Decisional carryover effects in interval timing: Evidence of a generalized response bias

Jordan J. Wehrman[1] · John Wearden[2] · Paul Sowman[1]

Abstract

Decisional carryover refers to the tendency to report a current stimulus as being similar to a prior stimulus. In this article, we assess decisional carryover in the context of temporal judgments. Participants performed a temporal bisection task wherein a probe between a long and short reference duration (Experiment 1) was presented on every trial. In Experiment 2, every other trial presented a duration the same as the short or long reference duration. In Experiment 3, we concurrently varied both the size and duration of stimuli. Experiment 1 demonstrated the typical decisional carryover effect in which the current response was assimilated towards the prior response. In Experiment 2, this was not the case. Conversely, in Experiment 2, we demonstrated decisional carryover from the prior probe decision to the reference duration trials, a judgment which should have been relatively easy. In Experiment 3, we found carryover in the judgment of both size and duration, and a tendency towards decisional carryover having a larger effect size when participants were making size judgments. Together, our findings indicate that decisional carryover in duration judgments occur given relatively response-certain trials and that this effect appears to be similar in both size and duration judgments. This suggest that decisional carryover is indeed decisional in nature, rather than due to assimilative effects in perception, and that the difficulty of judging the previous test stimuli may play a role in whether assimilation occurs in the following trial when judging duration.

Keywords Adaptation and Aftereffects · Categorization · Decision making · Duration judgment · Carryover · Time perception · Temporal bisection · Sequential processing

Assimilative carryover (or attractive serial dependence) refers to a bias towards judging a current stimulus as similar to a previous one. This effect has been shown in a variety of tasks, from numerosity (Cicchini, Anobile, & Burr, 2014; Fornaciai & Park, 2018) and orientation judgments (Fischer & Whitney, 2014), to the perceived taste of cheese (Muir & Hunter, 1991). These carryover effects are important to how we interact with the world; we understand events and stimuli in both a global context ('an ant is a small insect') and a local context ('that ant is big compared with that flea'). Most investigations of carryover effects involve one of the 'five senses': orientation is visual; cheese tasting is gustatory. However, carryover effects can also be seen in the judgment of duration (Bausenhart, Dyjas, & Ulrich, 2014; Jazayeri & Shadlen, 2010; Wehrman, Wearden, & Sowman,

2018a; Wiener, Thompson, & Coslett, 2014). Similar to in other domains, decisions related to prior durations tend to carry over; deciding one duration is short can lead to another duration also being judged as short (e.g., Brown, McCormack, Smith, & Stewart, 2005; Wiener et al., 2014; see Wearden, 2016, for an introduction to temporal decision-making). In the current article, we further investigate the parameters within which this 'decisional carryover' occurs in the judgment of duration. Finding similarities and differences between, for example, visual and temporal carryover effects can be informative about the mechanisms by which timing differs from other tasks, and how the processing/judgment of duration is (dis)similar to other carryover processes.

In the first two experiments performed here, we further investigate the parameters within which decisional carryover occurs. In the third experiment, we present participants with stimuli which are judgable in terms of their size and duration. By asking participants to attend specifically to one of these dimensions, we compare the carryover effects between the visual and temporal modalities. These experiments are largely based on the findings of carryover effects on duration judgments found in Wehrman et al. (2018a) and

✉ Jordan J. Wehrman
jordan.wehrman@mq.edu.au

[1] Macquarie University, Macquarie Park, Australia

[2] Keele University, Newcastle, UK

Wiener et al. (2014). It is worth acknowledging the existence of carryover effects found in other temporal tasks, such as reproduction (see Bausenhart et al., 2014; Jazayeri & Shadlen, 2010). However, in the current article, we are primarily interested in decisions about duration rather than duration reproductions.

To examine these effects, we performed a series of three temporal bisection experiments. The reasoning for using temporal bisection is presented below; however, in brief, this task requires participants to judge the duration of a stimulus in relation to preestablished standard (reference) durations. At the outset of the experiment, participants are taught to recognize a short and long extreme; for example, in Experiment 1, we use 400 ms and 1,100 ms, respectively. Following this learning phase, participants are asked to categorize intermediary intervals—for example, 800 ms—as closer to either the short or long reference. By analyzing participants' resultant responses, we can determine the duration at which the subjective midpoint occurs between these two references, as well as other features relating to perception, as will be discussed below.

Temporal decisional carryover

In both the Wehrman et al. (2018a) and Wiener et al. (2014) studies, a modified version of the temporal bisection task was used. In the Wehrman et al. (2018a) study, a 'reminder' task was employed in which, prior to a test duration being shown, participants were reminded of either the long or short reference duration. Participants were not required to judge this reference duration. Following the reminder, participants judged whether the test duration, which fell somewhere between the short and long reference durations, was closer to the short or long reference. In the Wiener et al. (2014) study, on the other hand, a 'reference-free' bisection method was used. In this task, participants were shown three examples of the objective middle duration at the beginning of the experiment. Subsequently, participants judged whether a test duration was of a longer or shorter duration than the average duration presented in the experiment.

In both these experiments, a decisional carryover effect was found. In the Wehrman et al. (2018a) study, participants judged the test duration as closer to the immediately preceding reference duration. For example, a short duration reminder led to an increased probability of the participant judging the test duration as 'short'. In the Wiener et al. (2014) study, the decision the participant made regarding the prior test duration tended to be repeated; a prior 'short' choice resulted in a higher probability of a subsequent 'short' decision. Interestingly, in addition to the carryover effect of the prior decision, Wiener et al. (2014) also showed a contrast effect to the objective prior duration; an objectively longer stimulus

resulted in a subsequent stimulus being perceived as shorter, though behaviourally this effect only occurred when the stimulus was auditory and not when it was visual. Modelling by Wiener et al. (2014), however, described a contrast effect in the visual domain as well, but this was overshadowed by the decisional carryover component. The pattern of *objective* contrast and *subjective* assimilation is similar to findings in nontemporal visual judgment tasks such as those obtained in Hampton, Estes, and Simmons (2005) and Jones, Love, and Maddox (2006). Further, Brown et al. (2005), in investigating judgments of auditory duration and frequency, similarly found a strong carryover effect from the prior trial decision.

Both the Wehrman et al. (2018a) and Wiener et al. (2014) studies proposed that assimilation effects were larger when the duration of the test stimulus was difficult to judge compared with when it was easy. Wehrman et al. (2018a) propose that this is due to an anchoring effect from the prior reminder reference. When a participant was unsure of the duration of the subsequent test duration, the boundaries of what constituted an acceptable guess for that duration were larger. Such conditions lead to a larger anchoring effect,[1] and thus stronger assimilative carryover. In the Wiener et al. (2014) study, decision bias was proposed to be due to response uncertainty. When participants were unsure of the duration of the test stimulus, they tended to repeat their most recent response. When participants were uncertain, responses were sticky. Interestingly, while Brown et al. (2005) also found decisional carryover, they used similar assimilation patterns to argue for a common judgment pattern across auditory frequency and temporal judgments (a proposal we support here).

In the current article, we firstly attempt to replicate the standard finding of decisional carryover found in Brown et al. (2005) and Wiener et al. (2014). Additionally, we expand these findings to the visual domain (not assessed in Brown et al., 2005) and with a more standard bisection task than that used in Wiener et al. (2014). In the second experiment, we further examine how decisional carryover is realized when extreme durations are presented on every other trial. Finally, in the third experiment, visual size and duration are modulated in a single experiment, allowing us to assess the similarities between these two judgment tasks. This third experiment expands on the findings by Brown et al. (2005) by assessing the visual modality, and, more interestingly, examining the carryover effects in a single experimental design, allowing a more thorough comparison of performance. To begin with, we discuss the use of the bisection task.

---

[1] For information about anchoring, see Bravo and Mayzner (1961), Chapman and Johnson (1994), Furnham and Boo (2011), Larimer (1965), Parducci and Marshall (1962), and Sherif, Taub, and Hovland (1958). The reason anchoring has a difficulty dependent effect is due to the proposed adjustment until reaching a plausible bound for the stimulus property. The more uncertain, the further the bounds are, and the stronger the anchor-and-adjust heuristic will be. See Lieder, Griffiths, Huys, and Goodman (2017).

Bisection

All the experiments performed in the current article use a bisection task, as introduced above (e.g., Church & Deluty, 1977; Wearden, 1991; Wearden & Ferrara, 1995, 1996). By fitting a psychometric function to the resulting decisions, we can determine both the bisection point (BP) and Weber ratio (WR). The BP is the duration at which participants are equally likely to decide that a given duration is either short or long. This is the primary measure used in the current experiments, as it can indicate a general shift in the perceived duration of a stimulus. If decisional carryover occurs, a prior 'short' response should lead to a higher chance of another 'short' response in the current trial, leading to a longer BP.

The WR is an index of discriminability, with a smaller WR indicating higher sensitivity to duration.[2] While not as indicative of carryover at the group level, Wiener et al. (2014) correlated WR differences between 'short' and 'long' prior decisions with BP point differences between those prior decisions. They found a positive correlation, indicating that as the difference in discriminability raised between these two conditions, so too did the BP. This was used as evidence indicating that when participants found the task more challenging, carryover effects were larger. Thus, the WR, while interesting in its own right, is also useful in the theories put forward by Wehrman et al. (2018a) and Wiener et al. (2014) in quantifying decision difficulty. Additionally, other theories related to serial dependencies, such as provided in Mori (1989), speculate that carryover is more likely when a decision is made with relatively little information, an issue to which WR is indicative.

Further, from the bisection data we can also examine the point of maximal uncertainty (PMU; see Birngruber, Schröter, Schütt, & Ulrich, 2017; Birngruber, Schröter, & Ulrich, 2014; Dyjas & Ulrich, 2014; Wehrman, Wearden, & Sowman, 2018b, for discussion and applications of the method). This is defined as the maximum of the reaction time (RT) distribution plotted over the tested durations. The maximum is found by using waveform moment analysis[3] (Cacioppo & Dorfman, 1987) and allows identification of the duration at which participants are least confident in their response. This point normally corresponds to the BP, providing added support to bisection results (see Balcı & Simen, 2014; Simen, Balci, Cohen, & Holmes, 2011). Like changes in the WR, the difficulty of making a decision, as indexed by the PMU, is informative of those conditions in which decisional carryover is likely to occur.

The bisection method has been used here for several reasons. Specifically, in the reference-free method used by Wiener et al. (2014), the mean duration used for judgment is altered by the durations presented throughout the experiment. In the model by Wiener et al. (2014), the mean duration criterion was proposed to be leaky, resulting in more recent trials having a stronger effect on the mean (i.e., standard) duration used for comparison of a test stimulus as 'long' or 'short'. This means that the judgment criterion in the Wiener et al. (2014) study was updated in relation to the durations experienced most recently. Using a reference-free task may result in more judgment uncertainty, due to its basis in recent experience, leading to larger carryover effects.

