A generalized magnitude system for space, time, and quantity? A cautionary note

Mark J. Yates
Psychological Sciences, The University of Melbourne, Melbourne, Australia

Tobias Loetscher
School of Psychology, Flinders University, Adelaide, Australia

Michael E. R. Nicholls
School of Psychology, Flinders University, Adelaide, Australia

We investigated the claim that larger stimuli are perceived to last longer (Xuan, Zhang, He, & Chen, 2007). This claim, along with other similar claims of interactions between magnitude representations, is frequently used to support the generalized magnitude system hypothesis—the suggestion that the brain represents information about different magnitudes (e.g., time, space, and quantity) via a common mechanism. It is not clear, however, whether the size of a stimulus genuinely affects the perceived duration of the stimulus or simply biases decisions about duration. This was addressed using duration equality judgments, which have been proposed to measure perceived duration unconfounded by decisional bias—in contrast to comparative judgments, which are generally considered bias-prone. Using equality judgments, we failed to find support for the claim that larger stimuli are perceived to last longer, despite replicating the original effect reported by Xuan et al. (2007) using comparative judgments. Instead, unexpectedly, larger stimuli were judged—though not necessarily perceived—as shorter in duration. This result casts doubt on the conclusions of a significant body of behavioral interference studies using comparative judgments, which support a generalized magnitude system. We also identify a hitherto unrecognized potential source of decisional bias associated with equality judgments.

Keywords: generalized magnitude system, temporal perception, size-duration interference


Introduction

The brain’s ability to encode information about magnitudes such as size, distance, duration, and quantity is fundamental for successful interaction with the external environment. Much work has been done in recent decades to uncover the mechanisms underlying this ability (for reviews, see Bueti & Walsh, 2009; Cohen Kadosh, Lammertyn, & Izard, 2008). A central idea within this literature is the proposal put forward by Walsh (2003) that the brain represents magnitudes across different dimensions (e.g., space, time, and numerosity) using a common abstract magnitude code rather than via distinct specific representations unique to each particular dimension.

A significant line of evidence put forward in support of this proposal is a class of behavioral interference experiments which demonstrate that, when participants are required to select which of two stimuli is greater in magnitude on a particular dimension (e.g., duration), their judgments are influenced by the magnitude of the stimuli on another dimension not relevant to the task (e.g., size or luminance; Dormal, Seron, & Pesenti, 2006; Levin, 1979; Oliveri et al., 2008; Stavy & Tirosi, 2000; Xuan et al., 2007).

A difficulty, however, with much of this work is that results of this nature are consistent with two alternative accounts. The existence of these two possible explanations has, until recently, rarely been considered within this literature—much less addressed—yet resolving this ambiguity is critical for identifying both the “where” and the “when” of any putative interactions between the processing of different dimensions. The first of these two accounts is that the magnitude of stimuli in the irrelevant dimension directly affects the perception of stimuli in the relevant dimension (e.g., brighter stimuli are perceived as longer). A second possibility is that the magnitude of stimuli in the irrelevant dimension merely biases decisions about the stimuli in the relevant dimension without necessarily affecting the perception of stimuli in that dimension.
In recent years, a number of purportedly perceptual effects have been demonstrated to be largely or entirely due to decisional bias (e.g., Frey, 1990; Nicholls, Lew, Loetscher, & Yates, 2011; Prinzmetal, Long, & Leonhardt, 2008. Also see Anton-Erxleben, Abrams, & Carrasco, 2010; Carrasco, Fuller, & Ling, 2008; Schneider & Bavelier, 2003; Schneider & Komlos, 2008; Schwarz & Eiselt, 2009; Shore, Spence, & Klein, 2001; Yates & Nicholls, 2011). Yet, corresponding attempts to separate perceptual versus decisional effects have been largely absent in investigations of the generalized magnitude system hypothesis, despite this being almost a model case where perceptual effects might be expected to be confounded with decisional effects, given the strong conceptual and linguistic similarities between magnitudes across different dimensions (e.g., the word “longer” can refer to both duration and distance; Casasanto & Boroditsky, 2008; Srinivasan & Carey, 2010).

