The effect of nontemporal stimulus size on perceived duration as assessed by the method of reproduction

Thomas H. Rammsayer
Department of Psychology, University of Bern, Bern, Switzerland
Center for Cognition, Learning, and Memory, University of Bern, Bern, Switzerland

Martin Verner
Department of Psychology, University of Bern, Bern, Switzerland

Perceived duration is assumed to be positively related to nontemporal stimulus magnitude. Most recently, the finding that larger stimuli are perceived to last longer has been challenged to represent a mere decisional bias induced by the use of comparative duration judgments. Therefore, in the present study, the method of temporal reproduction was applied as a psychophysical procedure to quantify perceived duration. Another major goal was to investigate the influence of attention on the effect of visual stimulus size on perceived duration. For this purpose, an additional dual-task paradigm was employed. Our results not only converged with previous findings in demonstrating a functional positive relationship between nontemporal stimulus size and perceived duration, but also showed that the effect of stimulus size on perceived duration was not confined to comparative duration judgments. Furthermore, the effect of stimulus size proved to be independent of attentional resources allocated to stimulus size; nontemporal visual stimulus information does not need to be processed intentionally to influence perceived duration. Finally, the effect of nontemporal stimulus size on perceived duration was effectively modulated by the duration of the target intervals, suggesting a hitherto largely unrecognized role of temporal context for the effect of nontemporal stimulus size to become evident.

Introduction

Numerous studies reported that perceived duration is positively related to various aspects of nontemporal stimulus magnitude. For example, longer judged duration as a function of increasing stimulus magnitude has been shown for nontemporal stimulus attributes such as brightness (e.g., Brigner, 1986; Long & Beaton, 1980; Xuan, Zhang, He, & Chen, 2007), numerosity (e.g., Oliveri et al., 2008; Vicario, 2011; Xuan et al., 2007), stimulus complexity (e.g., Ornstein, 1969; Schiffman & Bobko, 1974), or stimulus size (e.g., Ono & Kawahara, 2007; Xuan et al., 2007). It still remains unclear, however, whether these effects of nontemporal stimulus magnitude on judged duration are mediated by different processes associated with distinct elementary time experiences (cf., Fraisse, 1978, 1984; Grondin, 2010), by a common cognitive mechanism as, for example, coding efficiency (Eagleman & Pariyadath, 2009), or by a generalized magnitude system (Walsh, 2003). Recently, the question has been raised of whether nontemporal stimulus magnitude actually affects perceived duration of a stimulus or simply biases decisions about duration (Yates, Loetscher, & Nicholls, 2012). Against this background, the focus of the present study was on the effect of stimulus size on perceived duration—a phenomenon relatively unattended by past and present research on the experience of time.

The first scientific account of the effect of stimulus size on perceived duration can be ascribed to Mo and Michalski (1972). When presenting two circles, one smaller than the other, for the same duration of either 450 or 510 ms, their participants consistently judged the larger circle to be presented longer than the smaller one. In a series of experiments, Thomas and Cantor (1975, 1976) also demonstrated that large visual stimuli presented for the same duration as small stimuli appeared to have been presented longer. In their experiments, filled circles with diameters of 8.33 mm and 10.32 mm, for small and large stimuli, respectively, were presented for either 30 or 70 ms. Participants were

Citation: 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, 1–10, http://www.journalofvision.org/content/14/5/17, doi:10.1167/14.5.17.

doi: 10.1167/14.5.17 Received January 11, 2014; published May 30, 2014 ISSN 1534-7362

required to categorize the duration of each presentation of a circle as "short," "medium," or "long." Although there was a significant main effect of stimulus size, a significant interaction between target duration and stimulus size indicated that the increase in perceived duration was larger for the long than for the short target duration. These findings were confirmed by subsequent studies employing basically the same methodology but display times of either 20 and 50 ms (Cantor & Thomas, 1976) or 40 and 70 ms (Long & Beaton, 1980).

More recently, Xuan et al. (2007) investigated whether judgments of duration in the range of 600 to approximately 900 ms were influenced by the size of an open square. Applying a Stroop-like interference paradigm, their participants judged the presentation time of open squares while stimulus size was systematically varied. Their data indicated that larger squares were judged temporally longer than smaller ones and, thus, caused Xuan et al. (2007) to conclude that larger stimuli are perceived to last longer. This claim, however, has been questioned by pointing out that the mere fact that larger stimuli were judged longer than smaller ones does not necessarily imply that the size of a nontemporal stimulus effectively affects perceived duration (Yates et al., 2012). In fact, it is also conceivable that nontemporal stimulus size simply biases decisions about duration, especially when using a task contingent upon a comparative judgment (cf., Anton-Erxleben, Abrams, & Carrasco, 2010; Nicholls, Lew, Loetscher, & Yates, 2011). In a first experiment employing comparative judgments (i.e., participants judged whether the first or the second of two stimuli was presented longer), Yates et al. (2012) replicated the effect of stimulus size on duration judgment reported by Xuan et al. (2007). However, when using equality judgments (i.e., participants judged whether two stimuli were presented for the same duration or for different durations) proposed to measure perceived duration less confounded by decisional bias (cf., Anton-Erxleben et al., 2010; Nicholls et al., 2011; Schneider & Komlos, 2008), larger nontemporal stimuli were judged as shorter in duration. Rather than providing converging evidence for the general notion that larger stimuli are perceived to last longer, this unexpected pattern of results emphasized that nontemporal stimulus size may differentially bias decisions about duration as a function of comparative and equality judgments, respectively (Yates et al., 2012).

