Lisi, M. & Gorea, A. (2016). Time constancy in human perception. Journal of Vision, 16(14), 3.. doi: 10.1167/16.14.3

City Research Online

Time constancy in human perception

Matteo Lisi

Université Paris Descartes, Sorbonne Paris Cité, Paris, France
CNRS, Laboratoire Psychologie de la Perception, UMR 8242, Paris, France ✉

Andrei Gorea

Université Paris Descartes, Sorbonne Paris Cité, Paris, France
CNRS, Laboratoire Psychologie de la Perception, UMR 8242, Paris, France

Estimated time contracts or dilates depending on many visual-stimulation attributes (size, speed, etc.). Here we show that when such attributes are jointly modulated so as to respect the rules of perspective, their effect on the perceived duration of moving objects depends on the presence of contextual information about viewing distance. We show that perceived duration contracts and dilates with changes in the retinal input associated with increasing distance from the observer *only* when the moving objects are presented in the absence of information about the viewing distance. When this information (in the form of linear perspective cues) is present, the time-contraction/dilation effect is eliminated and time constancy is preserved. This is the first demonstration of a perceptual *time constancy*, analogous to size constancy but in the time domain. It points to a normalization of time computation operated by the visual brain when stimulated within a quasi-ecological environment.

Introduction

Time (or duration) perception has been known to be context dependent for at least a century (Allman, Teki, Griffiths, & Meck, 2014; Buhusi & Meck, 2005; Eagleman, 2008; Fraisse, 1963; Gorea, 2011; Hass & Durstewitz, 2014; James, 1890; van Wassenhove, 2009). It fluctuates with speed, with temporal frequency or the number of events per time unit, with numerosity and size, and with sensory adaptation, not to mention factors extrinsic to the sensory stimulation (e.g., attention, emotion, drugs). Many of these contextual influences arise from modulations of spatiotemporal features of the stimulation (e.g., size, speed). For

example, the perceived temporal interval between two stimulations increases with their spatial separation (Abe, 1935), and the perceived duration of a dynamic visual event (e.g., a moving object) dilates with increasing speed (Gorea & Kim, 2015; Kanai, Paffen, Hogendoorn, & Verstraten, 2006; Kaneko & Murakami, 2009; Linares & Gorea, 2015).

In ecological conditions, the spatial features of a visual event are largely dependent on the viewing distance: We live in a perspective world, where faraway objects tend to have smaller retinal projections, move more slowly, and cover shorter retinal distances. As they cross the retinas, the perceived duration of dynamic visual events would therefore be expected to vary with their distance from the observer, as a consequence of the changes in their retinal features. Moreover, on the basis of the studies already mentioned, some of the changes in the retinal stimulation associated with different viewing distances would be expected to contract perceived duration and others to dilate it. For example, the slower speed of a distant object should lead to time contraction (Gorea & Hau, 2013; Kanai et al., 2006), while its smaller size (Mcgraw, Roach, Badcock, & Whitaker, 2012) and the well-known Ponzo illusion (Ponzo, 1911) should enhance the apparent length of the trajectory it covers, which in turn should lead to time dilation (Gorea & Hau, 2013). Also, the larger size of near objects should lead to time dilation (Rammsayer & Verner, 2014; Xuan, Zhang, & Chen, 2007). Are all these synergic and conflicting modulators of our time perception eliminated in the perspective world we live in? Do we live in a unified time-perspective world?

It is well known that humans show *size constancy*: Our assessment of the objective (or *distal*) size of objects remains largely unaffected by changes in the

retinal (or *proximal*) size entailed by the viewing distance (Boring, 1942). Size constancy depends on the presence of information about the viewing distance (Holway & Boring, 1941), to which linear perspective provides a strong cue (Aks & Enns, 1996; Fineman, 1981). Based on a conjecture of Gorea and Hau (2013), we designed a series of experiments to test whether our perceptual system also shows *time constancy*[1] in a 3-D world as it is represented by its 2-D projection.