Though the subjective midpoint may similarly be affected by recent experience in the standard bisection task, the criteria used for judgment in the bisection task are supposedly the short and long reference durations which should be unaffected by recent experience. If assimilative carryover is seen in the current experiments, while also adding general support for the effect, it would show some robustness to the type of task used to find such effects. Relatedly, finding an assimilative carryover effect would demonstrate that the assimilative effect is in the judgment of the target duration, rather than in the judgment of the mean duration (though unlikely, this is nonetheless a possibility in the reference-free task).

Finally, due to its prominence in the interval timing literature, carryover effects in the standard bisection task should be assessed. The Wiener et al. (2014) and Wehrman et al. (2018a, b) methods are nonstandard for time perception research. Using a reference-free method is somewhat similar to that of Wearden and Ferrara (1995), where an implicit mean duration is theorized to provide the basis for bisection. However, Wiener et al. (2014) provided an explicit central duration, and participants were instructed to use the mean duration of the stimuli presented for their judgment. Due to this instruction, the subjective midpoint is likely to shift from trial to trial, which, as mentioned above, may have contributed to the carryover effects reported. Wehrman et al. (2018a, b), on the other hand, presented reference durations prior to the test stimuli. While this was aimed at biasing responses, it is again a nonstandard bisection method. These differences may influence whether assimilative carryover occurs and thus are worth considering by running a standard bisection task.

Current experiments

In Experiment 1, we simply investigated assimilative carryover effects using a standard temporal bisection procedure. Aside from using the bisection task, we also did not control for trial-to-trial sequences (which was done in the Wiener et al., 2014, study by the use of a de Bruijn sequence). This was done to establish whether assimilative carryover was task

---

[2] Specifically, it measures the slope of the distribution: $[(p(long) = .75) - (p(long) = 0.25)] / 2 / BP$
[3] In brief, this involved normalizing each mean RT for each duration in each condition by dividing by the sum of the mean RTs for each participant. Each duration was then multiplied by its corresponding weight, and the results were summed within each condition for each participant.

dependent, only occurring under a strict parameterization of the bisection task. Further, we extend Brown et al. (2005) by investigating the visual modality with a 'long' to 'short' (L:S) ratio intermediary to those used previously.

In Experiment 2, every other test trial had the same duration as one of the two reference durations from the initial learning phase of the task. In other words, the trial duration sequence was Reference (long or short) ➔ Test ➔ Reference ➔ Test. Participants were not informed of this pattern and did not report noticing it. Participants were still required to judge the 'easy' reference duration trials as well as the more difficult test duration trials. This experiment essentially used the reminder methodology from Wehrman et al. (2018a, b), but inserted a response after the reminder.

Asking participants to judge reference durations on every other trial examines whether a highly probable response (i.e., a highly likely 'long' response when shown a long reference) affects subsequent judgments. Given that assimilative carryover was found when judging the prior test duration (Brown et al., 2005; Wiener et al., 2014), and when not judging the prior reference duration (Wehrman et al., 2018a), we expect that judging the prior reference duration should also lead to assimilative carryover. However, it is possible that when the judgment is easier, and a response is added between the reference and test durations, assimilative carryover does not occur. This would be in line with theories such as Mori (1989) and Ward and Lockhead (1971) regarding the judgment of nontemporal stimuli. Additionally, we also examine whether the judgment of reference durations is affected by the prior test duration. Because the reference duration judgments are relatively easy, if an assimilative carryover effect occurs only when participants are unsure of the duration of the current stimulus, then the judgment of reference durations should not be prone to decisional carryover

In Experiment 3, we varied the size and duration of a visual stimulus simultaneously; a single stimulus could be both 'large' or 'small', and simultaneously 'long' or 'short'. Participants only attended to either size or duration—that is, we used a between-subjects design, though a single stimulus could vary along both dimensions. In the test phase, each trial presented a to-be-judged duration/size stimulus varying between both the reference durations and the reference sizes. This allowed us to directly compare the sequential effects of both size and duration in one experiment. This study aimed to replicate previous findings while varying the physical dimensions of the reference/target stimulus. A minor issue addressed in Experiment 3 is whether duration carryover effects are only seen when the stimulus does not vary. It could be that carryover effects are only seen when stimuli are unchanging; perhaps varying the size of the stimulus as the irrelevant dimension, even though participants judged the duration of the stimulus, would diminish the effect.

Of more interest, judgment carryover effects are proposed to occur primarily when participants are unsure of their response. In the current experiment, the ratio of the longest to shortest duration, and the ratio of the largest to smallest circle, were the same (4:1), and thus the task should be equally difficult despite the judged dimension. However, if one judgment dimension is more accurate than another (i.e., size estimation is more accurate than duration judgments or vice versa), as measured by the WR, there may also be an effect on the decisional carryover. This is expected to be the case given that other research has found the judgment of time to be more difficult than the judgment of spatial dimensions (e.g., Ogden, Samuels, Simmons, Wearden, & Montgomery, 2018). Brown et al. (2005) previously found qualitatively the same carryover effect in the bisection results of auditory duration and frequency; however, the ratio between the extremes in these experiments varied. Further, separate experiments were run such that only one of the elements varied. In Experiment 3, we expand on Brown et al. (2005) by requiring different judgments on the same task.

Method

Participants

Twenty participants were tested in each experiment. If the BP fell outside the tested durations, then that participant was replaced. Participants were paid or given course credit in exchange for their participation. Experiments were conducted according to the Declaration of Helsinki.

In Experiment 1, one participant was replaced. The mean age of those used was 22.3 years (SD = 4.3), with 10 males and two left-handed participants. In Experiment 2, four participants were replaced. The mean age of those used was 25.2 years (SD = 8.7), with 12 males and one left-handed participants. In Experiment 3, no participants were replaced. The mean age was 21.8 years (SD = 4.6), with four males and one left-handed participant.

Materials

Experimental stimuli were presented on a Samsung SyncMaster SA950 (27 inch) monitor controlled by a Dell Optiplex 9010 PC running 64-bit Windows 7. All experiments took place in dimly lit rooms, with participants seated 0.8 m away from the monitor. Neurobehavioral Systems' Presentation (Version 18.3) software was used to present the experiments.

The exact procedure for each experiment varied and is described in the relevant section. However, generally, participants completed a learning phase in which they were shown the reference standards with which to compare the test stimuli.

Following this, participants performed a test phase where they judged stimuli as closer to one or the other reference.

Analysis

The first five trials of each experiment and the first trial of each block were removed from analysis. Trials with RTs exceeding 3,000 ms were excluded from analysis.[4] Psychometric functions were calculated based on the proportion of times that a participant chose 'long' in all experiments, and also 'large' in Experiment 3. A cumulative Gaussian distribution was fit using the psignifit software package, Version 4 (see http://https://github.com/wichmann-lab/psignifit/wiki/; see Schütt, Harmeling, Macke, & Wichmann, 2016, for details on the implementation of this program). In brief, psignifit uses numerical integration to estimate psychometric functions from Bayesian inference. From the output of this program, we retrieved the BP and the WR.

We additionally found the PMU of the RT distribution across the durations or sizes, defined as the point of the objective measure corresponding to the maximum of the RT distribution. The maximum was found by using waveform moment analysis, as mentioned above (Cacioppo & Dorfman, 1987).

For both the psychometric functions, and the PMU calculation, we separately analyzed the data from the current trial grouped by both the objective prior stimulus and the response to the prior stimulus. Thus, we generated psychometric functions for whether the prior response was 'short' or 'long' and separately for whether the objective prior duration was short, middle, or long. The data were analyzed for the objective prior duration, and the prior response separately.

Statistical analysis was performed using R (R-Core-Team, 2015) and the package 'ez' (Lawrence, 2013). We ran separate tests on three dependent measures: the BP, the WR, and the PMU. Cohen's d is reported for t tests and partial eta squared (η_p^2) is reported for ANOVA results. The Holm correction is used for post hoc analysis (Holm, 1979). Any additional analysis is described in the relevant section.

Experiment 1: Sequential judgments

In this experiment, participants completed a standard bisection task. This was done to establish whether decisional carryover occurs in a standard rendition of the bisection task with uncontrolled trial to trial contingencies.

---

[4] Analyzing trials with RTs only under 1,000 ms, as done in Wiener et al. (2014), due to a possible RT limit on carryover effects (Wichmann & Hill, 2001), yielded a similar result.

Procedure

Both references, and all targets, consisted of solid white circles (125-pixel diameter) presented for various durations. The learning phase consisted of 30 trials, 15 of each reference duration, presented in a random order. The short reference was 400 ms and the long reference was 1,100 ms in duration. Each of these trials started with a fixation cross ('+') presented centrally for 500 ms, followed by a 300-ms blank screen. One of the two references was then presented, followed by another 300-ms blank screen prior to the next trial. At the end of the 30 trials, participants were shown two of each of the references, in the same format as during the learning phase, with an indefinite interval after the reference. They were asked to judge whether these were the long or short references. This was done to ensure that participants were able to discriminate the two reference durations.

The test phase consisted of four blocks of 72 trials with self-paced breaks between blocks. Each trial was in the same format as above, except that after the presentation of the circle an indefinite gap followed until a response was made. Participants responded with the 'S' key if they judged that the test interval was closer to the 'short' reference, and the 'L' key if they judged it as closer to the 'long' reference. Stimulus durations were arithmetically spaced at 500 ms, 600 ms, 700 ms, 800 ms, 900 ms, and 1,000 ms. Trials were programmed in pairs such that each pair order (e.g., 500 ms followed by 700 ms) was set to occur once per block, but we did not control for the order of pairs, resulting in a variable trial-by-trial experience overall (see Fig. 1).

For analysis, a 500-ms or 600-ms trial was defined as 'short,' a 700 ms or 800 ms trial was defined as 'medium', and a 900 ms or 1,000 ms trial was defined as 'long.' The psychometric functions analyzing the effects of the prior objective duration were built around which of these groups the current trial was preceded by.