Taken together, based on the available data, the effect of stimulus size on perceived duration remains unresolved. Most studies, so far, employed stimulus durations in the lower subsecond range. Therefore, it remains unclear to what extent these findings also hold for longer durations. In addition, and even more important, all these previous studies used methodological approaches that only allow for rather indirect quantification of the effect of stimulus size. In the majority of studies (Cantor & Thomas, 1976; Long & Beaton, 1980; Mo & Michalski, 1972; Thomas & Cantor, 1975, 1976) category ratings had been used to quantify the effect of stimulus size on perceived duration. With this duration scaling procedure, the experimenter presents a temporal interval and the participant locates its perceived duration in one of n predefined categories, which are ordered by temporal magnitude (Allan, 1979). More specifically, to quantify perceived duration as the dependent variable, the values of 0, 1, and 2 were given to "short," "medium," and "long" responses, respectively. For each participant, a mean response value was then obtained for each combination of stimulus duration and stimulus size. Thus, strictly speaking, category ratings provide just rank-order information with regard to perceived duration.

Also the methodological approach applied by Xuan et al. (2007) could not completely resolve these reservations. By using a Stroop-like paradigm, error rate of temporal judgments served as an indirect measure of perceived duration. Yates et al. (2012) remedied this shortcoming by increasing the range of duration pairs with different levels of task difficulty. This approach made it possible to estimate the individual psychometric function and, thus, each participant's point of subjective duration equality as a quantitative measure of the stimulus size effect. As already pointed out, however, this improved psychophysical procedure did not help to allay the concerns regarding a decisional bias as a potential source of the effect of stimulus size on judged duration.

In research on human timing, there are three major temporal judgment methods for a direct assessment of perceived duration (cf., Allan, 1979; Doob, 1971; Grondin, 2010; Zakay, 1990): (a) verbal estimation, i.e., the duration of a target interval is estimated verbally in terms of temporal units; (b) temporal production, i.e., an interval is produced equal to a duration that is verbally indicated; and (c) temporal reproduction, i.e., after presentation of a target interval, an attempt is made to reproduce one of equal duration by means of some operation. Both verbal estimation and temporal production show more intersubject variability than the method of reproduction (Block, 1989; Zakay, 1990). Furthermore, they use a translation of duration into socially learned time units (e.g., seconds or minutes) and, thus, their results depend on the relation of subjective time to clock time (Block, 1989; Clausen, 1950). By contrast, the method of reproduction appears to provide a more direct measure of the subjective experience of time (Danziger & Du Preez, 1963). Based on these considerations, the most direct and sensitive temporal judgment method for psychophysical assessment of perceived duration represents the method of reproduction (cf., Doob, 1971; Zakay, 1990): In the case of a positive effect of nontemporal stimulus size on perceived duration, larger stimulus sizes should lead to longer reproduced durations. Relatively few methodological studies have been performed to evaluate the method of reproduction. The most comprehensive review of the three major temporal judgment methods by Doob (1971) highlighted that, although the material on reliability is rather copious, the available studies yielded highly inconclusive results; depending on the specific study, reliability of the reproduction method varied from least to most reliable when compared to verbal estimation and temporal production.

Based on these considerations, the present study was designed to investigate whether the effect of nontemporal stimulus size on perceived duration can also be demonstrated with the method of temporal reproduction. This method not only enables the most direct psychophysical assessment of perceived duration, but also represents a temporal judgment method neither based on a category rating nor utilizing a comparative judgment. In order to prove reliability and accuracy of measurement (cf., Mueller & Martorell, 1988), estimates of split-half reliability, as an indicator of internal consistency (Cohen, Swerdlik, & Smith, 1992), and estimates of test-retest reliability, as an indicator of stability over time, were obtained in the present study. This latter aspect of reliability necessitates two testing sessions separated by a predefined period of time referred to as a test-retest interval.

Another question to be addressed in the present study was whether the effect of nontemporal stimulus size on perceived duration depends on the amount of attention paid to nontemporal stimulus magnitude. Allocation of attention onto a visual stimulus generally increases the salience of this stimulus (cf., Carrasco, Williams, & Yeshurun, 2002; Hopfinger & Mangun, 2001). Several studies indicated perceived size distortion due to changes in attention, although the direction of this distortion appeared to be highly ambiguous. While some studies reported attention-induced increases in perceived size (e.g., Anton-Erxleben, Henrich, & Treue, 2007; Masin, 2003, 2008), others found an opposite effect (e.g., Tsal & Shalev, 1996; Tsal, Shalev, & Zakay, 2005) or arrived at the conclusion that attention does not affect appearance of a target stimulus at all (e.g., Blaser, Sperling, & Lu, 1999; Schneider, 2006; Schneider & Komlos, 2008). Therefore, to identify a possible mediating influence of attention on the effect of stimulus size on perceived duration, a dual-task paradigm was applied in the present study. In addition to temporal reproduction as the primary task, a secondary task was added where participants were required to focus their attention on particular stimulus features. In the *salience condition*, participants were explicitly told to pay attention to stimulus size, whereas in the *control condition*, their attention was directed to stimulus shape. If attention is crucial for the effect of stimulus size to occur, no or at least a reduced effect of stimulus size should be the expected outcome for the control condition.