The time-constancy hypothesis predicts that the perceived duration of visual events should be largely unaffected by changes in their distance from the observer, and therefore of their proximal visual attributes, provided that information about the viewing distance is present. However, when information about viewing distance is absent, the same changes of the proximal attributes should induce systematic distortions of the perceived duration.

We tested this hypothesis in a series of experiments by asking observers to judge the duration of simple dynamic events consisting of 3-D rendered balls rolling along horizontal paths placed at different distances from the observer in fronto-parallel planes. Each of the durations (that is, the intervals between the appearance and disappearance of the rolling ball) was rendered as one uniquely defined event in the virtual dimension. The same duration (or visual event) could be rendered as if it were placed at different distances from the observer, so that each distance resulted in a different level of *foreshortening* of the rolling ball (i.e., different combinations of the ball's proximal attributes: size, speed, and length of the motion path). Our stimuli therefore can be described as different combinations of temporal duration and foreshortening of the rolling ball. The stimuli were presented in two different contextual conditions: either on a uniform gray background (hereafter referred to as the *flat* condition; for details, see Figure 1a, c, and d and Methods) or on a linear perspective projection of a checkerboard "floor" below a blue "sky" (hereafter, the *perspective* condition; Figure 1b and e). The flat condition allows testing whether the foreshortening of the moving ball, in the absence of other information about viewing distance, does systematically distort the perceived duration. The perspective condition, on the other hand, is meant to reveal whether contextual information about the viewing distance (in the form of simple linear perspective cues) can eliminate these distortions, provided that perspective rules are respected.

In the experiments, each event (the ball travelling along its path) could last for one of four durations, logarithmically spaced (600, 780, 1014, and 1318 ms). These durations were obtained in two distinct experiments (see Methods for details) by either slightly modulating the ball's distal speed so as to maintain the length of the motion path constant in the virtual dimension (*constant-length* experiment) or by letting the distal trajectory length vary while keeping the distal speed constant (*constant-speed* experiment). This allowed us to control for the potential use of the trajectory length (or speed) as a proxy for duration judgments. As anticipated, the rolling ball assumed one of four levels of foreshortening. In the perspective conditions (Figure 1e), each level was always associated with one of the four vertical locations on the screen (corresponding to four depth planes in the virtual dimension), so that speed, path length, and size scaled in agreement with the perspective rules. In two of three flat conditions, however, the different levels of foreshortening of the rolling balls were randomly associated with the balls' vertical locations on the screen (Figure 1c). As a consequence, the distal *event constancy* (i.e., the ability to recognize the event as having the same properties in the virtual dimension despite changes in its 2-D representation on the screen) over locations was compromised. (Note that we will hereafter use the term *foreshortening* to indicate proximal changes in the stimulus in both the flat and perspective conditions.) In a third flat condition the level of foreshortening of the rolling ball scaled systematically with the vertical location of the ball, just like in the perspective condition (Figure 1e). In this case, the four different durations were always obtained by varying the length of the trajectory of the rolling ball (hence keeping constant its distal speed). This flat condition was meant to isolate the pure contribution of the perspective-rendered background to duration perception.

To assess the contribution of the foreshortening of the rolling balls to perceived duration, we used the method of conjoint measurement (Ho, Landy, & Maloney, 2008; Luce & Tukey, 1964). The method is ideally suited for assessing how judgments made on a single perceptual dimension (here, perceived duration) are affected by changes along multiple, heterogeneous physical dimensions (here, physical duration and foreshortening). The method requires that stimuli be presented in pairs and that observers order them in each pair according to some criterion—here, perceived duration (i.e., which of the two stimuli lasted longer; for details, see Figure 1a and b and Methods).

