duration. *The Quarterly Journal of Experimental Psychology, 55B*(3), 193-211.

Gil, S., Rousset, S. & Droit-Volet, S. (2009). How liked and disliked foods affect time perception. *Emotion, 9*(4), 457-463.

Hagura, N., Kanai, R., Orgs, G., & Haggard, P. (2012). Ready steady slow: action preparation slows the subjective passage of time. *Proceedings of the Royal Society B: Biological Sciences,* doi:10.1098/rspb.2012.1339

Heister, G., Ehrenstein, W.H., & Schroeder-Heister, P. (1986). Spatial S-R compatibility effects with unimanual two-finger choice reactions for prone and supine hand positions. *Perception & Psychophysics, 40*(4), 271-278.

Ivry, R.B. & Schlerf, J.E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences, 12*(7), 273-280.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility - A model and taxonomy. *Psychological Review, 97*(2), 253-270.

Lalanne, L., Van Assche, M. & Giersch, A. (2012a). When predictive mechanisms go wrong: Disordered visual synchrony thresholds in schizophrenia. *Schizophrenia Bulletin, 38*(3), 506-513.

Lalanne, L., Van Assche, M., Wang, W., & Giersch, A. (2012b). Looking forward: An impaired ability in patients with schizophrenia? *Neuropsychologia 50*, 2736-2744.

Lu, C.-H., & Proctor, R.W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review, 2*(2), 174-207.

Matthews, W. (2011). Can we use verbal estimation to dissect the internal clock? Differentiating the effects of pacemaker rate, switch latencies, and judgment processes. *Behavioral Processes, 86*(1),68-74.

Ono, F., & Kawahara, J.-I. (2007). The subjective size of visual stimuli affects perceived duration of their presentation. *Perception & Psychophysics, 69*, 952-957.

Rakitin, B.C. (2005). The effect of spatial stimulus-response compatibility on choice time production accuracy and variability. *Journal of Experimental Psychology: Human Perception and Performance, 31*(4), 685-702.

Spencer, R.M.C, Karmarker, U., & Ivry, R.B. (2009). Evaluating dedicated and intrinsic models of temporal encoding by varying context. *Philosophical Transactions of the Royal Society: B, 364*, 1853-1863.

Varakin, D.A., Klemes, K.J., & Porter, K.A. (2013). The effect of scene structure on time perception. *Quarterly Journal of Psychology, 66*(8), 1639-1652.

Vicario,C.M., Pecoraro,P., Turriziani,P., Koch,G., Caltagirone,G., & Oliveri,M. (2008). Relativistic compression and expansion of experiential time in the left and right space. *PLoS One 3*, e1716.

Vicario, C.M., Rappo, G., Pepi, A.M. & Oliveri, M. (2009). Timing flickers across sensory modalities. *Perception, 38,* 1144-1151.

Wearden, J.H., & Bray, S. (2001). Scalar timing without reference memory? Episodic temporal generalization and bisection in humans. *The Quarterly Journal of Experimental Psychology, 54B*(4), 289-309.

Wearden, J.H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of the model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology, 51B*(2), 97-120.

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

Young, L.N., & Cordes, S. (2013). Fewer things, lasting longer: The effects of emotion on quantity judgments. *Psychological Science 24*(6). 1057-1059.

Zakay, D., & Block, R. A. (1995). An attentional-gate model of prospective time estimation. In M. Richelle, V. De Keyser, G. d'Ydewalle, & A. Vandierendonck (Eds.), *Time and the dynamic control of behavior* (pp. 167-178). Liège , Belgium : Universite de Liege.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science, 6*(1), 12-16.

(Manuscript Received: 17 October 2017; accepted: 5 February 2018)

# APPENDIX

The slope of logit function and the difference limen (dL) were analyzed as measures of temporal acuity. The slope indicates the steepness of the line relating actual duration to log(p(long)/p(short)), with a larger slope indicating greater temporal acuity. The dL is half the difference between the durations predicted to elicit 25% and 75% long-responses (see Droit-Volet and Wearden, 2002). For dL, smaller values indicate greater temporal acuity.

For Experiment 1 both were used as DVs in separate ANOVAs with response mapping as a between-subject variable, and screen side presentation as a within-subject variable. There were no significant effects for either measure of temporal acuity: slope as the DV, response mapping, $F(1, 50) < 1$; screen side, $F(1, 50) < 1$; interaction, $F(1, 50) = 2.29$, $p = .14$; dL as the DV, response mapping, $F(1, 50) < 1$; screen side, $F(1, 50) < 1$; interaction, $F(1, 50) = 1.79$, $p = .19$ (see Table A1).

**Table A1. Descriptive statistics from Experiment 1. Slopes were from the logistic functions, thus the unit is log(p(long)/p(short)). For dL, the unit is milliseconds (ms).**

| Statistic | Response Mapping | Screen Left *M* (*SD*) | Screen Right *M* (*SD*) | *N* |
|---|---|---|---|---|
| Slope | Left-Key: Short<br>Right-Key: Long | 0.0073<br>(0.0035) | 0.0067<br>(0.0027) | 26 |
| | Left-Key: Long<br>Right-Key: Short | 0.0071<br>(0.0029) | 0.0075<br>(0.0030) | 26 |
| dL | Left-Key: Short<br>Right-Key: Long | 178<br>(71) | 188<br>(80) | 26 |
| | Left-Key: Long<br>Right-Key: Short | 184<br>(100) | 168<br>(61) | 26 |