Method

Participants

The participants were eight male and 32 female adult volunteers ranging in age from 18 to 29 years (mean age ± *SD*: 21.8 ± 2.5 years). All participants were undergraduate psychology students and received course credit for taking part in this experiment. They were naïve about the purpose of this study and had normal or corrected-to-normal vision. The study was approved by the local ethics committee, and informed consent was obtained from each participant prior to the experiment.

Stimuli and procedure

The presentation of stimuli was controlled by E-Prime 2.0 experimental software running on a Dell Optiplex 760 Computer with a 17 in. monitor. Participants' responses were logged by means of a Cedrus RB-730 response box. Visual stimuli indicating the target intervals were either filled squares or filled circles presented in two different sizes subtending a visual angle of 1.2° and 10.0°, respectively. Reproduction intervals were indicated by a fixation cross of a constant size subtending a visual angle of 2.0°. All stimuli were presented in black on a white background.

Each participant performed two versions of the reproduction task conforming to the salience and the control condition, respectively. Order of version was balanced across participants. On each version of the task, the participant was required to reproduce three different target intervals. Durations of the target intervals were 800, 1000, and 1200 ms. Because explicit counting becomes a useful timing strategy for intervals longer than approximately 1200 ms (Grondin, Meilleur-Wells, & Lachance, 1999), the longest target interval was chosen not to exceed this critical value. There were 16 presentations of each target interval resulting in a total of 48 trials for each version of the task. The 16 presentations of each target interval consisted of four trials of each possible factorial combination of stimulus shape (circles and squares) and stimulus size (small and large). All 48 trials were presented in random order.

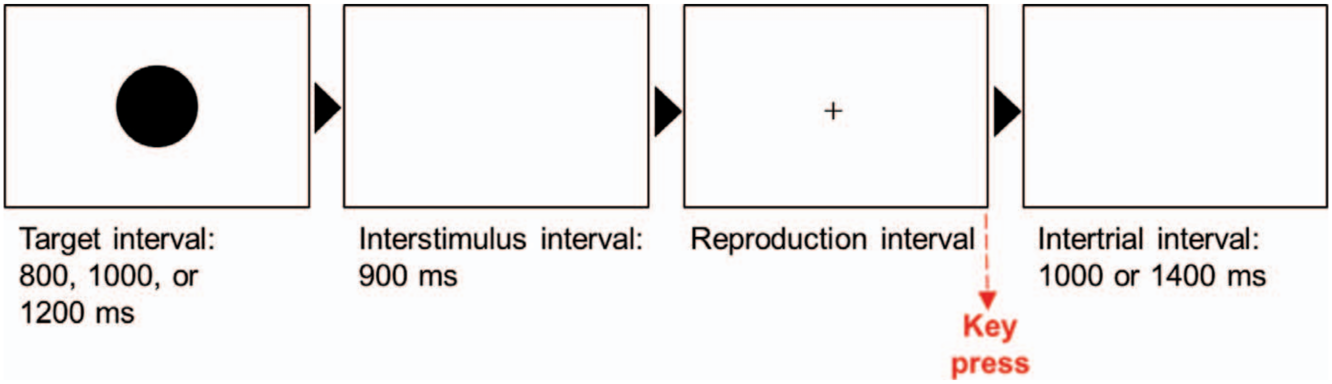

Figure 1. A sample trial of the temporal reproduction task. In the present example, the target interval consisted of a large filled circle presented for either 800, 1000, or 1200 ms. After a 900-ms interstimulus interval (blank screen), the reproduction interval marked by a fixation cross was started. The participant terminated the reproduction interval by pressing a designated response button when he/she perceived the reproduction interval as temporally identical to the immediately preceding target interval. The next trial began after an intertrial interval of either 1000 or 1400 ms.

On each trial, the target interval was followed by a blank screen for 900 ms. The start of the reproduction interval was marked by the appearance of a fixation cross. Participants were instructed to end the reproduction interval by pressing a designated response button when its duration was perceived as temporally identical to the corresponding target interval. After termination of the reproduction interval, a blank screen was presented for either 1000 or 1400 ms before the next trial was started. These two intertrial intervals were presented in a randomized manner in order to prevent a rhythmic response pattern. The time course of a complete trial is depicted in Figure 1.

In addition to the temporal reproduction task, participants were required to indicate whether the nontemporal target stimulus was either small or large (salience condition) or whether it was a circle or a square (control condition). More precisely, in the salience condition, participants had to press one of two designated response buttons in order to terminate the reproduction interval if the stimulus indicating the target interval was small and the other one if a large stimulus was displayed. In the control condition, stimulus size was irrelevant and response buttons corresponded to the geometrical shape (circle or square) of the stimulus. The assignment of response button to hand was held constant within each participant but was balanced across participants.

On each trial, the reproduced duration was logged with an accuracy of ±1 ms. As a quantitative measure of perceived duration, mean reproduced durations (MRDs) were computed for each experimental condition. The effect of stimulus size on perceived duration was defined as the difference between the MRD for the large stimulus size and the corresponding MRD for the small stimulus size. Each participant was tested in two experimental sessions separated by a test-retest interval of one week.