Methods

Stimuli and apparatus

Observers sat in a quiet, dimly lit room. Stimuli were presented on a Mitsubishi Diamond Plus 230SB CRT monitor (screen resolution = 1600 × 1200, vertical refresh rate = 85 Hz) and were generated in OpenGL using custom-made software running under the Xenomai real-
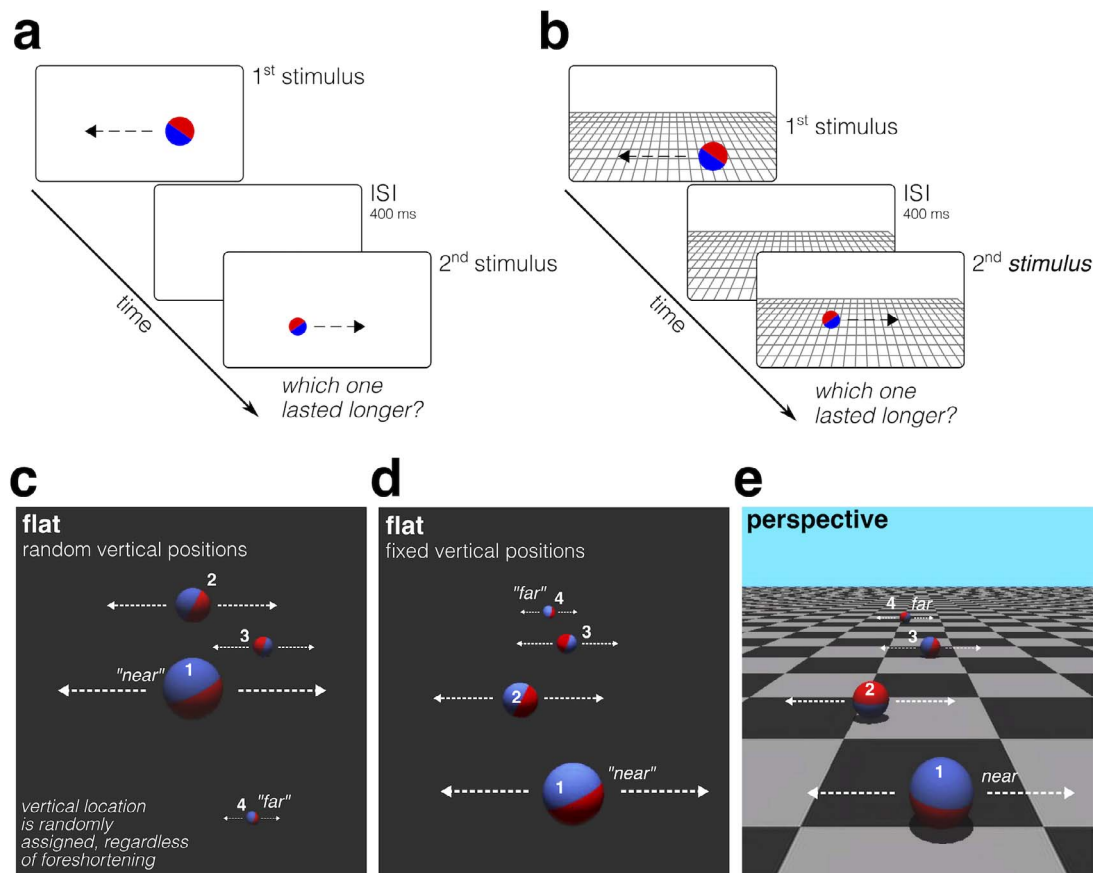
Figure 1. Experimental paradigms and stimuli. The experimental paradigm is illustrated for (a) the flat conditions (empty gray background) and (b) the perspective conditions (a perspective-rendered checkerboard floor and blue sky; see also e). In each trial, a rolling ball appeared at one of four random vertical locations, followed after 400 ms by a second rolling ball also randomly placed at one of the four vertical locations. (c) The four ball sizes (corresponding to four levels of foreshortening, from near to far), placed on the uniform (flat) background at one of four vertical locations chosen randomly and independently of the ball size. (d–e) The four ball sizes now placed at vertical locations corresponding to the correct depth planes (relative to their sizes) on (d) the uniform background and (e) on the perspective-rendered background. In the latter case, the four proximal ball sizes translate into only one distal size within the virtual dimension.