Results

Of the trials, 0.8% were discarded due to them having associated response RTs longer than 3,000 ms. Average responses are shown in Fig. 2.
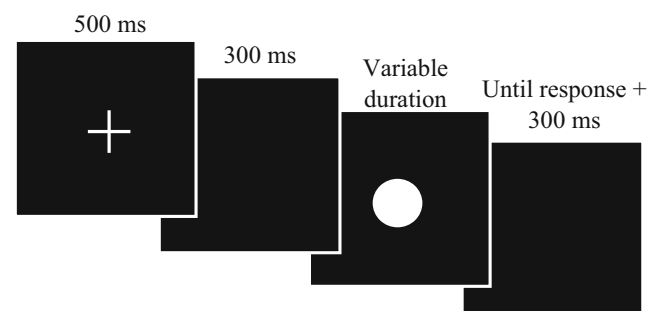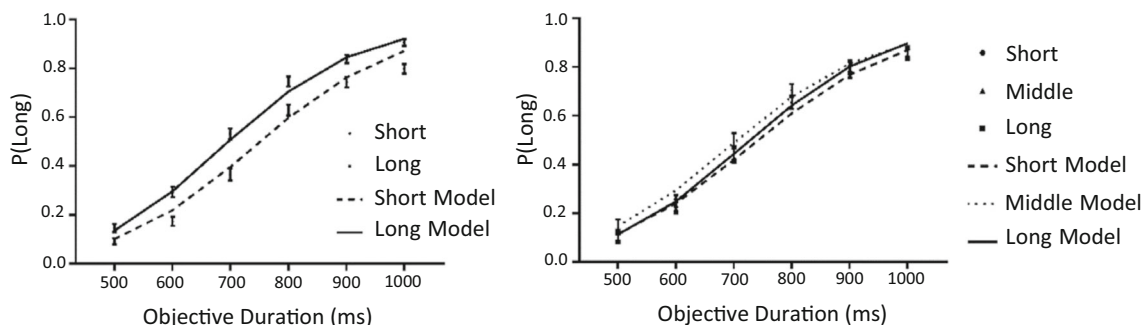

Fig. 1 Trial design of Experiments 1 and 2 during the test phase

Fig. 2 Left: Probability of choosing 'long', following either a 'short' or 'long' prior response. Right: Probability of choosing 'long', following a short, middle, or long objective prior to stimulus duration. Lines for both graphs represent the fit of the model (as described below). Error bars represent one standard error of the mean

BP: Prior response The BP was significantly longer following a 'short' prior choice (757 ms) compared with a 'long' prior choice (706 ms), $t(19) = 2.17$, $p = .043$, $d = .54$ (see Fig. 3, top left). This indicates that if a participant decided that the preceding trial was closer to the short reference, then the participant was more likely to perceive the subsequent trial as 'short' as well.
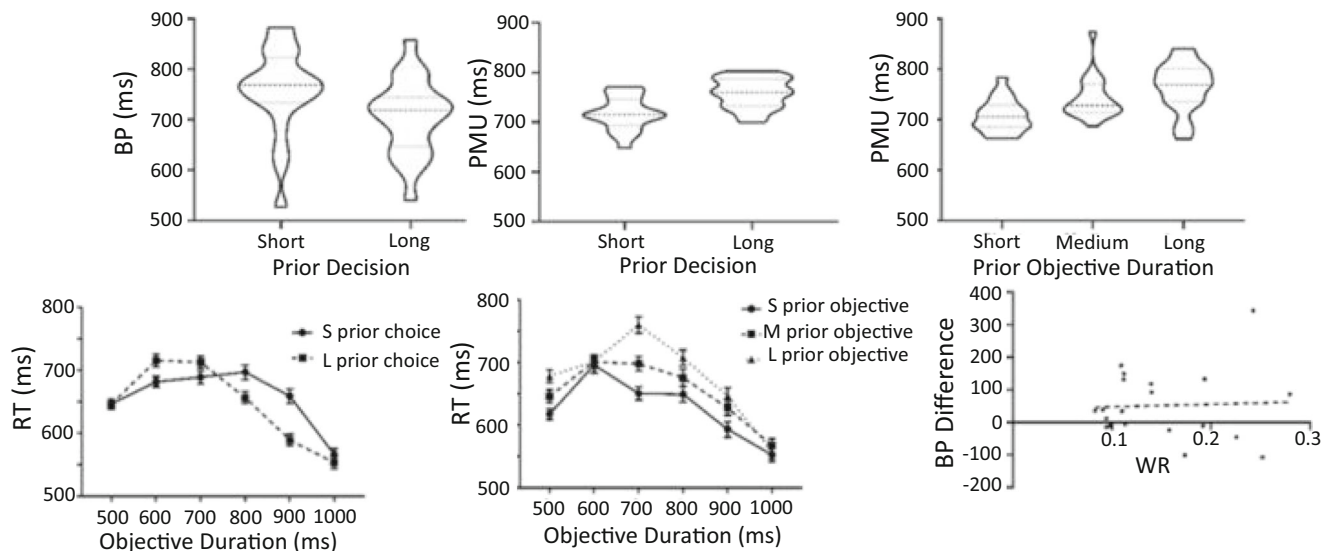
BP: Prior objective duration There was no significant effect of the objective prior duration on the perceived duration of the current target, $F(2, 38) = .188$, $p = .829$, $\eta_p^2 = .01$. The mean BP was at 734 ms.

WR Neither the prior response, $t(19) = .694$, $p = .496$, $d = .08$, nor the prior objective duration, $F(2, 38) = .149$, $p = .862$, $\eta_p^2 = .01$, significantly affected the WR. The mean WR was .151 and .155 in the respective tests.

In addition to testing the WR and BP separately, it is interesting to ask whether the WR correlates with the BP. To test this, we examined the Pearson correlation between the average WR with the difference between the BP given a 'short' or 'long' prior decision. In other words, this analysis asks whether participants are more likely to demonstrate decisional carryover if they were more unsure of response in general.

This analysis showed no correlation between BP difference and WR ($r = .05$, $p = .850$). Thus, it appeared that participants were demonstrating decisional carryover irrespective of their uncertainty in judging stimulus duration (see Fig. 3, bottom right).

Model fitting To confirm the results of the BP analysis, we fitted a model derived from scalar expectancy theory (SET; Gibbon, Church, & Meck, 1984; Meck, Church, & Olton, 1984; Wearden, 1991). The model was developed from Wearden and Jones (2013), used in Droit-Volet, Wearden, and Zélanti (2015), and applied to assimilative carryover in Wehrman et al. (2018a). These ideas were developed in Droit-Volet and Wearden (2001).

Fig. 3 Top left: BP following a short or long response in the prior trial. Middle: PMU following a short or long prior response. Right: PMU following a short, medium, or long prior objective duration. Bottom left: Chronometric curve based on prior decision. Middle: Chronometric curve based on the prior objective duration. Right: Correlation between BP difference and WR. All error bars represent one standard error of the mean. Dashed line in violin plots represent median, and dotted lines represent quartiles

This model combines four SET processes into two parameters: a sensitivity parameter, c, and a displacement parameter, d. Changes in the sensitivity parameter, c, flatten or steepen the curves shown in Fig. 2. Higher c values make the curve flatter and lower c values make it steeper. The d value shifts the curve on the x-axis, with smaller d values shifting it to the left and larger d values to the right. To fit the model, we transformed the short and long reference durations into Gaussian distributions with means equal to the reference duration, and a coefficient of variation, c. On each trial, a value of the short and long reference duration (S* and L*) were randomly chosen from these distributions and multiplied by d. We assume the probe duration (T) is timed without error, and any variance is absorbed by the reference duration. The decision rule for a given probe was if $T - dS^* < T - dL^*$ the model chose short, otherwise the model chose long. The two parameters (c and d) were varied over a wide range to obtain the best fit. We fit the model separately based on whether the prior trial was objectively long, medium, or short, and whether the previous decision was 'short' or 'long'. Each data point was simulated 10,000 times. Fits were quantified using mean absolute deviation (MAD), the absolute difference between the data points and model's predictions divided by the number of data points.

For the objectively long, medium, and short duration previous trials, MAD values were 0.01, 0.02, and 0.02, indicating good model fits. There did not appear to be a consistent pattern in terms of either c (0.34, 0.37, and 0.36, respectively) or d (0.94, 0.90, and 0.95, respectively) with respect to the duration of the stimulus on the previous trial. For the 'long' and 'short' prior responses, MAD values were good (0.01 and 0.02, respectively). The value of c did not vary depending on the prior response participants made ($c = .35$ for both). However, there was a larger bias value when participants previously chose 'short' ($d = .97$) compared with 'long' ($d = 0.90$). The direction of this bias was the same as that found in Wehrman et al. (2018a), and of a similar magnitude to the results of Experiment 2 from that article.

PMU: Prior response The PMU (see Fig. 3, top center) was significantly later following a 'long' (759 ms) rather than 'short' prior response (718 ms), $t(19) = 2.97$, $p = .008$, $d = .66$.

PMU: Prior objective duration The PMU (see Fig. 3, top right) was significantly later given a longer objective prior trial, $F(2, 38) = 6.63$, $p = .003$, $\eta_p^2 = .26$. A short objective duration led to a PMU of 708 ms, the medium objective duration led to a PMU of 743 ms, and a long objective duration led to a PMU of 765 ms. Subsequent Holm-corrected analysis showed that a previously short duration stimulus resulted in a significantly smaller PMU compared with both the medium, $t(19) = 2.62$, $p = .034$, $d = .59$, and long, $t(19) = 3.79$, $p = .004$, $d = .85$, objective prior stimuli. The medium and long objective prior

stimuli were not significantly different, $t(19) = 1.21$, $p = .241$, $d = .27$. Chronometric graphs depending on both the prior objective (see Fig. 3, bottom right) and prior choice (Fig. 3, bottom center) are included below.

RTs Because the effect of the prior trial on the PMU was in the opposite direction to the BP, we performed an extra exploratory analysis. This was to test whether the pattern of RTs was related to a phenomenon known as the variable foreperiod effect. The variable foreperiod effect describes how RTs tend to be shorter if people have longer to prepare for a response (see Los, 2010, 2013). For example, if a 'go' signal could occur after either 400 ms or 800 ms from the start of a trial, RTs will tend to be shorter if the signal is after the 800-ms interval. Further, RTs also tend to be shorter if the prior trial was shorter (i.e., had an interval of 400 ms rather than 800 ms). In the current experiment, perhaps waiting longer to be able to respond acts in a similar way, resulting in RTs being shorter after a longer duration stimulus. Alternatively, it could be that participants were specifically faster when repeating their choices (i.e., when responding 'short' in both trials).

For the exploratory ANOVA, mean RTs were the dependent variable, and 'short' and 'long' current and prior responses were the independent factors. There was a main effect of prior response, $F(1, 19) = 13.49$, $p = .002$, $\eta_p^2 = .42$, indicating that RTs were slower when the prior response was 'long' compared with 'short' (mean RTs = 667 ms and 627 ms, respectively). There was also a main effect showing that the current response was faster after a 'long' response (602 ms) rather than a 'short' one (692 ms), $F(1, 19) = 26.25$, $p < .001$, $\eta_p^2 = .58$. The interaction effect was not significant, $F(1, 19) = 1.72$, $p = .205$, $\eta_p^2 = .08$. This pattern was consistent with a variable foreperiod explanation.

Experiment 2: Judged references

The findings of Experiment 1 indicate that responses in a standard temporal bisection task do exhibit decisional carryover effects. In Experiment 2, we attempt to extend these findings by asking whether decisions regarding reference durations affect subsequent decisions regarding probe durations, and vice versa. Because Wehrman et al. (2018a) found that an unjudged reference duration results in assimilative carryover to a subsequent probe duration, and Experiment 1, as well as Wiener et al. (2014), found decisional carryover from one trial to the next, we expect that decisions regarding a reference duration should affect the reported duration of a subsequent probe duration. However, because both Wehrman et al. (2018a) and Wiener et al. (2014) propose some form of difficulty-based carryover, there should not be a carryover effect from the decision regarding the probe duration to the decision of the subsequent reference duration.