In the current study, we focused on the demonstration by Xuan et al. (2007) that larger stimuli are judged to last longer and sought to address whether this is because the size of a stimulus genuinely affects its perceived duration or merely biases decisions about duration in a duration judgment task. The Xuan et al. (2007) study was selected on the basis that it investigated interactions between spatial and temporal dimensions, both cardinal dimensions in the original formulation of the generalized magnitude system hypothesis (Walsh, 2003), and fundamental dimensions for perception and behavior more generally. In addition, the study is both elegant as well as broadly representative of the class of behavioral interference experiments described earlier. Our method for isolating perceptual versus decisional effects might equally be applied to other similar experiments, and our intention with this study is not just to address this question for the specific finding of interference between spatial and temporal dimensions but to flag this issue more broadly across structurally similar experiments investigating all possible different pairs of stimulus dimensions (including future experiments). It is worth noting that Xuan et al. (2007) describe their findings as concerning “time perception” [italics added] (Xuan et al., 2007). However, the authors appear not to make any distinction between “judged duration” and “perceived duration” (e.g., “For example, velocity of motion is found to influence time perception...Judgments of duration are also found to be related with the intensity of visual stimuli.” Xuan et al., 2007). Thus, their conclusion regarding the influence of size on time perception appears to rely on the implicit assumption that judged duration is synonymous with perceived duration. At any rate, irrespective of the particular interpretation put forward by the authors, the underlying question remains: Are larger stimuli judged to last longer because size affects perceived duration or because size biases decisions about duration?

In the original Xuan et al. (2007) experiment, participants judged which of two consecutively presented squares (one large and one small) was presented for longer or shorter. The squares varied in duration such that one square (either large or small) was always presented longer. To assess whether the size of the squares influenced judgments about the duration of the squares, trials were grouped into “congruent” and “incongruent” categories. The congruent category consisted of trials in which the spatial and temporal dimensions of each stimulus were “congruent,” i.e., when the large square was presented for longer and the small square for the shorter duration. The converse association applied for the incongruent category. If size influences duration judgments, error rates should decrease in the congruent condition and increase in the incongruent condition. This effect was observed, and on this basis, the authors concluded that larger stimuli are judged to last longer (a claim we do not dispute) but also that larger stimuli are perceived to last longer (which we set out to test in the current study).

As outlined earlier, an influence of size on duration judgments might occur because size genuinely influences perceived duration or because size biases decisions about duration. Unfortunately, these are experimentally indistinguishable alternatives in the Xuan et al. (2007) study because the predicted effect on error rates in the congruent versus incongruent conditions is the same for both accounts.

Differentiating between these two alternatives is, in principle, possible by substituting the “comparison judgment” (“Which square is presented for longer or shorter?”) for an “equality judgment” (“Are the squares the same or different duration?”). In the equality judgment, the “response options are not bound to individual targets” (Schneider & Komlos, 2008), rather the response options (i.e., same or different) apply to both stimuli as a pair. This feature of the equality judgment is considered critical as it means that decisional bias (i.e., a tendency to select one response option over the other, e.g., “same duration” response more than the “different duration” response) should operate in an orthogonal direction to perceptual bias (genuine influence of size on perceived duration), such that perceptual effects can be identified unconfounded by decisional bias (for a more complete description, refer to Schneider & Komlos, 2008). It is important to note that the data we obtained in the current study prompted us to reconsider the assumption that equality judgments are unaffected by decisional bias. Despite this, data from equality judgments, unlike data from comparative judgments, can still potentially provide unambiguous support for the claim that larger stimuli
are perceived to last longer (refer to the Discussion section for a full explanation).

With this in mind, we first sought to replicate the size-duration behavioral interference effect reported by Xuan et al. (2007) using a duration comparison judgment as per the original study (Experiment 1). Next, we addressed whether their finding could be attributed to a genuine influence of size on perceived duration using a duration equality judgment (Experiment 2).