time framework for Linux (http://www.xenomai.org). The viewing distance was about 45 cm. Rolling balls of four sizes (on-screen diameter of about 2.2, 1.2, 0.7, and 0.4 cm) covered horizontal trajectories of variable lengths and speeds at one of four vertical locations (7.8, 11.0, 12.9, and 14.0 cm from the bottom edge of the screen). The duration of the ball motion could take one of four logarithmically spaced temporal values (600, 780, 1014, and 1318 ms). The screen either was uniform dark gray (0.8 cd/m^2) or displayed a linear perspective projection of a checkerboard "floor" (mean luminance = 2.3 cd/m^2) below a blue "sky" (10.3 cd/m^2). In the latter case the four y-locations translated into four depth planes equally spaced in log units as seen by the observer (1.8, 3.2, 5.8, and 10.5 squares of the checkerboard floor). The width of the squares as rendered on the screen was 14.6 cm for the central square at the lower edge of the display, decreasing to 0.2 cm at the level of the horizon. When placed at the corresponding elevations, the four ball sizes translated

into a unique size in the virtual 3-D world: The largest ball placed in the nearest depth plane had the same size (in virtual dimension) as the smallest ball placed in the most remote depth plane. Note that in order to distinguish the characteristics of the moving ball in the virtual 3-D world from their 2-D projection on the screen, we use the terms *distal* and *proximal*, respectively.

Two different experiments were run, in which the four durations were obtained either by varying the distal trajectory length while keeping the distal speed constant (constant-speed experiment) or by varying the distal speed and keeping the distal length of the trajectory constant (constant-length experiment; see Figure 2). In the constant-length experiment the distal trajectory length was always equal to the mean of the distal lengths used in the constant-speed experiment (averaged over the four physical durations). Each experiment consisted of two separate conditions, with and without the perspective checkerboard background floor (perspective
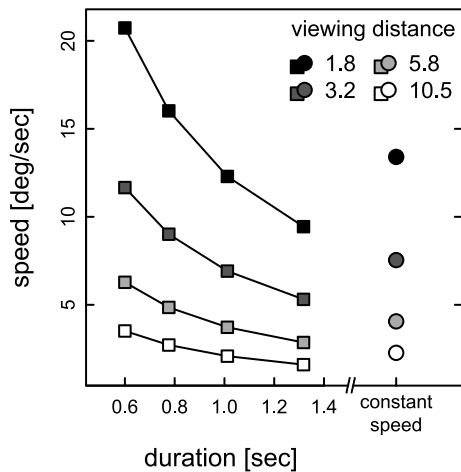
Figure 2. Relationship between proximal speed and duration in the constant-length experiment. The proximal speeds (in °/s) of the rolling balls used in the constant-length experiment are plotted as a function of the tested durations. Different symbol shades indicate the four foreshortening levels, which correspond to four different viewing distances in the virtual dimension (reported in the inset as the number of squares of the checkerboard plane). The different speeds were obtained by choosing one unique trajectory length (defined in the virtual dimension, and corresponding to the average of the trajectory lengths of all durations in the constant-speed experiment; see main text for details). In the constant-speed experiment, the speed of the ball was constant for all durations and translated into four proximal speeds depending on the level of foreshortening or viewing distance.

and flat conditions; respectively, Figure 1b and e and Figure 1a, c, and d). For the perspective background condition, the proximal sizes, speeds, and trajectory lengths of the rolling balls covaried with the location of the ball in the perspective plane so as to be commensurate with the foreshortening laws of linear perspective, resulting in four levels of foreshortening of the moving ball. For the flat background condition, the proximal ball size and speed or length also covaried, resulting in four foreshortening levels, identical to those of the perspective condition but randomly displayed at one of the four *y*-locations on the screen (Figure 1c). An additional control experiment with a flat (uniform) background was run, where the *y*-location of the moving ball on the screen covaried with the foreshortening of the ball just like in the perspective condition (flat with fixed vertical positions; Figure 1d).