Procedure

This experiment was run identically to Experiment 1, except that every other trial in the testing phase presented one of the reference durations from the learning phase of the task (i.e., on Trials 1, 3, 5, etc.). Participants experienced three of each reference–probe pairing (e.g., 400-ms reference followed by 900-ms test, or 1,100-ms reference followed by 800-ms test) per block, for 36 trials total. Participants completed five blocks of trials. Trial pairs were presented in a random order. Psychometric functions were generated as per Experiment 1, except that the objective prior duration could only be long or short.

Results

Of the trials, 0.4% were discarded due to RTs slower than 3,000 ms. Figure 4 shows mean responses.

BP There was no significant effect of either the objective duration of the prior reference, $t(19) = .657$, $p = .519$, $d = .08$, mean BP = 736 ms, nor the judgment of the prior reference, $t(19) = .832$, $p = .416$, $d = .05$, mean BP = 738 ms, on the BP. This lack of effect on the BP was unexpected; therefore, we calculated the Bayes factor for the comparison of the BP given a 'long' or 'short' prior response using the BayesFactor package for R (Morey, Rouder, Jamil, & Morey, 2015). The Bayes factor was $.249 \pm .02\%$[5]—moderate support for accepting no difference between the conditions (Jeffreys, 1961).

WR The WR was not significantly affected by either the objective duration of the prior reference, $t(19) = .696$, $p = .495$, $d = .14$, mean WR = .135, or the decision regarding that reference, $t(19) = .849$, $p = .406$, $d = .15$, mean WR = .134.

As per Experiment 1, we were interested in whether the differences between 'long' and 'short' prior decisions in BP and overall WR were correlated. In this experiment, there was a significant correlation between WR and BP difference ($r = 0.70$, $p = .001$). This medium-strength association indicated that the higher the WR, the larger the decisional carryover effect (see Fig. 5, top right).

Model fitting We fit the same model as above to the data from the current experiment. We fit separate models depending on if the prior reference duration was long or short, and whether participants chose 'long' or 'short'.

The model fits were reasonable for each of the four conditions (MAD values all 0.03). For the objective duration of the prior stimulus, there appeared to be very little variation in parameters, confirming the BP results: If the objective duration was long, then $c = 0.34$ and $d = 0.94$, whereas if the prior

objective duration was short, then $c = 0.35$ and $d = 0.94$. In terms of the subjective duration of the prior reference duration, $c = .32$ and $d = 0.95$ if the prior choice was 'long', and $c = .38$ and $d = 0.94$ if the prior choice was 'short'. This fit showed no difference in response bias (d) between the conditions, as per the BP results. Interestingly, there appeared to be slightly larger c values if the prior choice was 'long', indicating the fit was slightly flatter following a 'long' decision.

PMU The PMU was significantly later given a 'long' prior decision (763 ms) compared to a 'short' prior decision (721 ms), $t(19) = 2.58$, $p = .018$, $d = .58$. Similarly, an objectively longer prior duration anchor (764 ms) resulted in a longer PMU compared with a short prior objective duration (719 ms), $t(19) = 2.74$, $p = .013$, $d = .61$. Both these effects are shown in Fig. 5, top left.

RTs We again analyzed RTs as per Experiment 1. This was done to confirm the initially exploratory analysis. There was a main effect of current response, $F(1, 19) = 29.04$, $p < .001$, $\eta_p^2 = .60$, 'short' response RTs = 729 ms, 'long' = 649 ms, and of prior response, $F(1, 19) = 5.15$, $p = .035$, $\eta_p^2 = .21$, 'short' = 676 ms, 'long' = 707 ms). There was no interaction effect, $F(1, 19) = 1.04$, $p = .321$, $\eta_p^2 = .05$. These findings confirmed the results in Experiment 1.

Reference duration judgments We additionally analyzed the probability that a participant chose 'long' to the reference trial based on both the objective duration of the prior test interval and whether participants chose 'short' or 'long' to the preceding test interval. An ANOVA regarding the objective prior duration showed a main effect of whether the judged reference duration was short or long, $F(1, 19) = 1726.9$, $p < .001$, $\eta_p^2 = .99$, indicating that if a participant was shown a long reference they were more likely to choose a 'long' response (probability 'long' = .922, compared with .035). Neither the main effect of prior objective duration, $F(2, 38) = .185$, $p = .832$, $\eta_p^2 = .01$, nor the interaction between prior objective duration and the duration of the judged reference, $F(2, 38) = .112$, $p = .895$, $\eta_p^2 = .01$, reached significance.

When examining the effects of the prior response on the response regarding the reference duration, the ANOVA showed a main effect of whether the participant was presented a short or long reference stimulus, $F(1, 19) = 1855.1$, $p < .001$, $\eta_p^2 = .99$, indicating that if a participant was shown a long reference they were more likely to choose a 'long' response (probability 'long' = .927 compared with .034 if the reference was short). There was also a main effect of the response the participant made in the prior trial, $F(1, 19) = 11.91$, $p = .003$, $\eta_p^2 = .39$ (see Fig. 5, top center). This effect showed that if participants chose 'long' to the preceding test duration, they were more likely to choose 'long' again to the reference duration (probability long = .498, compared with .462 following

---

[5] Or, equivalently, the Bayes factor was 4.02 in favour of the null hypothesis.
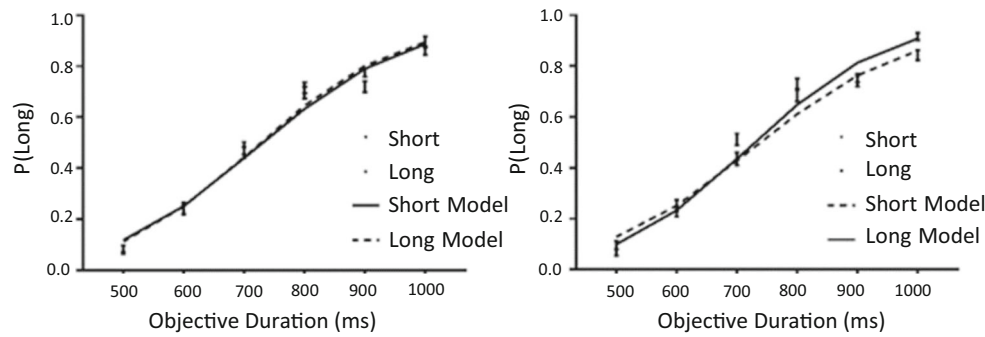
Fig 4 Left: Probability of choosing 'long' following either a 'short' or 'long' prior response. Right: Probability of choosing 'long' following a short or long objective prior stimulus duration. Lines represent the model fit to the data. Error bars represent one standard error of the mean

a 'short' prior response). Figure 5 shows these effects separated out between the two reference durations. There was no significant interaction effect, $F(1, 19) = 2.96$, $p = .102$, $\eta_p^2 = .13$.

Comparison of WRs in Experiments 1 and 2 Because the results between Experiments 1 and 2 were unexpectedly different, it is interesting to test whether the WRs were different between the experiments. This is because both Wiener et al. (2014) and Wehrman et al. (2018a, b) used explanations which employed an uncertainty-based decision effect. Perhaps in Experiment 2 decisions were less uncertain and therefore there was no effect of prior decisions on the BP. However, a between-subjects t test revealed no significant difference in the WRs, $t(32.0) = 1.02$, $p = .318$, $d = .31$. The Bayes factor weakly favoured the null hypothesis, but the amount of evidence provided is only anecdotal (BF = .464 ± .01%).[6]

Experiment 3: Sequential size versus time bisection

Experiment 1 demonstrated that decisional carryover occurred in a standard bisection task. However, unlike this finding, Experiment 2 demonstrated that when an 'easy' reference duration was judged prior to the judgment of an intermediary probe duration, decisions did not carryover. In Experiment 3, we attempt to extend these findings by asking participants to perform a bisection task with regards to either duration or size. If the difficulty of the current decision determines whether decisional carryover occurs or not, there might be differences in the carryover seen in the judgment of size, an easier judgment dimension, compared with duration, a more difficult judgment dimension. More generally, it is interesting to directly compare the decisional carryover in the bisection of time

and another type of stimulus. These results may be informative of differences in the judgment processes given a visual and temporal decision requirement.

Procedure

Both anchors, and all targets, consisted of solid white circles of varying sizes presented for various durations. Participants carried out one of two variations of this experiment, either judging the duration or the size of the stimuli. However, whichever judgment participants were required to make, the actual procedure was the same.

In the learning phase, participants were shown eight sets of each reference size–duration combination. There were two sizes (40-pixel and 160-pixel diameter) and two durations (300 and 1,200 ms), so, for example, a short anchor of 300 ms could be either 40 or 160 pixels in diameter. This resulted in 16 examples of each reference in the relevant dimension (e.g., eight examples of 300 ms in which the size was 40 pixels, and eight examples in which the size was 160 pixels). Note that the ratio of 'long'/'large' to 'short'/'small' is the same in both cases (4:1). Prior to each, the words 'long' or 'short' were presented for 500 ms if the participant was performing the duration judgment task. Otherwise, the words presented were 'large' and 'small.' As opposed to in Experiments 1 and 2, we thought it necessary to reinforce the stimulus aspect that the participants needed to attend to, due to both size and duration varying simultaneously. After the presentation of the reference, a 500 ms blank screen was presented, followed by the next trial. After the learning phase, each of the four anchors were presented once and the participant was required to identify them in the relevant way (i.e., size or duration) to ensure they had learned the references.