Assessment of reliability

As measures of reliability for the method of reproduction, test-retest and split-half reliability estimates were determined. Test-retest reliability estimates were obtained by correlating corresponding pairs of MRD values from the first and the second experimental session. Estimates of split-half reliability were obtained by splitting the trials of an experimental condition into two halves by adopting an odd-even split, in which the odd-numbered trials formed one half and the even-numbered trials formed the second half. Then, Pearson correlations between scores of the two halves were computed. Because coefficients of split-half reliability represent an estimate of internal consistency from a correlation of two halves of all trials, and because the reliability of a task is affected by its total number of trials, the Spearman-Brown formula was used for adjustment of the split-half coefficients (cf., Allen & Yen, 1979).

Results

To control for outliers, a procedure based on the one suggested by Chang, Tzeng, Hung, and Wu (2011) was applied. At first, for each participant, all reproduced durations that were more than ±2 *SDs* from that participant's mean reproduced duration for a given target interval were considered invalid trials and, thus, not included in further data analysis. By using this criterion, 4.1% of all trials were removed from data

	Target interval					
	800 ms		1000 ms		1200 ms	
	M	*SD*	*M*	*SD*	*M*	*SD*
Measurement 1						
Small stimuli	868	131	1010	140	1127	161
Large stimuli	868	149	1041	167	1187	166
Measurement 2						
Small stimuli	956	209	1153	235	1193	222
Large stimuli	1005	240	1193	222	1270	247

Table 1. Means (*M*) and standard deviations (*SD*) of reproduced durations for large and small stimuli as a function of target durations and time of measurement. Measurements 1 and 2 were separated by a one-week interval. All data in ms.

analysis. In a next step, each participant's remaining reproduced durations were submitted to a one-way analysis of variance with target intervals (800, 1000, and 1200 ms) as three levels of a repeated-measures factor. The lack of a significant main effect as well as any non-significant differences among the three factor levels would imply an individual's inability to follow the instruction to reproduce the target intervals. None of our participants, however, had to be excluded on the basis of this latter criterion.

Analysis of error rates on the two versions of the secondary task yielded faultless performance with error rates of 0.00 and, thus, indicated that all participants conformed to the instructions of the salience and control condition, respectively. There also was no indication of a statistically significant effect of stimulus shape (circles and squares) on MRD. Data were, therefore, collapsed across shapes of nontemporal stimuli. Means and *SDs* for reproduced durations as a function of target duration, stimulus size, and time of measurement are given in Table 1.

For statistical analysis, four-way within-subjects analysis of variance was performed with Target Interval (800, 1000, and 1200 ms), Stimulus Size (small and large stimuli), Stimulus Attribute Relevance (salience and control condition), and Time of Measurement (first and second testing session) as four repeated-measurement factors. To protect against violations of sphericity, Greenhouse-Geisser corrected *p* values are reported where appropriate (cf., Geisser & Greenhouse, 1958).

Four-way analysis of variance revealed statistically significant main effects of Target Interval, $F(2, 78) = 320.78$, $p < 0.001$, $\eta_p^2 = 0.892$, Stimulus Size, $F(1, 39) = 55.40$, $p < 0.001$, $\eta_p^2 = 0.587$, and Time of Measurement, $F(1, 39) = 16.50$, $p < 0.001$, $\eta_p^2 = 0.297$. The significant main effect of Target Interval indicated longer MRDs with increasing duration of the target intervals. Subsequent Scheffé post-hoc analysis revealed that MRDs of all three target intervals differed

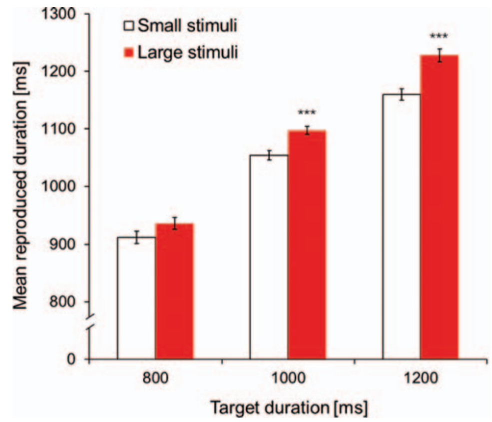

Figure 2. Reproduced duration as a function of stimulus size and target duration. The effect of stimulus size on reproduced duration was effectively moderated by the duration of the target interval. Scheffé post-hoc tests showed statistically significant longer reproduced duration with large compared to small nontemporal stimulus size for the 1000- and 1200-ms target interval. For the 800-ms target interval the effect of stimulus size on reproduced duration did not reach statistical significance. Error bars: 95% confidence interval calculated as recommended by Baguley (2012). *** $p < 0.001$.

significantly from each other ($p < 0.001$). The significant main effect of Stimulus Size on MRD clearly argued for an effect of nontemporal stimulus magnitude on perceived duration. Large target stimuli were reproduced longer than small target stimuli; MRDs were 1042 ms and 1087 ms for small and large target stimuli, respectively. Furthermore, MRDs differed significantly as a function of Time of Measurement; MRDs were reliably longer when obtained in the second testing session (1112 ms) than in the first testing session (1016 ms). No main effect of Stimulus Attribute Relevance on MRD could be established, $F(1, 39) = 0.67$, $p = 0.42$, $\eta_p^2 = 0.017$; MRDs for the salience and control condition were 1059 ms and 1070 ms, respectively.