Participants

Eight human observers (five women and three men; mean age = 35 years, SD = 11.87) paticipated in the perspective and flat conditions of Experiment 1

(constant-speed experiment), eight (two women and six men; mean age = 38.13 years, SD = 12.46) in the perspective and flat conditions of Experiment 2 (constant-length experiment), and eight (three women, five men; mean age = 32.37 years, SD = 14.68) in the control experiment with flat background and fixed positions; two observers (the authors) were shared by the three groups. The number of participants was chosen so as to permit us to obtain sensible estimates of standard deviations for the random-effects terms of the models used in the analysis (Bates, 2010; for details on the models, see the Conjoint measurement models subsection). All had normal or corrected-to-normal vision and gave their informed consent to perform the experiments. The study was conducted in accordance with French regulations and the requirements of the Helsinki Convention.

Procedure

We used the method of conjoint measurement (Ho et al., 2008; Knoblauch & Maloney, 2012; Luce & Tukey, 1964). The method calls for a simple psychophysical task: Two stimuli are presented in succession (or at different locations), and the observer is required to order them according to some criterion. In our task, observers were asked to report which of the two stimuli in a pair lasted longer (Figure 1a and b), allowing us to assess the contamination of duration judgments by contextual variables (for details, see the Conjoint measurement model subsection). The contextual variable here was the foreshortening of the moving ball (that is, the joint variation of the proximal speed and size of the ball and the length of its motion path), measured in different conditions where information about viewing distance (the checkerboard floor rendered in linear perspective) was present or absent.

One trial unfolded as follows: The empty or perspective background was presented for 400–500 ms (uniformly distributed) before the appearance of the first rolling ball for one of four durations at one of four vertical locations, followed by an interstimulus interval of 400 ms and then the appearance of the second rolling ball also for one of four durations at one of four vertical locations. The next trial was initiated immediately after the response. The trajectories covered by the moving ball were presented always centered with respect to the screen, with some horizontal jitter (uniformly distributed within a range of ±40% of the side of the checkerboard square on which the ball was located). The direction of motion (leftward or rightward) of the first and second rolling ball in a pair were independently randomized.

There were 16 possible stimuli in each experiment (4 durations × 4 levels of foreshortening), yielding 120

different stimulus pairs (with order discarded and repetitions forbidden within a pair). Each stimulus pair was presented four times (with the order of the stimuli balanced across repetitions of the same pair) in each experiment, yielding 480 trials per condition and per observer, translating into a duration of about 35 min. per condition. We did not manipulate all the properties of the stimulus independently (e.g., proximal size, proximal speed, vertical location), because the number of possible stimulus pairs would have then increased to a prohibitive number. The order of the flat and perspective experimental conditions (in both constant-length and constant-speed experiments) was counter-balanced across observers so that half of the observers started with the condition with the perspective back-ground and the other half started with the condition with no background.

Conjoint measurement models

We modeled the data within the framework proposed by Ho et al. (2008) and Knoblauch and Maloney (2012) for conjoint measurement experiments. We started by fitting an additive model to our data. The model assumes that perceived duration results from the linear sum of separate contributions of the physical duration and foreshortening. More specifically, the perceived duration ψ of a given stimulus was modeled as

$$\psi_{ij} = D_i + F_j \quad (1)$$

where the indices $i \in \{1, 2, 3, 4\}$ and $j \in \{1, 2, 3, 4\}$ indicate, respectively, the level of physical duration (D) and foreshortening (F) of the stimulus (see Stimuli and apparatus subsection). If perceived duration is not affected by variations in depth, then F should be equal to zero for all j. When two stimuli are compared, we assume that observers base their decisions on the noise-contaminated variable Δ:

$$\Delta_{ijkl} = \psi_{ij} - \psi_{kl} + \varepsilon = (D_i + F_j) - (D_k + F_l) + \varepsilon$$
$$\varepsilon \sim \mathcal{N}(0, 4\sigma^2)$$
$$(2)$$

so that ψ_{ij} will be judged as longer than ψ_{kl} when $\Delta > 0$ (with ε representing a normally distributed judgment error). For model identifiability, we anchored the perceptual scales by setting the scale values for the first stimulus level of each dimension to zero (i.e., $D_1 = F_1 = 0$). Following Gerardin, Devinck, Dojat, and Knoblauch (2014), we scaled the estimated values of D and F so that they would be on the same scale as the sensitivity index d' (Green & Swets, 1966).

In addition—and different from Ho et al. (2008), who fitted a model for each observer—we fitted a single

mixed-effects (or hierarchical) additive model for each condition and experiment, under the assumption that the scale values D and F are normally distributed across observers. Within the mixed-effects modeling frame-work, the perceived duration of a given stimulus is modeled as a linear combination of *fixed* (D and F) and *random*, or observer-specific, effects (u_s); for example, for a given observer s, the perceived duration of a stimulus with physical duration level i and foreshortening level j is

$$\psi_{ij,s} = D_i + u_{i,s}^D + F_j + u_{j,s}^F$$
$$u \sim \mathcal{N}(0, \Sigma) \quad (3)$$

where Σ indicates the variance-covariance matrix for the multivariate Gaussian distribution of the random (observer-specific) effects u. This model can be formu-lated as a generalized linear mixed-effects model (Knoblauch & Maloney, 2012) with observer as the grouping factor. We estimated its parameters by maximum likelihood using R (R Core Team, 2015) and the lme4 library (Bates, Maechler, Bolker, & Walker, 2014).

Since the estimated values for the perceptual scales (D and F) in our experiment showed a clear linear dependence on the stimulus levels, we introduced an additional simplification and fitted an *additive-linear* model, where these scales are modeled as linear functions of the stimulus indices. In the additive-linear model the decision variable can be notated as

$$\Delta_{ijkl} = (i\beta_D + j\beta_F) - (k\beta_D + l\beta_F) + \varepsilon$$
$$= (i - k)\beta_D + (j - l)\beta_F = \delta_{ik}\beta_D + \delta_{jl}\beta_F \quad (4)$$

where the β values are the linear slopes; i, j, k, and l are the indices indicating the levels of the stimuli; and δ indicates the differences between the indices. This model can also be formulated as a generalized linear mixed-effects model, where the slope for each observer s is modeled as a sum of fixed (β) and random (b_s) effects:

$$\Phi^{-1}\left(p(\Delta_{ijkl} > 0|s)\right) = \delta_{ik}(\beta_D + b_{D,s}) + \delta_{jl}(\beta_F + b_{F,s})$$
$$b \sim \mathcal{N}(0, \Sigma)$$
$$(5)$$

Here Φ indicates the cumulative distribution func-tion of the standard normal distribution, and $p(\Delta_{ijkl} > 0|s)$ corresponds to the probability that the observer s judges the stimulus with physical duration and foreshortening levels i and j, respectively, to last longer than the stimulus with physical duration and fore-shortening levels k and l.

In order to test the effect of foreshortening on the duration judgments, we compared for each condition and experiment the additive-linear model with an *independence* model where the fixed effect of fore-

shortening is set to zero (that is, a model where $\beta_F = 0$). Across the two models (additive-linear and independence), we kept constant the structure of the random components b. Since the independence model is nested within the additive-linear model, we compared the two models using a likelihood-ratio test. Additionally, we tested the effect of the perspective background at the within-observer level by fitting both conditions (flat and perspective) with a single model that included parameters for the changes in the slopes between the two conditions, and by comparing these models via likelihood-ratio tests with reduced models where the changes in slopes were forced to be zero.