During the testing phase, there were five possible sizes (60, 80, 100, 120, 140 pixel diameter), and five possible durations (450, 600, 750, 900, 1,050 ms). There were 50 trials in each of the seven test blocks, of which there were 10 of each in the 'of-interest' dimension (either duration or size). Each size–duration combination was presented twice in each block. Each trial

---

[6] Using an ANOVA, with prior response as a within-subjects factor, and the experiment as a between-subjects factor, showed no effects on either the BP or the WR. This was not included here for brevity.
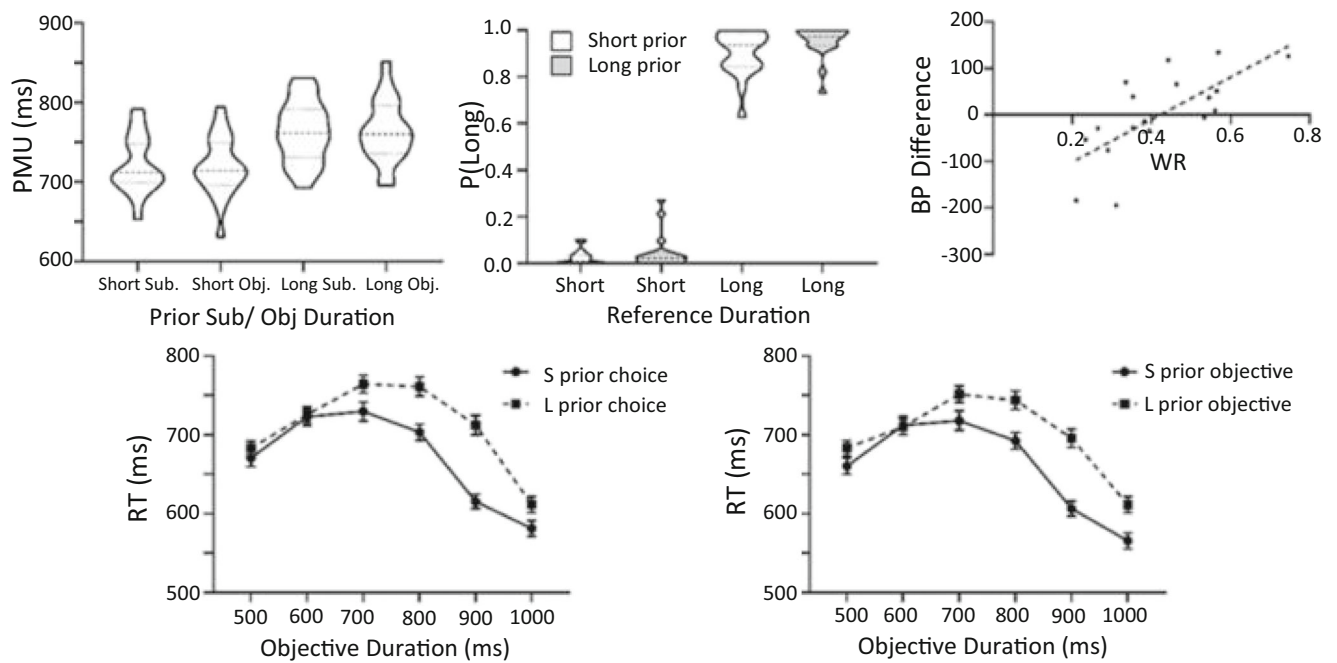
Fig. 5 Left: PMU based on objective prior duration (black) and the prior decision (grey). Top middle: Probability of a 'long' response to either a long or short reference, following a 'short' or 'long' response to the prior test duration (bar colour). Note, there was a main effect of prior response on responses—however, no interaction effect. Top right: Correlation between BP difference and WR. Bottom left and right: Chronometric curves given prior choice and prior duration, respectively. All error bars represent one standard error of the mean. In the violin plots, the dashed lines represent the median while the dotted lines represent the quartiles

consisted of the test stimulus, followed by a 300-ms blank screen. A question mark was then shown until participants responded. A 500-ms blank screen was shown after the response, then the next trial started (see Fig. 6). Like in Experiment 1, short (450 ms or 600 ms), medium (750 ms), and long (900 ms or 1,050 ms) were defined for each prior trial. The diameter of the prior circle was also defined as small (60–80-pixel diameter), medium (100 pixel), or large (120–140-pixel diameter).

Participants responded with the 'S' key if they judged that probe stimulus was closer to the 'small'/'short' reference, and with the 'L' key if the probe was judged to be closer to the 'large'/'long' reference. Because the frame of reference given either a size or duration judgement was different (i.e., size is measured in pixels and time in milliseconds), we normalized the objective duration/size by dividing the objective duration/size by the mean objective duration/size. This meant that a
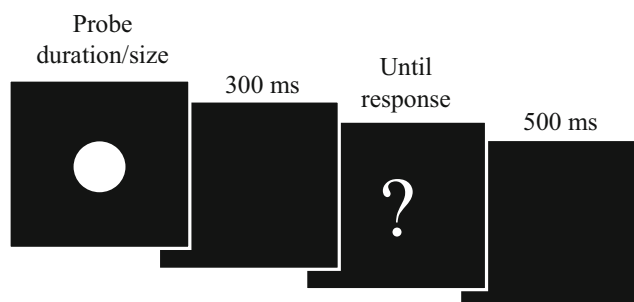

Fig. 6 Task procedure for the testing phase of Experiment 3

value of 1 corresponded to a size/duration objectively in the middle of the two references.
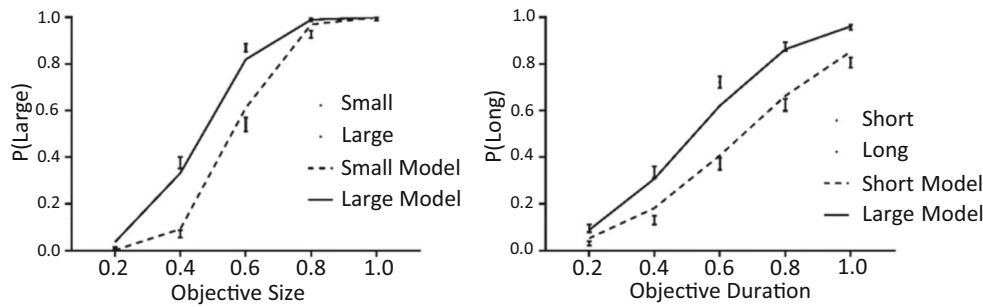
Results

Of the trials, 0.3% were discarded due to RTs longer than 3,000 ms.

BP: Prior response See Fig. 7. We analyzed prior response ('S' or 'L' for 'small'/'short' and 'large'/'long') as a within-subjects factor, and the task (judging size or duration) as a between-subjects factor in an ANOVA. There was no main effect of whether the participant judged size or duration, though this effect approached significance, $F(1, 18) = 3.25$, $p = .088$, $\eta_p^2 = .15$. However, there was a main effect of whether the participant responded 'S' or 'L' in the prior trial, showing that after an 'S' response the BP was 1.032, and after an 'L' response the BP was .896, $F(1, 18) = 43.71$, $p < .001$, $\eta_p^2 = .71$ (see Fig. 9, top left). The interaction between the prior response and the task approached significance, $F(1, 18) = 4.41$, $p = .050$, $\eta_p^2 = .20$.

Because the statistical significance of the interaction effect was borderline, we further investigated the effect of whether participants responded 'S' or 'L' for duration and size judgments separately. The difference between an 'S' and 'L' prior choice when judging duration was .180, $t(9) = 4.73$, $p = .001$, Holm-corrected, $d = 1.49$, while the difference when judging size was half this size at .093, $t(9) = 5.81$, $p < .001$, Holm-

Fig. 7 Left: Probability of choosing 'large' following either a small, middle, or large prior objective size. Right: Probability of choosing 'long' following a short, middle, or long objective prior stimulus

duration. Lines represent the fit of the models used. Error bars represent one standard error of the mean

correct, $d = 1.84$. While the difference in subsequent judgements between 'S' and 'L' prior choices were larger when judging duration, the actual effect size of prior choice was larger when judging size, indicating a more consistent effect in this case.

Prior objective stimulus See Fig. 8. There was no main effect of the objective prior stimulus value on the BP, $F(2, 36) = .376$, $p = .689$, $\eta_p^2 = .02$. The main effect of whether participants judged time or size again approached significance, $F(1, 18) = 3.46$, $p = .080$, $\eta_p^2 = .16$. The interaction effect was not significant, $F(1, 18) = .098$, $p = .907$, $\eta_p^2 = .01$. The mean BP was .961.

WR: Prior response The main effect of judgment criteria was significant (size: .071, time: .157), $F(1, 18) = 43.88$, $p < .001$, $\eta_p^2 = .71$ (see Fig. 9, top center). Neither the effect of anchor, $F(1, 18) = .142$, $p = .711$, $\eta_p^2 = .01$, nor the interaction between anchor and judgment reached significance, $F(1, 18) = .104$, $p = .750$, $\eta_p^2 = .01$.

Testing the Pearson correlation of BP differences and WR, as per Experiments 1 and 2, between the subjective prior duration ($r = .52$, $p = .127$), size ($r \approx 0$, $p \approx 1$), or the combination of both size and duration ($r = -.04$, $p = .857$) showed no correlations in this task (see Fig. 9, top right).

Prior objective stimulus There was a main effect of whether participants judged the duration or the size of the stimulus

(size: .076, time: .155), $F(1, 18) = 52.84$, $p < .001$, $\eta_p^2 = .75$. Neither the effect of anchor, $F(2, 36) = .374$, $p = .690$, $\eta_p^2 = .02$, nor the interaction between anchor and judgement criteria reached significance, $F(2, 36) = .072$, $p = .930$, $\eta_p^2 = .00$.
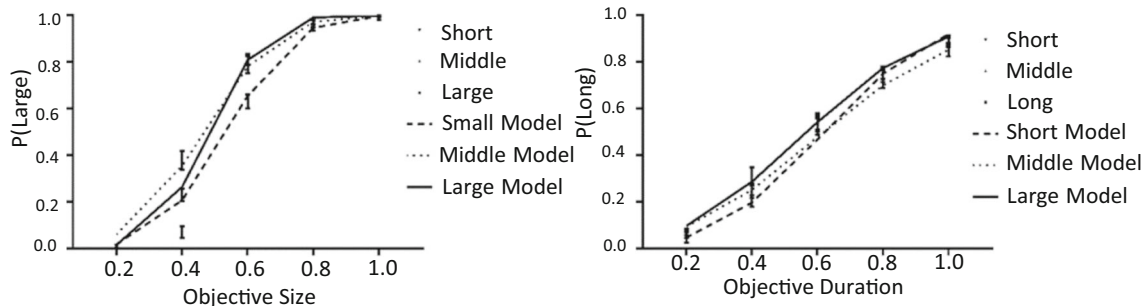
Model fitting We fitted the same model used in Experiments 1 and 2 to the data presented here. All model fits were good (MAD all below 0.02). For size judgments, if the prior choice was 'large', then $c = .21$ and $d = .86$, whereas if the prior choice was 'small', then $c = .16$ and $d = .96$. These values of d indicate an assimilative carryover effect. For time judgments, if the prior choice was 'long', then $c = .33$ and $d = .91$, whereas if the prior choice was 'short', then $c = .35$ and $d = 1.06$. Again, these values of d indicate an assimilative carryover effect, and as per the comparisons of the BP shown above, the assimilative carryover effect appears larger when judging duration. Further, the values of c indicate that judgments of duration are less sensitive to change than judgments of size.