Statistically significant two-way interactions between Stimulus Size and Target Interval, $F(2, 78) = 8.98$, $p < 0.001$, $\eta_p^2 = 0.187$, and between Stimulus Size and Time of Measurement, $F(1, 39) = 8.90$, $p < 0.01$, $\eta_p^2 = 0.186$, suggested that the effect of Stimulus Size was effectively moderated by the duration of the target interval as well as by the number of testing sessions. Post hoc analysis of the two-way interaction between Stimulus Size and Target Interval yielded a statistically significant effect of nontemporal stimulus magnitude for the 1000- and 1200-ms target interval (both $p < 0.001$) but not for the 800-ms target interval (see Figure 2). In addition, the effect of stimulus size increased from the first to the second testing session; while large target stimuli were reproduced 30 ms longer than small target stimuli in the first testing session ($p < 0.01$), this difference virtually doubled to 60 ms ($p < 0.001$) in the second testing session. Additionally, there was a statistically significant two-way interaction of Target Interval and Time of Measurement on MRD, $F(2, 78) = 3.49$, $p < 0.05$, $\eta_p^2 = 0.082$. Further analysis of this interaction revealed that all target intervals were reproduced longer in the second compared to the first testing session, but this effect tended to decrease with increasing duration of the target interval and was least pronounced, although still significant, for the longest target interval. No other interactions reached the 5% level of statistical significance.

In a final step, split-half and test-retest reliability estimates of MRD values were determined for each target interval and each stimulus size (see Table 2). Test-retest reliability estimates of the stimulus size effect are also given in Table 2. As can be seen from Table 2, estimates of split-half reliability ranging from $r = 0.85$ to $r = 0.98$ were extremely high with an average split-half coefficient of approximately $r = 0.94$. Test-retest coefficients for MRDs, ranging from $r = 0.65$ to $r = 0.73$, were smaller but still satisfactory.

	Target interval		
	800 ms	1000 ms	1200 ms
Split-half reliability			
Measurement 1			
Small stimuli	0.85	0.92	0.96
Large stimuli	0.89	0.89	0.91
Measurement 2			
Small stimuli	0.98	0.96	0.95
Large stimuli	0.97	0.97	0.97
Test-retest reliability			
Small stimuli	0.68	0.73	0.65
Large stimuli	0.65	0.69	0.72
Stimulus size effect	0.04	−0.08	0.21

Table 2. Split-half and test-retest reliability estimates for mean reproduced durations as a function of stimulus size. Estimates of split-half reliability were obtained by splitting the trials of an experimental condition into two halves by adopting an odd-even split. In a next step, Pearson correlations between scores of the two halves were computed and adjusted by means of the Spearman-Brown formula. Test-retest reliability estimates were obtained by correlating corresponding mean reproduced durations from the first (Measurement 1) and the second (Measurement 2) experimental session separated by a one-week interval. Although not meaningful (see Discussion), for the sake of completeness, also test-retest coefficients for the effect of stimulus size are reported.

Discussion

The present study was designed to investigate whether the effect of nontemporal stimulus size on perceived duration reported for comparative judgments can be generalized and proved to also hold for another temporal judgment method. Another major goal of the present study was to examine to what extent this effect depends on the amount of attention paid to stimulus size.

In a recent study, Yates et al. (2012) challenged the notion of a direct effect of nontemporal stimulus size on perceived duration. When employing comparative judgments, Yates et al. (2012) found a positive effect of stimulus size on duration judgments. However, when using equality judgments, larger stimuli were judged as shorter in duration compared to smaller stimuli. This pattern of results led them to conclude that "at least one of the two types of duration judgments is prone to some form of decisional bias" (p. 5). Because no further conclusions could be drawn on Yates et al.'s (2012) data, the method of temporal reproduction was applied in the present study.

Although never used before to examine the effect of stimulus size on perceived duration, this temporal judgment method can be considered the most direct psychophysical assessment of perceived duration and does not utilize a comparative judgment. The very high split-half reliability coefficients, obtained in the present study, indicated good instrument reliability and accuracy of the temporal reproduction task for quantification of perceived duration. In other words, our reproduction task can be considered a reliable measuring instrument that was minimally affected by random error. On the other hand, stability over time for MRDs, as indicated by test-retest coefficients, was somewhat smaller but still satisfactory. Split-half reliability coefficients can, in general, be expected to be higher than test-retest reliability estimates. This difference is because more factors contribute to measurement error and, thus, inflate error variance when test-retest reliability is determined than when split-half coefficients are computed. Possible intervening factors, such as temporal instability of the attribute being measured, current levels of fatigue or motivation, day-to-day fluctuations in mood, or effects of practice may affect test-retest reliability estimates, but have no effect on split-half estimates. Therefore, the observed moderate test-retest coefficients may not indicate that the reliability of the measuring instrument is poor, but may, instead, signify that perceived duration has been changed due to intervening factors (cf., Carmines & Zeller, 1987).