**Experiment 1**

Experiment 1 was a modified replication of Xuan et al. (2007). The original study employed a simple Stroop-like paradigm in which the presented pairs of squares were either “congruent” (i.e., a short small square and a long large square, in either order) or “incongruent” (i.e., a short large square and a long small square, in either order). The level of task difficulty did not vary across trials (all duration pairs had the same ratio, 1:1.25), and the dependent variable in their study was the number of errors made in each of these two conditions. While this method allows an initial assessment of whether size affects duration judgments, a key limitation is that it does not yield an estimate of the magnitude of the size/duration interference effect.

An improvement which allows the magnitude of the effect to be quantified is to incorporate a greater range of duration pairs than used in the original study. With multiple duration discriminations at various increments of difficulty, it is possible to generate individual psychometric functions. By using curve-fitting techniques, points of subjective duration equality can be derived, thereby quantifying effects of interest.

**Method**

**Participants**

Sixteen healthy participants from the University of Melbourne took part in this experiment in return for course credit (10 females, six males, ranging in age from 18 to 28, mean age of 19.8). All participants were righthanded with normal or corrected-to-normal vision. Participants gave informed, written consent prior to participating in the experiments in this study. Experiments received approval from the University of Melbourne Ethics Committee.

**Stimuli and procedure**

Participants sat at a desk facing a monitor interfaced with a PC computer, with their chin placed in a chinrest, located approximately 57 cm from the monitor. They were asked to judge which of two consecutively presented squares was displayed for the longer, or shorter, duration. (Participants were randomly assigned to the two judgment conditions.) The two squares (consisting of white outlines on a black background) were presented at the same central location (Figure 1). The squares varied in both size and duration. There were two square sizes: a “large” square (3.2° × 3.2°) and a “small” square (1.0° × 1.0°). The square pairs could be presented in one of eight possible size-duration pairings (i.e., duration of small/large squares was either 640/960, 680/920, 720/880, 760/840, 840/760, 880/720, 920/680, or 960/640 ms, respectively). The monitor refresh rate was set to 75 Hz (13.33 ms per frame, all stimuli durations were multiples of 40 ms, i.e., 3 × 13.33 ms). The summed duration of all square pairs was 1,600 ms. On all trials, the two squares were separated by a brief interstimulus interval of 200 ms during which a small white central fixation cross was present. On each trial, both the small and the large square were presented. Small squares were presented first on half the trials and large squares first on the remainder. This resulted in 16 unique stimulus configurations: 8 size-duration pairings × 2 size presentation orders (small square first or large square first). For each unique stimulus configuration, there were 24 individual trials, yielding a total of 384 trials, which were presented in randomized order.

Participants judged whether the former or latter square was presented for longer (or shorter, i.e., comparison/comparative judgment) via buttons located on the desk within comfortable reach. The two response buttons (one left and one right) were located immediately adjacent to each other on either side of participants’ midsagittal plane. Assignment of former/latter longer (or former/latter shorter) judgments to left/right response buttons was balanced across participants. Trials with response times >5 seconds were
automatically excluded and represented at a later point in the experiment.

Results

For each participant, the proportion of trials for which the large square was judged as longer in duration—\(P\) (large square judged longer)—was plotted as a function of the duration difference between the squares (large square duration minus small square duration). Averaged data is shown in Figure 2. Cumulative normal distribution functions were fitted to the data for each participant. Visual inspection of the plots for each participant suggested a good to very good fit, confirmed by corresponding \(R^2\) values that ranged from 0.795 to 0.998, with a mean of 0.937. The fitted functions were used to estimate the point of subjective duration equality (PSE) for each participant (derived as the duration difference at which the fitted functions crossed 0.5 on the \(y\)-axis).

The mean of the PSE values for all participants was \(-43\) ms, indicating that, on average, participants judged the two squares as equal in duration when the large square was \(43\) ms shorter in duration than the small square. A one-sample \(t\)-test performed on the PSE values demonstrated that participants’ PSE values were significantly shifted from 0 ms (\(t[15] = 2.248, p = 0.04\)). Data were collapsed across task type (“Which square is longer?” or “Which square is shorter?”), duration presentation order (longer square first or shorter square first), size presentation order (large square first or small square first), and response side assignment (“former”/“latter” responses left/right versus right/left) as none of these factors showed significant effects.