When considering the objective prior size, a large prior trial had a $c = .18$ and $d = .88$, a medium prior trial had a $c = .24$ and a $d = .86$, and a small prior trial had a $c = .21$ and a $d = .93$. These results did not show a consistent pattern across prior stimulus size. Considering prior stimulus duration, a long prior trial gave a $c = .38$ and a $d = .94$, a medium prior trial gave a $c = .41$ and a $d = .98$, and a short prior trial gave a $c = .31$ and a
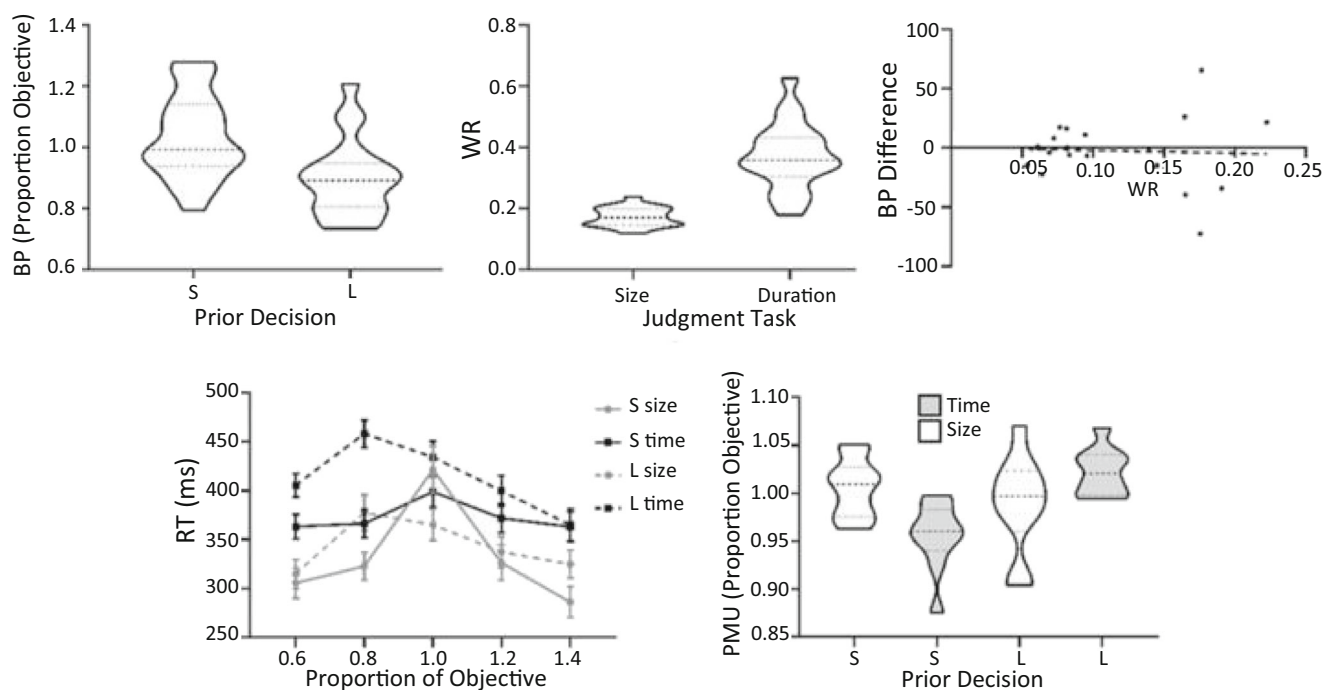

Fig. 8 Left: Probability of choosing 'large' following either a small, middle or large prior objective size. Right: Probability of choosing 'long' following a short, middle, or long objective prior stimulus

duration. Lines represent the fit of the models used. Error bars represent one standard error of the mean

Fig. 9 Top left: BP given whether the prior trial acted as an 'S' or 'L' anchor. Top middle: WR depending on judgement task. Top right: Correlation between WR and BP difference, across both conditions. Bottom left: Chronometric function given each prior decision under each judgment condition. Bottom right: PMU given whether the prior choice was an 'S' or an 'L' and whether participants were judging size or duration. All error bars represent one standard error of the mean. Dashed line in violin plots represent median, and dotted lines represent quartiles

$d = 1.02$. These model fits did show a pattern of a larger bias following an objectively longer prior stimulus. However, it is worth noting that this pattern is in the opposite direction to previous findings regarding the prior objective stimulus duration (Wiener et al., 2014) and further that this effect was not significant in terms of BP. As per the models fitted to the prior decision data, the c parameters show more sensitivity to the judgment of size rather than duration.

While the models fit to each experiment are interesting in their own right, it is also informative to compare the results of each model across each experiment. Therefore, Table 1 in the appendix concatenates this information for reference.

PMU: Prior response There was no main effect of whether people responded 'S' or 'L' in the prior trial, $F(1, 18) = 1.98$, $p = .177$, $\eta_p^2 = .09$, or of the task they were performing, $F(1, 18) = 1.98$, $p = .116$, $\eta_p^2 = .13$. However, the interaction between these terms was significant, $F(1, 18) = 7.44$, $p = .014$, $\eta_p^2 = .29$ (see Fig. 9, bottom). Post hoc analysis of the interaction term showed that the prior response significantly affected the PMU if the participant was judging duration, $t(15.7) = 4.94$, $p < .001$, $d = 2.21$, Holm-corrected. After an 'S' response, the PMU was earlier (.955) compared with if the response was 'L' (1.022). However, if the participant was attending to the size dimension, there was no effect of the subjective size of the prior trial, $t(14.6) = .873$, $p = .397$, $d = .39$, Holm-corrected. Note that in terms of duration judgments, the

PMU pattern was similar to that of Experiments 1 and 2, while the judgment of size showed no effect. This may be added support to a duration-based modulation of the PMU, rather than a decision-based effect.[7]

Prior objective stimulus Again there were no main effects of the judgment task they performed ($F(1, 18) = .025$, $p = .623$, $\eta_p^2 = .01$). The main effect of prior objective duration approached significance ($F(2, 36) = 2.76$, $p = .077$, $\eta_p^2 = .13$) as did the interaction effect ($F(2, 36) = 2.53$, $p = .094$, $\eta_p^2 = .12$).

Discussion

Review of findings

In Experiment 1, the main finding of interest was that the bisection point (BP) was significantly higher if the participant reported that the prior stimulus duration was 'short' rather

---

[7] Note that in this experiment, unlike in Experiments 1 and 2, a question mark was added to prompt response. This was done because it was found when judging size participants tended to respond prior to the conclusion of the stimulus. Generally, the RT findings in the current experiment replicate the results of Experiments 1 and 2 when participants were judging duration, indicating, perhaps, the question mark did not significantly affect PMU in this case. RTs to a size judgment (mean = 339 ms) were faster than to a time judgment (mean = 393 ms), but this difference was not significant, $t(14.0) = 1.42$, $p = .177$, $d = .64$.

than 'long'. This was supported by the SET model which was fit to the data, indicating a decisional carryover effect, as was expected given the findings by Brown et al. (2005) and Wiener et al. (2014). However, while Wiener et al. (2014) found a positive correlation between the degree of uncertainty participants showed and the size of the decisional carryover, this finding was not replicated here. Rather, it appeared that decisional carryover occurred irrespective of the WR.

In Experiment 2, when every other trial was of either the short or long reference duration, there were no indications of a decisional carryover effect from the prior reference trial to the current test trial. Bayes factor analysis and SET modelling both further supported the conclusion that there was no carryover effect. Further, in opposition to Experiment 1, there was a positive correlation between the WR differences and BP differences; participants who showed stronger decisional carryover also tended to be more uncertain of response overall. Interestingly, there *was* an assimilative carryover effect of the prior test judgment on the judgment of the subsequent reference duration. Notably, the only difference between Experiment 1 and 2 was whether every second trial was a reference duration or another test duration. In both experiments, there was no effect of objective prior stimulus duration on subsequent judgments.

In Experiment 3, we replicated the decisional carryover effect, shown by a larger BP following an 'S' (i.e., 'short'/'small') judgment of the prior test stimulus, and by a higher value of the d parameter following a 'small'/'short' prior decision. Decisional carryover was present whether judging the duration or size of the test stimulus. The difference in the carryover effect when participants were judging either size or duration approached significance in terms of the BP ($p = .050$). Further analysis of the carryover effect showed that the effect size was larger when participants judged size rather than duration, though the difference in BP was larger when participants judged duration, as were the difference in the values of d. This finding largely confirms and extends the findings by Brown et al. (2005) by demonstrating decisional carryover given a single task in which two possible stimulus dimensions could be judged, rather than in separate tasks with varying stimuli. Further, while Brown et al. (2005) used auditory stimuli, visual stimuli were used here, a factor which has previously been shown to affect timing (e.g., Droit-Volet, Tourret, & Wearden, 2004).

We also found that participants were more easily able to identify changes in size rather than duration, shown by a lower Weber ratio, BPs closer to the objective mean when size judgments were required, and higher c parameter values given time judgments. This indicates that the judgment of duration is indeed more difficult than the judgment of size, despite the ratio of the 'L' stimulus to the 'S' stimulus being the same under both judgment criteria (Droit-Volet, 2010; Droit-Volet, Clément, & Fayol, 2008; Ogden et al., 2018). Again, as per

Experiments 1 and 2, we did not find a contrast effect of the prior trial objective dimensions. Rather, if anything, there was a tendency towards an assimilation effect, as shown by the values of d in the models fit to time judgments. The reason for this is likely due to the correlation between how people respond and the objective stimulus: If presented with a large stimulus, the participant is more likely to respond 'large', and that response in turn results in an assimilative carryover effect. Though Wiener et al. (2014) found a contrast effect with the objective prior stimulus duration in the auditory domain, and other studies have found contrast effects in terms of the prior objective stimulus (e.g., Jones et al., 2006), here, it could be the case that the contrast effects of the objective prior stimulus were not large enough to overcome the decision effects of the prior response. This suggestion is in line with the modelling of Wiener et al. (2014): Objective contrast does occur in the visual domain, but is not strong enough to overcome the decision bias.

In all the experiments, as long as participants were judging duration, the PMU tended to be longer following a 'long' prior trial compared with a 'short' prior trial. In Experiment 3, we showed that there was a significant effect of prior objective and subjective duration on the PMU, but no significant effect of the objective or subjective prior *size* on the speed of the current size judgment. In prior research, the PMU has been shown to correspond to the BP (e.g., Birngruber et al., 2014; Birngruber, Schröter, & Ulrich, 2015; Wehrman, et al. 2018b); however, in the present research, we showed an assimilative BP effect and a contrastive PMU.[8] Wiener et al. (2014) found a similar pattern of results. While the PMU may be an interesting metric in categorizing perceptual decision difficulty, the lack of a PMU effect in the visual modality in Experiment 3 questions its relevance for making such claims, at least in terms of decisional carryover effects. Instead, we found that the PMU pattern found could be explained in terms of classical RT experiments (see Los, 2010, 2013). This was initially identified in an exploratory RT analysis in Experiment 1, and confirmed by repeating this analysis in Experiments 2 and 3. This RT pattern is an interesting finding in its own right: How long you have to wait to be able to respond in a bisection task affects how quickly you respond (i.e., there may be some form of 'surprise' when the duration finishes), and could be informative for those using RT measures in future. However, further research is required to systematically investigate such effects.

Overall, our findings make three points about the decisional carryover effect in the judgment of duration. Largely, these findings are supportive of a general, judgment-based effect of prior responses on perceived duration. Further, it appears that

---

[8] This is likely due to prior research examining PMU in relation to what was happening in the current trial, rather than looking at the duration of the current and prior trial, as done here.

the judgment of duration and the judgment of other stimuli are similar processes.