Unlike quantification of perceived duration for the large (MRD_large) and the small (MRD_small) stimulus size, quantification of the magnitude effect is based on the difference between MRD_large and MRD_small of corresponding target intervals. Such difference scores present a special problem with regard to test-retest reliability (Murphy & Davidshofer, 2005). The higher the correlation between MRD_large and MRD_small, the lower the test-retest coefficients of the resulting difference scores. This relationship is because two factors entered into this difference: first, differences in true MRD_large and MRD_small scores and, second, differences due to measurement error. In the present study, corresponding MRD_large and MRD_small values were highly correlated with each other as indicated by a mean Pearson correlation of $r = 0.92$ across all experimental conditions. Thus, MRD_large and MRD_small values shared a substantial portion of true variance. As a consequence, variability in differences between scores on MRD_large and MRD_small was due almost entirely to measurement error (cf., Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Lord & Novick, 1968). When errors of measurement were responsible for much of the variability observed in the difference scores representing the effect of stimulus size on perceived duration, no correlational relationship between difference scores beyond chance level can be expected; and, therefore, no meaningful statement on the reliability of the effect of stimulus size on perceived duration can be derived from this methodological approach.

Our findings obtained by means of the reproduction method clearly confirmed the positive effect of nontemporal stimulus size on perceived duration as reported from previous studies using target duration in the subsecond range and comparative judgments. Large stimuli presented at durations ranging from 800 to 1200 ms were reproduced approximately 4.3% longer than small stimuli presented for the same durations. It should be noted, however, that this effect of nontemporal stimulus size on reproduced duration was effectively modulated by target duration and became reliably more pronounced with increasing duration of the target interval: Although large stimuli were perceived 2.6% longer than small stimuli when presented for 800 ms, a statistically significant influence of stimulus size on reproduced duration could be established only for the 1000- and the 1200-ms target durations. At the latter target duration, the effect of stimulus size doubled, relative to the 800-ms target duration, to 5.9%. Such a moderating effect of target duration might suggest that a stimulus has to be presented longer than 800 ms for the effect of stimulus size to become effective. This interpretation, however, is challenged by earlier studies in the subsecond range. Larger effects of stimulus size on judged duration were also found when target duration was 70 ms or 50 ms rather than when target duration was 30 ms or 20 ms, respectively (Cantor & Thomas, 1976; Thomas & Cantor, 1976). A similar tendency has been reported by Long and Beaton (1980) for 70- and 40-ms target intervals.

Given that the effect of stimulus size was least pronounced for the shortest duration in a series of target durations, irrespective of whether the presented target duration was in the order of milliseconds or seconds, a range or context effect (e.g., Helson, 1948; Parducci, 1968; Poulton, 1975) could also account for the observed increase in the effect of stimulus size with increasing target duration. Psychophysical judgments are often dependent on context and range of the stimuli applied (e.g., Kowal, 1993; Marks, 1992). This dependency also holds for duration judgments (Bausenhart, Dyjas, & Ulrich, 2014; Jazayeri & Shadlen, 2010; Ryan, 2011). Although the mechanisms and neural underpinnings underlying range and context effects are still not well understood (Jazayeri & Shadlen, 2010; Ryan, 2011), the interaction between stimulus size and duration of the target intervals observed in former studies as well as in the present experiment suggests that the effect of nontemporal stimulus size on perceived duration is a function of more than just the mere stimulus input on the current trial. Converging evidence for such a conclusion comes from a study by Gomez and Robertson (1979). In their study, a substantial effect of nontemporal stimulus size on judged duration occurred only when size was varied within session, but not when held constant within a given session. Furthermore, they reported considerably reduced effects of stimulus size when the difference in duration between the target intervals was increased.

In the present study, for the first time, a repeated measurement design was applied to explore the effect of stimulus size on perceived duration. This procedure provided some new insights in the time course of both MRDs as well as the influence of nontemporal stimulus size on duration judgments across two successive measurements separated by a test-retest interval of one week. All MRDs were reliably longer when obtained in the second compared to the first testing session. This lengthening in MRD from the first to second testing session, however, tended to decrease with increasing duration of the target interval. Independent of the mutual influence of time of measurement and duration of the target interval on MRD, the effect of stimulus size across all three target intervals virtually doubled from the first to the second testing session. This latter finding arose from the fact that the lengthening effect on MRD observed in the second testing session was much more pronounced for the large than for the small target stimuli. Thus, more extensive practice and/or increased familiarity with a given set of nontemporal stimuli might represent an important, hitherto unrecognized contributing factor for larger stimuli to more effectively induce longer perceived duration.

A cautionary note refers to the finding that the stimulus size effect increased with longer target interval. This effect may be partially contributed by a longer afterimage for longer target intervals. From a purely theoretical point of view, even within conceptual frameworks proposed to account for the effect of nontemporal stimulus size on perceived duration, such as coding efficiency (Eagleman & Pariyadath, 2009) or a generalized magnitude system (Walsh, 2003), a confounding effect of afterimage cannot be definitely excluded. There is, however, converging evidence supporting the notion that afterimages do not underlie the effect of nontemporal stimulus size on perceived duration. For example, employing identical stimuli, Yates et al. (2012) showed that larger stimuli are perceived to last longer when using comparative judgments but failed to do so when using equality judgments. Because duration of stimulus presentation was the same for both comparative and equality judgments, different afterimages could not account for the observed effects of stimulus size on perceived duration. Furthermore, Ono and Kawahara (2007) measured the perceived duration of a visual object whose apparent area was altered by the Ebbinghaus illusion while its physical size remained invariant. They found that the perceived duration for apparently larger stimuli was longer than that of apparently smaller stimuli presented for the same duration. This finding is indicative of a timing process influenced by a size-contrast illusion that operates at higher levels of the visual system (Ono & Kawahara, 2007), whereas afterimages arise from more peripheral physiological factors (e.g., Craik, 1940; Zaidi, Ennis, Cao, & Lee, 2012). Nevertheless, future studies should be designed to systematically and directly examine a possible mediating influence of afterimages on the effect of nontemporal stimulus size on perceived duration.