In Experiment 2, both measures of temporal acuity were entered as DVs in ANOVAs with response mapping, standard duration magnitude (200ms/800ms vs. 400ms/1600ms; between-subjects), SRC timing (during vs. post interval; between-subjects), and screen side (left vs. right; within-subjects) as factors (see Tables A2 and A3). For both DVs, the main effects of standard duration magnitude (slope as DV: $F\ (1, 207) = 80.75$, $p < .001$, $\eta^2_{generalized} = .26$; dL as DV: $F(1, 207) = 50.23$, $p < .001$, , $\eta^2_{generalized} = .18$) and screen side (slope as DV: $F\ (1, 207) = 7.09$, $p = .008$, $\eta^2_{generalized} = .004$; dL as DV: $F(1, 207) = 4.98$, $p = .03$, , $\eta^2_{generalized} = .002$) were significant. The main effect of standard duration is expected: the slopes should be steeper for short standards ($M = .0010$) than long standards ($M = .0062$), and the dL should be smaller for short standards (*M* = 128ms) than the dL for long standards (*M* = 212ms).

**Table A2. Slope descriptive statistics from Experiment 2. Slopes were from the logistic functions, thus the unit for the means (M) and standard deviations (SD) is log(p(long)/p(short)).**

| SRC Timing | Standard Duration Magnitude | Response Mapping | Screen Left *M* (*SD*) | Screen Right *M* (*SD*) | N |
|---|---|---|---|---|---|
| During Interval | 200ms/800ms | Left-Key: Short Right-Key: Long | 0.010 (0.0034) | 0.011 (0.0035) | 26 |
| | | Left-Key: Long Right-Key: Short | 0.010 (0.0044) | 0.0085 (0.0034) | 28 |
| | 400ms/1600ms | Left-Key: Short Right-Key: Long | 0.0064 (0.0026) | 0.0061 (0.0030) | 23 |
| | | Left-Key: Long Right-Key: Short | .0067 (.0020) | 0.0059 (.0020) | 22 |
| Post Interval | 200ms/800ms | Left-Key: Short Right-Key: Long | 0.012 (0.0041) | 0.010 (0.0035) | 23 |
| | | Left-Key: Long Right-Key: Short | 0.010 (0.0036) | 0.010 (0.0036) | 33 |
| | 400ms/1600ms | Left-Key: Short Right-Key: Long | 0.0060 (0.0023) | 0.0062 (0.0025) | 25 |
| | | Left-Key: Long Right-Key: Short | 0.0065 (0.0039) | 0.0060 (0.0023) | 35 |

**Table A3. Difference limen (dL) descriptive statistics from Experiment 2. Means (M) and standard deviations (SD) reported in milliseconds (ms).**

| SRC Timing | Standard Duration Magnitude | Response Mapping | Screen Left *M* (*SD*) | Screen Right *M* (*SD*) | N |
|---|---|---|---|---|---|
| During Interval | 200ms/800ms | Left-Key: Short Right-Key: Long | 124 (57) | 117 (41) | 26 |
| | | Left-Key: Long Right-Key: Short | 141 (90) | 152 (70) | 28 |
| | 400ms/1600ms | Left-Key: Short Right-Key: Long | 204 (99) | 223 (116) | 23 |
| | | Left-Key: Long Right-Key: Short | 192 (118) | 241 (233) | 22 |
| Post Interval | 200ms/800ms | Left-Key: Short Right-Key: Long | 107 (41) | 124 (47) | 23 |
| | | Left-Key: Long Right-Key: Short | 131 (49) | 123 (47) | 33 |
| | 400ms/1600ms | Left-Key: Short Right-Key: Long | 210 (80) | 212 (110) | 25 |
| | | Left-Key: Long Right-Key: Short | 210 (92) | 205 (63) | 35 |

The main effects of screen side were unpredicted. They reflect steeper slopes and smaller dLs for stimuli appearing on the left compared to the right (slopes of M = .0083 and *M* = .0080 respectively for left and right; dLs of *M* = 164ms and *M* = 173ms, respectively for left and right). This effect suggests slightly better temporal acuity for stimuli that appear on the left side of the screen compared to the right. This effect might be interpretable in terms of the MTL hypothesis: perhaps the MTL represents shorter durations (on the left) with greater precision than longer durations (on the right), which would be expected if the MTL followed something like Weber's Law (see, e.g. Dehaene, 2003 for a discussion of Weber's Law and the mental number line). However, the effect of screen side was small and qualified by higher order interactions. With slope as the DV, the three-way interaction between response mapping, SRC timing, and screen-side ($F$ (1, 207) = 7.55, $p$ = .006, $\eta^2_{generalized}$ = .004) and the four-way interaction ($F$ (1, 207) = 9.76, $p$ < .002, $\eta^2_{generalized}$ = .006) were both significant. With dL as the DV, three-way interactions between response mapping x SRC timing x screen side ($F$(1, 207) = 5.01, $p$ = .03, , $\eta^2_{generalized}$ = .003) and response mapping x standard duration magnitude x screen side ($F$(1, 207) = 4.70, $p$ = .03, , $\eta^2_{generalized}$ = .003) were significant.

The interaction effects do not have clear implications for the main hypotheses under investigation here. However, it is interesting that the effect of screen-side suggested increased temporal acuity on the left side of the screen. Given the inconsistent pattern of the effects of screen side as reflected by the higher-order interactions, and a lack of an effect in Experiment 1, any account we can offer is highly speculative, and should be considered with extreme caution.