The results in this experiment replicate the finding of Xuan et al. (2007), which showed larger stimuli are judged to last longer. These results are consistent with both the hypothesis that the size of a stimulus influences its perceived duration, with larger stimuli perceived as longer than smaller stimuli, and the hypothesis that the size of a stimulus influences duration judgments about that stimulus without necessarily influencing its perceived duration.

Experiment 2

In Experiment 2, participants judged whether the two squares were presented for the “same duration” or “different durations”—an equality judgment. We reasoned that this type of judgment should allow us to test the hypothesis that the size of a stimulus influences the perceived duration of that stimulus, unconfounded by any potential influence of decisional bias. If larger stimuli are genuinely perceived to last longer, this effect should be observed using an equality judgment. If, instead, the effect identified in Xuan et al. (2007) is wholly, or partly, attributable to decisional bias associated with the duration comparison judgment, an absence or reduction of the effect should be observed. As highlighted earlier, the results from this experiment prompted us to re-evaluate the reasoning described earlier. Despite this, the rationale for conducting the equality judgment experiment remained intact (see Discussion section).

Method

Participants

Sixteen additional healthy participants (eight females, eight males, ranging in age from 18 to 23, mean age of 19.4) took part in this experiment. All participants except one were right-handed and had normal or corrected-to-normal vision.

Stimuli and procedure

Experiment 2 was identical to Experiment 1 except for two changes described here.

First, participants in Experiment 2 were instructed to judge whether the two squares were presented for the same duration or different durations. Assignment of same/different judgments to left/right responses was balanced across participants. Second, a new duration pair was added to the duration pairs employed in Experiment 1. The new duration pair was 800/800 ms for the small/large squares. This was to ensure that, on some trials, the “same duration” response was correct. The addition of the new duration pair resulted in the
creation of 18 unique stimulus configurations for this experiment: 9 size-duration pairings × 2 size presentation orders. For each unique stimulus configuration, 24 individual trials were presented, yielding a total of 432 trials.

Results

For each participant, the proportion of trials for which the two squares were judged as having been presented for equal durations—\(P(\text{squares judged same duration})\)—was plotted as a function of the duration difference between the large and small squares. Averaged data are shown in Figure 3. Normal distribution functions (with a constant scaling factor) were fitted to the data for each participant. Visual inspection of the plots for each participant suggested a good to very good fit, confirmed by corresponding \(R^2\) values that ranged from 0.786 to 0.983, with a mean of 0.91. The fitted functions were used to estimate the PSE for each participant, derived as the point at which the proportion of “same duration” responses was at its maximum.

The mean of the PSE values for all participants was \(+16\) ms, indicating that, on average, participants judged the two squares as equal in duration when the large square was 16 ms longer in duration than the small square. A one-sample \(t\)-test performed on the PSE values demonstrated that participants’ PSE values were significantly shifted from 0 ms (\(t[15] = 3.838, p = 0.002\)). This finding was unanticipated and is discussed more fully in the next section. Data were collapsed across duration presentation order, size presentation order, and response side assignment as none of these factors showed significant effects.

Discussion

This study sought to establish whether the size of a stimulus influences the perceived duration of that stimulus. Previous work has demonstrated that stimulus size can influence duration judgments such that larger stimuli are judged to last longer (Xuan et al., 2007). Yet this finding can be attributed to either a genuine effect of size on perceived duration—which in turn influences duration judgments—or by an effect of size on duration decisions which leaves perceived duration unaffected.

Experiment 1 replicated the original finding of Xuan et al. (2007), namely that larger stimuli are judged to last longer, using a duration comparison judgment as per the original study. Using stimuli comparable to the stimuli employed in the original study, the large and small squares were judged as equal in duration when the large square was shorter than the small square by around 43 ms.