The lack of an effect of Stimulus Attribute Relevance indicated that the effect of nontemporal stimulus size on perceived duration does not depend on the amount of attention paid to nontemporal stimulus magnitude. Furthermore, this finding is consistent with the notion that attention does not affect perceived appearance of a target stimulus (Blaser et al., 1999; Schneider, 2006; Schneider & Komlos, 2008). It also supports Xuan et al.'s (2007) assumption that stimulus magnitude need not be processed intentionally to effectively modulate perceived duration. Rather, magnitude information of a stimulus appears to be processed automatically and beyond cognitive control (cf., Dehaene & Akhavein, 1995; Tzelgov, Meyer, & Henik, 1992) but still effectively influences perceived duration.

Because in the present study a dual-task procedure was applied, one may argue that the additional nontemporal task of reporting size or shape may have influenced participants' temporal reproductions. It should be noted, however, that there was no indication of either a significant main effect of Stimulus Attribute Relevance or a statistically significant interaction including Stimulus Attribute Relevance. This pattern of results suggests that the observed effect of nontemporal stimulus size on reproduced duration can be considered to be independent of the additional nontemporal task.

Conclusions

Complementing previous studies, we showed that the effect of nontemporal stimulus size on perceived duration was not confined to comparative duration judgments. Moreover, this effect was independent of attentional resources allocated to stimulus size. Thus, nontemporal stimulus information has not to be processed intentionally to influence perceived duration. Finally, the effect of nontemporal stimulus size on perceived duration was effectively modulated by the duration of the target intervals. This latter finding suggests an important role of temporal context for the effect of stimulus size to become evident. Taken together, our results not only converge with previous findings in demonstrating a functional positive relationship between nontemporal stimulus size and perceived duration, but also cast some doubt on the notion of a decisional bias underlying this relationship.

Keywords: stimulus magnitude, perceived duration, temporal reproduction, temporal context, attention

Acknowledgments

Commercial relationships: none.
Corresponding author: Thomas H. Rammsayer.
Email: thomas.rammsayer@psy.unibe.ch.
Address: Department of Psychology, University of Bern, Bern, Switzerland.

References

Allan, L. G. (1979). The perception of time. *Perception & Psychophysics, 26,* 340–354.

Allen, M., & Yen, W. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

Anton-Erxleben, K., Abrams, J., & Carrasco, M. (2010). Evaluating comparative and equality judgments in contrast perception: Attention alters appearance. *Journal of Vision, 10*(11):6, 1–22, http://www.journalofvision.org/content/10/11/6, doi:10.1167/10.11.6. [PubMed] [Article]

Anton-Erxleben, K., Henrich, C., & Treue, S. (2007). Attention changes perceived size of moving visual patterns. *Journal of Vision, 7*(11):5, 1–9, http://journalofvision.org/7/11/5/, doi:10.1167/7.11.5. [PubMed] [Article]

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods, 44,* 158–175.

Bausenhart, K. M., Dyjas, O., & Ulrich, R. (2013). Temporal reproductions are influenced by an internal reference: Explaining the Vierordt effect. *Acta Psychologica, 147,* 60–67.

Blaser, E., Sperling, G., & Lu, Z.-L. (1999). Measuring the amplification of attention. *Proceedings of the National Academy of Sciences, USA, 96,* 11681–11686.

Block, R. A. (1989). Experiencing and remembering time: Affordances, context, and cognition. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life span perspective* (pp. 333–363). Amsterdam: North-Holland.

Brigner, W. L. (1986). Effects of perceived brightness on perceived time. *Perceptual and Motor Skills, 63,* 427–430.

Cantor, N. E., & Thomas, E. A. C. (1976). Visual masking effects on duration, size, and form discrimination. *Perception & Psychophysics, 19,* 321–327.

Carmines, E. G., & Zeller, R. A. (1987). *Reliability and validity assessment*. Beverley Hills, CA: Sage publications.

Carrasco, M., Williams, E. M., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masking: Support for signal enhancement. *Journal of Vision, 2*(6):4, 467–479, http://www.journalofvision.org/content/2/6/4, doi:10.1167/2.6.4. [PubMed] [Article]

Chang, A. Y.-C., Tzeng, O. J. L., Hung, D. L., & Wu, D. H. (2011). Big time is not always long: Numerical magnitude automatically affects time reproduction. *Psychological Science, 22,* 1567–1573.

Clausen, J. (1950). An evaluation of experimental methods of time judgment. *Journal of Experimental Psychology, 40,* 756–762.

Cohen, R. J., Swerdlik, M. E., and Smith, D. K. (1992). *Psychological testing and assessment*. Mountain View, CA: Mayfield Publishing Company.

Craik, K. J. W. (1940). Origin of visual after-images. *Nature, 145,* 512.

Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Danziger, K., & Du Preez, P. D. (1963). Reliability of time estimation by the method of reproduction. *Perceptual and Motor Skills, 16,* 879–884.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 314–326.

Doob, L. W. (1971). *Patterning of time*. New Haven, CT: Yale University Press.

Eagleman, D. M., & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B, 364,* 1841–1851.

Fraisse, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. VIII, pp. 203–254). New York: Academic Press.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology, 35,* 1–36.

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *Annals of Mathematical Statistics, 29,* 885–891.

Gomez, L. M., & Robertson, L. C. (1979). The filled-duration illusion: The function of temporal and nontemporal set. *Perception & Psychophysics, 25,* 432–438.

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

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in time-intervals discrimination task: A critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 993–1004.

Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychological Review, 55,* 297–313.

Hopfinger, J. B., & Mangun, G. R. (2001). Tracking the influence of reflexive attention on sensory and cognitive processing. *Cognitive, Affective, & Behavioral Neuroscience, 1,* 56–65.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13,* 1020–1026.

Kowal, K. H. (1993). The range effect as a function of stimulus set, presence of a standard, and modulus. *Perception & Psychophysics, 54,* 555–561.

Long, G. M., & Beaton, R. J. (1980). The contribution of visual persistence to the perceived duration of brief targets. *Perception & Psychophysics, 28,* 422–430.

Lord, F. M., & Novick, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addisson-Wesley.

Marks, L. E. (1992). The slippery context effect in psychophysics: Intensive, extensive, and qualitative continua. *Perception & Psychophysics, 51,* 187–198.

Masin, S. C. (2003). Test of the effect of attention on judged length of a line. *Perceptual and Motor Skills, 96,* 599–606.

Masin, S. C. (2008). Voluntary visual attention and phenomenal line length. *Perceptual and Motor Skills, 107,* 273–287.

Mo, S. S., & Michalski, V. A. (1972). Judgement of temporal duration of area as a function of stimulus configuration. *Psychonomic Science, 27,* 97–98.

Mueller, W. H., & Martorell, R. (1988). Reliability and accuracy of measurement. In T. G. Lohman, A. F. Roche, & R. Martorell (Eds.), *Anthropometric standardization reference manual* (pp. 83–86). Champaign, IL: Human Kinetics Books.

Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing. Principles and applications.* Upper Saddle River, NJ: Pearson Education.

Nicholls, M. E. R., Lew, M., Loetscher, T., & Yates, M. J. (2011). The importance of response type to the relationship between temporal order and numerical magnitude. *Attention, Perception, & Psychophysics, 73,* 1604–1613.

Oliveri, M., Vicario, C. M., Salerno, S., Koch, G., Turriziani, P., Mangano, R., … Caltagirone, C. (2008). Perceiving numbers alters time perception. *Neuroscience Letters, 438,* 308–311.

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

Ornstein, R. E. (1969). *On the experience of time*. Harmondsworth: Penguin Books.

Parducci, A. (1968). The relativism of absolute judgments. *Scientific American, 219,* 84–90.

Poulton, E. C. (1975). Range effects in experiments on people. *American Journal of Psychology, 88,* 3–32.

Ryan, L. R. (2011). Temporal context affects duration reproduction. *Journal of Cognitive Psychology, 23,* 157–170.

Schiffman, H. R., & Bobko, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology, 103,* 270–277.

Schneider, K. A. (2006). Does attention alter appearance? *Perception & Psychophysics, 68,* 800–814.

Schneider, K. A., & Komlos, M. (2008). Attention biases decisions but does not alter appearance. *Journal of Vision, 8*(15):3, 1–10, http://journalofvision.org/content/8/15/3/, doi:10.1167/8.15.3. [PubMed] [Article]

Thomas, E. A. C., & Cantor, N. (1975). On the duality of simultaneous time and size perception. *Perception & Psychophysics, 18,* 44–48.

Thomas, E. A. C., & Cantor, N. E. (1976). Simultaneous time and size perception. *Perception & Psychophysics, 19,* 353–360.

Tsal, Y., & Shalev, L. (1996). Inattention magnifies perceived length: The attentional field hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 233–243.

Tsal, Y., Shalev, L., & Zakay, D. (2005). The lengthening effect revisited: A reply to Prinzmetal and Wilson (1997) and Masin (1999). *Psychonomic Bulletin & Review, 12,* 185–190.

Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 166–179.

Vicario, C. M. (2011). Perceiving numbers affects the subjective temporal midpoint. *Perception, 40,* 23–29.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space, and quantity. *Trends in Cognitive Science, 7,* 483–488.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision, 7*(10):2, 1–5, http://www.journalofvision.org/content/7/10/2, doi:10.1167/7.10.2. [PubMed] [Article]

Yates, M. J., Loetscher, T., & Nicholls, M. E. R. (2012). A generalized magnitude system for space, time, and quantity? A cautionary note. *Journal of Vision, 12*(7):9, 1–7, http://www.journalofvision.org/content/12/7/9, doi:10.1167/12.7.9. [PubMed] [Article]

Zaidi, Q., Ennis, R., Cao, D., & Lee, B. (2012). Neural locus of color afterimages. *Current Biology, 22,* 220–224.

Zakay, D. (1990). The evasive art of subjective time measurement: Some methodological dilemmas. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 59–84). Hillsdale, NJ: Lawrence Erlbaum Associates.