Decisional carryover In all three experiments, we found some form of decisional carryover effect. In Experiment 1, there was a continuous carryover from the judgment of one test duration to the next. In Experiment 2, this occurred from test duration judgments to the judgment of reference durations, and not vice versa. In Experiment 3, we replicated the finding of Experiment 1, and found that a similar effect was present in the judgment of size. Our findings thus indicate that decisional carryover is a general phenomenon. Few researchers have investigated carryover effects in duration judgments. However, Brown et al. (2005), Wiener et al. (2014) and Wehrman et al. (2018a), as well as the current study, support decisional carryover effects in duration judgments. Further, others have described similar carryover effects in terms of numerosity (Cicchini, Anobile, & Burr, 2014; Fornaciai & Park, 2018) and orientation judgments (Fischer & Whitney, 2014). In the current research, specifically in Experiment 3, we additionally examined the carryover effect from either size or duration in a single experiment.

As mentioned in the introduction, size can be ascertained instantaneously while duration is judged based on the input from other senses. It is possible that the actual precept of judged duration or size are affected by previous exposure however it seems more likely that it is simply the judgment of the stimulus that is affected. It could be argued that this bias is purely motoric in nature; if you just pressed 'left' you press 'left' again. Indeed, throughout the experiments we did not counterbalance responses per participant. The reason for using a consistent mapping is because time is often spatially localized such that the left is associated with earlier times or shorter durations (see, for example, Bonato, Zorzi, & Umiltà, 2012, for review). Using the reverse mapping (i.e. left as long) may have resulted in a Simon-type of effect which we wished to avoid.

However, it is worth noting that some research suggests a response *alternation* bias in motor responses, such that participants who responded 'left' in the prior trial may be more likely to choose 'right' in the current trial (e.g., Pape & Siegel, 2016). Furthermore, response repetition bias has also been shown when motor bias is accounted for (Akaishi, Umeda, Nagase, & Sakai, 2014). Here, though a motoric response bias is an important consideration in future research, perhaps the effect of decisional carryover is actually stronger than estimated due to a motoric alternation effect.

_____

[0] This is likely due to prior research examining PMU in relation to what was happening in the current trial, rather than looking at the duration of the current and prior trial, as done here.

Decisional carryover and response difficulty Experiment 1 largely replicated the findings by Brown et al. (2005) and Wiener et al. (2014). However, in Experiment 2, we did not find a decisional carryover effect from the judgment of reference duration to the judgment of test duration. In fact, we found moderate evidence *against* such an effect existing. Further, we found a decisional carryover effect from the judgment of the test duration to the judgment of the reference duration. In Experiment 3, the L:S ratio was the same whether participants were judging duration or size, and indeed the assimilative carryover effect only approached significance in differentiating between whether participants judged size or duration. Despite this, the WR and c parameter provided evidence that participants found the judgment of size easier than the judgment of duration, supporting prior research (Droit-Volet, 2010; Droit-Volet et al., 2008; Ogden et al., 2018).

Together, these findings indicate the need for some changes to the theories by Wiener et al. (2014) and Wehrman et al. (2018a): It did not appear that the difficulty of the current judgment was the sole determiner of whether decisional carryover would occur. Difficulty-based decisional carryover effects have been shown more widely in the serial dependency literature (e.g., Fornaciai & Park, 2018; Larimer, 1965). Thus, particularly the carryover effect found in Experiment 2 is interesting. It is possible either that the judgment of duration is removed enough from physical features that decisional carryover can occur on any given trial with some chance, or, perhaps, that duration is ambiguous enough that decisional carryover can occur in even the easiest of trials. This second point can be supported by Experiment 3, where the extreme *durations* appeared to still show some decisional carryover, while the extreme *sizes* did not.

Another possible explanation for this pattern of effects is provided by Akaishi, et al. (2014), who found that decisions based on less sensory evidence are more prone to repetition. Here, while the sensory evidence could be quantified as the same despite response certainty, perhaps the ease of decision-making in the prior trial has a similar effect. In Akaishi et al.'s (2014) terms, a more difficult decision results in a strong update of the response bias towards repetition. This leads to a carryover effect in Experiment 2 when going from the judgment of a test duration to a reference judgment (which in time perception is still somewhat ambiguous), but not vice versa. Perhaps, if the current stimulus is adequately difficult to judge, *and* the prior stimulus was adequately difficult to judge, then decisional carryover occurs. Further supporting this line of logic, the values of c across all experiments were highest given an objectively middle duration and/or size, indicating less sensitivity to change in the following trial. If this is the case, we expect stronger carryover effects following an objectively middle duration stimulus. Future research could test this by providing central durations on every other trial and examining whether the effects of the decision regarding those durations carries over to the next trial.

Alternatively, Stewart, Brown, and Chater (2005) presented the relative judgment model (RJM). Though originally theorized in relation to absolute judgements, a similar process could occur in bisection-type tasks. This model suggests that judgment is performed on the relative characteristics of the current stimulus in relation to previous experience, affecting specifically the judgment of a stimulus and not its perception. The judgment-based locus of effect is supported here; there appears to be a common, judgment-driven effect on reported durations and stimulus sizes in the current experiments. Further, the RJM model proposes that because we lack a stable long-term measure of the absolute magnitudes presented in the past, we rely instead on the relative difference between the current and prior stimulus. This leads to contrasting perceptual carryover effects, as found in Wiener et al. (2014) in the auditory domain, though this was not an effect found here. This also leads to decisional carryover effects: If a participant previously judged a stimulus as 'short', and the next stimulus was either shorter or near the same duration as that prior stimulus, then it will also lead to a 'short' decision.[9] This could explain the effects of Experiment 2: When a participant judged a short reference duration, the next stimulus would always be relatively longer, leading to a lack of a carryover effect (or if present, it would lead to a higher probability of judging a shorter objective probe duration as 'short' again, which seems to be the case). Further, decisions carried over in Experiment 2 from difficult (probe) to easy (reference) judgments. This could be due to, for example, some shorter probe durations being judged as 'long', and then in the next trial, because the reference duration was relatively close to that probe duration, decisional carryover occurred. Despite this apparent match between the RJM model predictions and our findings, the RJM model has encountered difficulties in accounting for absolute identification and matching of visual stimuli (see Guest, Adelman, & Kent, 2016). Further research, perhaps combining the roles of short-term influences along with longer-term exemplars, is required in this regard.

The question arises why Wehrman et al. (2018a) did find an assimilative carryover effect from a reference duration when this duration was *not* judged. One possibility is that the strength of the carryover is inversely proportional to the duration from one stimulus to the next. In the current experiment, perhaps the addition of a judgment resulted in too long a duration between stimuli and thus a reduction in the carryover effect. This is similar to the fading bias proposed by Wiener et al. (2014). For stimuli at the extremes (i.e., the references)

from which decisional carryover is already weak, this increase in intertrial time may be enough to eliminate the effect. Alternatively, perhaps in Wehrman et al. (2018a), when there is no judgment between the reference and test stimuli, the stimuli are considered in conjunction with one another. Participants are then judging the *combined* duration of the stimuli in relation to a midpoint established by the reference durations When a short reference is presented, the average duration of the stimuli will be shorter, while when the long reference is presented the average duration will be longer. This could result in a general bias in response.

Decisional carryover is robust Finally, the current experiments demonstrate that the decisional carryover effect in time perception is robust to various modifications from the tasks previously used. Firstly, response assimilation still occurs when using a standard bisection task and uncontrolled trial-to-trial contingency in the visual modality (see Brown et al., 2005). Further, varying the physical size of the stimulus did not stop the carryover effect in the temporal modality. The size of the stimulus (Alards-Tomalin, Leboe-McGowan, Shaw, & Leboe-McGowan, 2014; Rammsayer & Verner, 2014, 2016), as well as the spatial localization of a stimulus (Johnston, Arnold, & Nishida, 2006), which size could affect, have been shown to affect perceived duration. Thus, it is possible that the various dimensions of a stimulus could ameliorate the decisional carryover effect in the temporal domain. However, the current study demonstrates that the decisional carryover effect occurs despite differences in the stimulus dimensions. Further, in both Wehrman et al. (2018a) and Wiener et al. (2014), each stimulus used to represent duration was identical. Repetition of stimuli results in a reduced perceived duration of the second stimulus (Birngruber et al., 2014; Fromboluti, Jones, & McAuley, 2013; Schindel, Rowlands, & Arnold, 2011; Wehrman et al., 2018b).[10] Again, decisional carryover is robust to such an effect. Together, these features seem to indicate decisional carryover is robust to various low-level features of a stimulus (e.g., size, visual field, repetition). This adds some credence to the idea that temporal assimilation is a poststimulus, judgment-based effect.

Conclusion

In this article we replicated the decisional carryover effect of a prior response in Brown et al. (2005) and Wiener et al. (2014), and of an unjudged reference duration in Wehrman et al. (2018a). Further, we found that a carryover effect can still be found when varying the size of the target and reference stimulus. Generally, these effects are similar to those found

---

[9] This suggestion has interesting ramifications for examining RT data in future: A comprehensive model could examine the effects of how close a prior and current stimulus are in objective terms, and whether a decision was repeated or not. It would seem likely that RTs should indicate faster decisions in the current trial if a decision was repeated and the current and prior trial were closer rather than further apart.

[10] Though size and repetition are of interest more generally in time perception research, the effects of these on sequential processing is not the primary concern of the current article.

previously. However, the key difference here was that a response to an ambiguous stimulus did not seem to be required for decisional carryover to occur. Instead, our findings are supportive of general decisional carryover effects across tasks; perhaps any response may be assimilated if the prior trial was adequate to elicit such a carryover effect and there is at least a modicum of ambiguity. In additional findings, we showed that the RT pattern of duration judgments found here is in line with a variable foreperiod interpretation, and that decisional carryover is a robust feature of time perception judgments when considering other concurrent factors.

Compliance with ethical standards

Appendix

Table 1 Comparison of the model parameters fit in each experiment, condition and prior subjective/objective duration (or in Experiment 3, size)

Experiment	Type	Prior	Sensitivity (c)	Displacement (d)
1	Objective	Short	0.34	0.94
		Medium	0.37	0.90
		Long	0.36	0.95
	Subjective	Short	0.35	0.90
		Long	0.35	0.97
2	Objective	Short	0.34	0.94
		Long	0.35	0.94
	Subjective	Short	0.32	0.95
		Long	0.38	0.94
3	Objective size	Small	0.18	0.88
		Medium	0.24	0.86
		Large	0.21	0.93
	Subjective size	Small	0.21	0.86
		Large	0.16	0.96
	Objective duration	Short	0.38	0.94
		Medium	0.41	0.98
		Long	0.31	1.02
	Subjective duration	Short	0.33	0.91
		Long	0.35	1.06

References

Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous mechanism of internal choice estimate underlies decision intertia. *Neuron, 81*(1), 195–206. https://doi.org/10.1016/j.neuron.2013.10.018

Alards-Tomalin, D., Leboe-McGowan, J. P., Shaw, J. D., & Leboe-McGowan, L. C. (2014). The effects of numerical magnitude, size, and color saturation on perceived interval duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 555–566. https://doi.org/10.1037/a0035031

Balcı, F., & Simen, P. (2014). Decision processes in temporal discrimination. *Acta Psychologica, 149*, 157–168. https://doi.org/10.1016/j.actpsy.2014.03.005

Bausenhart, K. M., Dyjas, O., & Ulrich, R. (2014). Temporal reproductions are influenced by an internal reference: Explaining the Vierordt effect. *Acta psychologica, 147*, 60-67. https://doi.org/10.1016/j.actpsy.2013.06.011

Birngruber, T., Schröter, H., Schütt, E., & Ulrich, R. (2017). Stimulus expectation prolongs rather than shortens perceived duration: Evidence from self-generated expectations. *Journal of Experimental Psychology: Human Perception and Performance.* Advanced online publication. https://doi.org/10.1037/xhp0000433

Birngruber, T., Schröter, H., & Ulrich, R. (2014). Duration perception of visual and auditory oddball stimuli: Does judgment task modulate the temporal oddball effect? *Attention, Perception, & Psychophysics, 76*(3), 814–828. https://doi.org/10.3758/s13414-013-0602-2.