The critical experiment was Experiment 2, which, using a duration equality judgment, sought to determine whether this effect of size on duration judgments could be attributed to a genuine effect of size on duration perception. If large squares were truly perceived as longer, the leftward shift in PSE values observed in Experiment 1 would also be predicted for Experiment 2. The results of this second experiment were surprising. Not only was the original effect from Experiment 1 (leftward shift in PSE) not replicated in the second experiment, there was in fact a clear shift in the opposite direction. This result did not support the claim that larger stimuli are perceived to last longer. Rather, based on the results for Experiment 2, the point at which large and small squares were judged as equal in duration was estimated to occur when the large square was 16 ms longer than the small square. If PSE values in this experiment were based entirely on perceived duration, this would be expected to occur when the large square was 16 ms longer than the small square. If PSE values in this experiment were based entirely on perceived duration, this would be expected to occur if in fact larger stimuli are perceived to last shorter than smaller stimuli. Using two different approaches to assess the influence of stimulus size on perceived duration, two apparently contradictory conclusions emerged.

From this, it can be concluded that at least one of the two types of duration judgments is prone to some form of decisional bias, though no further conclusion can be drawn based on the data presented here. The possible decisional bias operating in Experiment 1 (which would shift PSE values leftward in a manner indistinguishable from perceptual bias) has already been well described. Here, we propose an additional, previously unconsidered form of decisional bias operating in Experiment 2.

Figure 3. Averaged data for Experiment 2 showing proportion of responses in which the squares were judged as equal in duration as a function of the duration difference between the squares. The hypothetical unbiased function was modeled on the averaged data but symmetrical around the line \(x = 0\).
When judging whether the squares are the same or different durations, participants’ responses may be affected not only by whether the squares are perceived as the same or different durations but also by whether the squares are the “same” or “different” in a more abstract sense. More specifically, participants may judge whether the spatial and temporal dimensions of each square in the square pair are “congruent” or “incongruent.” If the small square is presented for the shorter duration, this square has spatial and temporal dimensions that are the “same” in the sense that both are the “lesser” magnitudes. Conversely, if the small square is presented for the longer duration, this square has spatial and temporal dimensions that are “different.” On congruent trials, both squares have spatial and temporal dimensions that are the “same” in this sense, even though one square is small and short and the other is large and long. Similarly on incongruent trials, both squares have “different” spatial and temporal dimensions. The effect of such a decisional bias would therefore be to raise the proportion of “same duration” responses for pairs of squares whose spatial and temporal dimensions are congruent (x-axis values > 0 ms, right half of x-axis, refer to Figure 3) and lower the proportion of “same duration” responses for pairs of squares belonging to the incongruent condition (x-axis values < 0 ms, left half of x-axis). This asymmetrical raising and lowering of the response function would translate into a rightward shift of PSE values, equivalent to what would be observed if larger stimuli are perceived as shorter than smaller stimuli. It is worth noting that, despite this potential source of decisional bias in the equality duration judgment task, this task nevertheless remains an improvement on the comparative duration judgment task for the purposes of testing the claim that larger stimuli are perceived to last longer. This is because the decisional bias proposed to operate in the equality duration judgment task (Experiment 2) would shift PSE values to the right, therefore any observed leftward shift of PSE values in this experiment would provide unambiguous evidence that larger stimuli are perceived to last longer given that a PSE shift in this direction cannot be attributed to decisional bias. This is not the case for the comparative duration judgment task (Experiment 1) in which a leftward PSE shift could represent either a genuine perceptual effect or a decisional bias.

The actual PSE shift in Experiment 2 was to the right. Regardless of the underlying cause of this rightward shift, the original claim we set out to test—that larger stimuli are perceived to last longer—was not supported. Indeed, based on the data from the two experiments presented in this study, it is equally valid to claim that larger stimuli are perceived to last shorter rather than longer. While the original claim pertains to one specific pairing of stimulus dimensions (size and duration) proposed to share a common magnitude representation, similar claims have been made for other pairs of stimulus dimensions. Together, such claims constitute a major plank of evidence for the generalized magnitude system hypothesis. In light of the results of the current study, we suggest that such claims should be regarded with caution.

Acknowledgments

This work was supported by Australian Research Council Discovery Project DP0986118.

Commercial relationships: none.
Corresponding author: Mark J. Yates.
Email: mjyates@unimelb.edu.au.
Address: Psychological Sciences, The University of Melbourne, Melbourne, Australia.

References


dence supporting and rejecting the doctrine of prior entry. Human Movement Science, 9, 481–530.