Birngruber, T., Schröter, H., & Ulrich, R. (2015). The influence of stimulus repetition on duration judgments with simple stimuli. *Frontiers in Psychology, 6*, 1213. https://doi.org/10.3389/fpsyg.2015.01213

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience & Biobehavioral Reviews, 36*(10). 2257–2273. https://doi.org/10.1016/j.neubiorev.2012.08.007

Bravo, L., & Mayzner, M. (1961). Assimilation and contrast effects of anchoring stimuli on judgments: A partial replication of the Sherif, Taub, and Hovland study. *The Journal of Psychology, 52*(2), 333–334. https://doi.org/10.1080/00223980.1961.9916533

Brown, G. D., McCormack, T., Smith, M., & Stewart, N. (2005). Identification and bisection of temporal durations and tone frequencies: Common models for temporal and nontemporal stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 919–938. https://doi.org/10.1037/0096-1523.31.5.919

Cacioppo, J. T., & Dorfman, D. D. (1987). Waveform moment analysis in psychophysiological research. *Psychological Bulletin, 102*(3), 421. https://doi.org/10.1037/0033-2909.102.3.421

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making, 7*(4), 223–242. https://doi.org/10.1002/bdm.3960070402

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 3*(3), 216–228. https://doi.org/10.1037/0097-7403.3.3.216

Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences, 111*(21), 7867–7872. https://doi.org/10.1073/pnas.1402785111

Dyjas, O., & Ulrich, R. (2014). Effects of stimulus order on discrimination processes in comparative and equality judgements: Data and models. *The Quarterly Journal of Experimental Psychology, 67*(6), 1121–1150. https://doi.org/10.1080/17470218.2013.847968

Droit-Volet, S. (2010). Speeding up a master clock common to time, number and length? *Behavioural Processes, 85*(2), 126–134. https://doi.org/10.1016/j.beproc.2010.06.017

Droit-Volet, S., Clément, A., & Fayol, M. (2008). Time, number and length: Similarities and differences in discrimination in adults and children. The Quarterly Journal of Experimental Psychology, 61(12), 1827-1846. https://doi.org/10.1080/17470210701743643

Droit-Volet, S., Tourret, S., & Wearden, J. (2004). Perception of the duration of auditory and visual stimuli in children and adults. *The Quarterly Journal of Experimental Psychology, Section A, 57*(5), 797–818. https://doi.org/10.1080/02724980343000495

Droit-Volet, S., & Wearden, J. H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology, 80*(2), 142–159. https://doi.org/10.1006/jecp.2001.2631

Droit-Volet, S., Wearden, J. H., & Zélanti, P. S. (2015). Cognitive abilities required in time judgment depending on the temporal tasks used: A comparison of children and adults. *The Quarterly Journal of Experimental Psychology, 68*(11), 2216–2242. https://doi.org/10.1080/17470218.2015.1012087

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience, 17*, 738. https://doi.org/10.1038/nn.3689

Fornaciai, M., & Park, J. (2018). Attractive serial dependence in the absence of an explicit task. *Psychological Science, 29*(3), 437–446. https://doi.org/10.1177/0956797617737385

Fromboluti, E. K., Jones, K. B., & McAuley, J. D. (2013). Temporal preparation contributes to the overestimation of duration of 'oddball' events. *Frontiers in Human Neuroscience Conference Absract: 14th Rhythm Production and Perception Workshop Birmingham.* https://doi.org/10.3389/conf.fnhum.2013.214.00013

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*(1), 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Science, 423*(1), 52–77.

Guest, D., Adelman, J. S., & Kent, C. (2016). Relative judgement is relatively difficult: Evidence against the role of relative judgement in absolute identification. *Psychonomic Bulletin & Review, 23*(3), 922–931. https://doi.org/10.3758/s13423-015-0940-2

Hampton, J. A., Estes, Z., & Simmons, C. L. (2005). Comparison and contrast in perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1459-1476. https://doi.org/10.1037/0278-7393.31.6.1459

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

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

Johnston, A., Arnold, D. H., & Nishida, S. (2006). Spatially localized distortions of event time. *Current Biology, 16*(5), 472–479. https://doi.org/10.1016/j.cub.2006.01.032

Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: Separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 316–332. https://doi.org/10.1037/0278-7393.32.3.316

Larimer, G. S. (1965). Ambiguity and nearness of anchors as factors in assimilation. *The American Journal of Psychology, 78*(3), 414–422.

Lawrence, M. A. (2013). Easy analysis and visualization of factorial experiments (Version 4.2-2) [Computer software package]. Retrieved from https://rdrr.io/cran/ez/

Lieder, F., Griffiths, T. L., Huys, Q. J., & Goodman, N. D. (2017). The anchoring bias reflects rational use of cognitive resources.

Psychonomic Bulletin & Review. 1–28. Advance online publication. https://doi.org/10.3758/s13423-017-1286-8

Los, S. A. (2010). Foreperiod and the sequential effect: Theory and data. In A. C. Nobre & J. T. Coull (Eds.), *Attention and time* (pp. 289–302). Oxford, England: Oxford University Press.

Los, S. A. (2013). The role of response inhibition in temporal preparation: Evidence from a go/no-go task. *Cognition, 129*(2), 328–344. https://doi.org/10.1016/j.cognition.2013.07.013

Meck, W. H., Church, R. M., & Olton, D. S. (1984). Hippocampus, time, and memory. *Behavioral Neuroscience, 98*(1), 3–22.

Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). Package 'BayesFactor': Computation of Bayes Factors for Common Designs (Version 0.9.12-2) [Computer software package]. Retrieved from https://cran.r-project.org/

Mori, S. (1989). A limited-capacity response process in absolute identification. *Perception & Psychophysics, 46*, 167–173. https://doi.org/10.3758/bf03204977

Muir, D. D., & Hunter, E. A. (1991). Sensory evaluation of cheddar cheese: Order of tasting and carryover effects. *Food Quality and Preference, 3*(3), 141–145. https://doi.org/10.1016/0950-3293(91)90050-O

Ogden, R. S., Samuels, M., Simmons, F., Wearden, J., & Montgomery, C. (2018). The differential recruitment of short-term memory and executive functions during time, number, and length perception: An individual differences approach. *The Quarterly Journal of Experimental Psychology*, 1–14. Advance online publication. https://doi.org/10.1080/17470218.2016.1271445

Pape, A., & Siegel, M. (2016). Motor cortex activity predicts response alternation during sensorimotor decisions. *Nature Communications, 7*(1). https://doi.org/10.1038/ncomms13098

Parducci, A., & Marshall, L. M. (1962). Assimilation vs. contrast in the anchoring of perceptual judgments of weight. *Journal of Experimental Psychology, 63*(5), 426–437. https://doi.org/10.1037/h0048727

R-Core-Team. (2015). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rammsayer, T. H., & Verner, M. (2014). The effect of nontemporal stimulus size on perceived duration as assessed by the method of reproduction. *Journal of Vision, 14*(5), 17-17. https://doi.org/10.1167/14.5.17

Rammsayer, T. H., & Verner, M. (2016). Evidence for different processes involved in the effects of nontemporal stimulus size and numerical digit value on duration judgments. *Journal of Vision, 16*(7), 13. https://doi.org/10.1167/16.7.13

Schindel, R., Rowlands, J., & Arnold, D. H. (2011). The oddball effect: Perceived duration and predictive coding. *Journal of Vision, 11*(2), 17–17.

Schütt, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research, 122*, 105–123. https://doi.org/10.1016/j.visres.2016.02.002

Sherif, M., Taub, D., & Hovland, C. I. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *Journal of Experimental Psychology, 55*(2), 150–155.

Simen, P., Balci, F., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *Journal of Neuroscience, 31*(25), 9238–9253. https://doi.org/10.1523/JNEUROSCI.3121-10.2011

Stewart, N., Brown, G. D., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review, 112*(4), 881–911. https://doi.org/10.1037/0033-295X.112.4.881

Ward, L. M., & Lockhead, G. R. (1971). Response system processes in absolute judgment. *Perception & Psychophysics, 9*(1-B), 73–78. https://doi.org/10.3758/BF03213031

Wearden, J. (1991). Human performance on an analogue of an interval bisection task. *The Quarterly Journal of Experimental Psychology, 43*(1), 59–81.

Wearden, J. (2016). *The psychology of time perception*. London, England: Palgrave Macmillan UK.

Wearden, J., & Ferrara, A. (1995). Stimulus spacing effects in temporal bisection by humans. *The Quarterly Journal of Experimental Psychology Section B, 48*(4), 289–310. https://doi.org/10.1080/14640749508401454

Wearden, J., & Ferrara, A. (1996). Stimulus range effects in temporal bisection by humans. *The Quarterly Journal of Experimental Psychology: Section B, 49*(1), 24–44. https://doi.org/10.1080/713932615

Wearden, J., & Jones, L. A. (2013). Explaining between-group differences in performance on timing tasks. *The Quarterly Journal of Experimental Psychology, 66*(1), 179–199. https://doi.org/10.1080/17470218.2012.704928

Wehrman, J. J., Wearden, J., & Sowman, P. (2018a). The expected oddball: Effects of implicit and explicit positional expectation on duration perception. *Psychological Research*, 1–15. Advance online publication. https://doi.org/10.1007/s00426-018-1093-5

Wehrman, J. J., Wearden, J., & Sowman, P. (2018b). Short-term effects on temporal judgement: Sequential drivers of interval bisection and reproduction. *Acta psychologica, 185*, 87-95. https://doi.org/10.1016/j.actpsy.2018.01.009

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293–1313.

Wiener, M., Thompson, J. C., & Coslett, H. B. (2014). Continuous carryover of temporal context dissociates response bias from perceptual influence for duration. *PLOS ONE, 9*(6), e100803. https://doi.org/10.1371/journal.pone.0100803