Finally, in order to evaluate whether the additive-linear model provides an adequate description of the data, we fitted a *saturated* model where we introduced an additional coefficient β_{FD} that was applied to a product of the indices to obtain an additional interaction term that could account for interactions between physical duration and foreshortening. As with the independence model, this saturated model was compared to the additive-linear models with likelihood-ratio tests (keeping the structure of the random effects constant).

Results

The effect of foreshortening on the observers' responses averaged over all durations and observers (that is, the marginal effect of foreshortening) is represented as different shades of the discs in the top panels of Figure 3. Duration judgments in the flat conditions (left and middle top panels in Figure 3) were clearly dependent on the magnitude of the proximal attributes: On average, observers tended to judge "near" events (i.e., larger balls moving faster over a longer path) to last longer than "far" events. This was not the case for the perspective condition (right top panel in Figure 3), where duration judgments appear to have been unaffected by the foreshortening of the rolling balls.

To quantify the interaction between physical duration and the foreshortening of the moving balls, we initially fitted the data with an additive model (Ho et al., 2008; Knoblauch & Maloney, 2012; see Conjoint measurement models subsection). The parameters of the model represent the additive contributions to the perceived duration of each level of stimulus physical duration and of foreshortening (except for the shortest and smallest values, respectively, which were set to 0 for model identifiability), expressed in units of d' (Gerardin et al., 2014; Green & Swets, 1966). The strength of this modeling approach is that it allows measurement of the effect of different, heterogeneous physical variables on

a common perceptual scale, even when these variables are defined on different domains (here, temporal and spatial). Hereafter, the notation D refers to the coefficients coding for the contribution of physical duration, and F to the coefficients coding for the contribution of foreshortening. They denote the contribution to the perceived duration of their physical counterparts. We estimated the set of scale values that best capture observer's judgments of the perceived duration difference between the stimuli in each pair by maximizing the likelihood of observer's responses under the additive model (for details, see Methods). The estimated scale values (averaged over observers) are represented in Figure 3 (bottom panels) as a function of the stimulus level (1 to 4), where stimulus level refers to either physical duration (the four durations used, in increasing order; open symbols) or foreshortening (the four combinations of proximal attributes, ordered from near to far; shaded symbols). The values for F show different trends depending on the experimental condition: They decrease in the flat conditions, indicating a contraction of perceived duration with increasing levels of foreshortening (from near to far), while they stay around 0 in the perspective conditions. This indicates that the net effect of perspective foreshortening, in the flat condition, is to make the perceived duration contract, the more so as the speed, size, and length of the moving ball decrease. This effect is abolished in the perspective condition.

An interesting feature of the estimated scale values D and F is their linear dependence on the stimulus levels. This linearity is likely a consequence of the logarithmic spacing of the stimuli, which (in accordance with Weber's law) made them perceptually equidistant (Rogers, Knoblauch, & Franklin, 2016). We took advantage of this linearity to model the data at the group level with a generalized linear mixed-effects model, following the approach described by Rogers et al., where the estimated scale values are treated as linear functions of the stimulus levels. As a consequence, each contribution to perceived duration (physical duration or foreshortening) can be specified by a single parameter, the slope of the linear function, with the advantage of largely reducing the number of parameters (for details, see Methods). We called this model additive-linear because it is based on the assumption that all observers share the same underlying linear shape of the perceptual scales, although they might differ in sensitivity (the slope). We tested this assumption by comparing the more parsimonious additive-linear model to a hierarchical version of the additive model of Ho et al. (2008), which was fitted at the group level and does not make any assumptions about the underlying shape of the perceptual scales (for details, see the Conjoint measurement models subsection). We compared the models using the Akaike
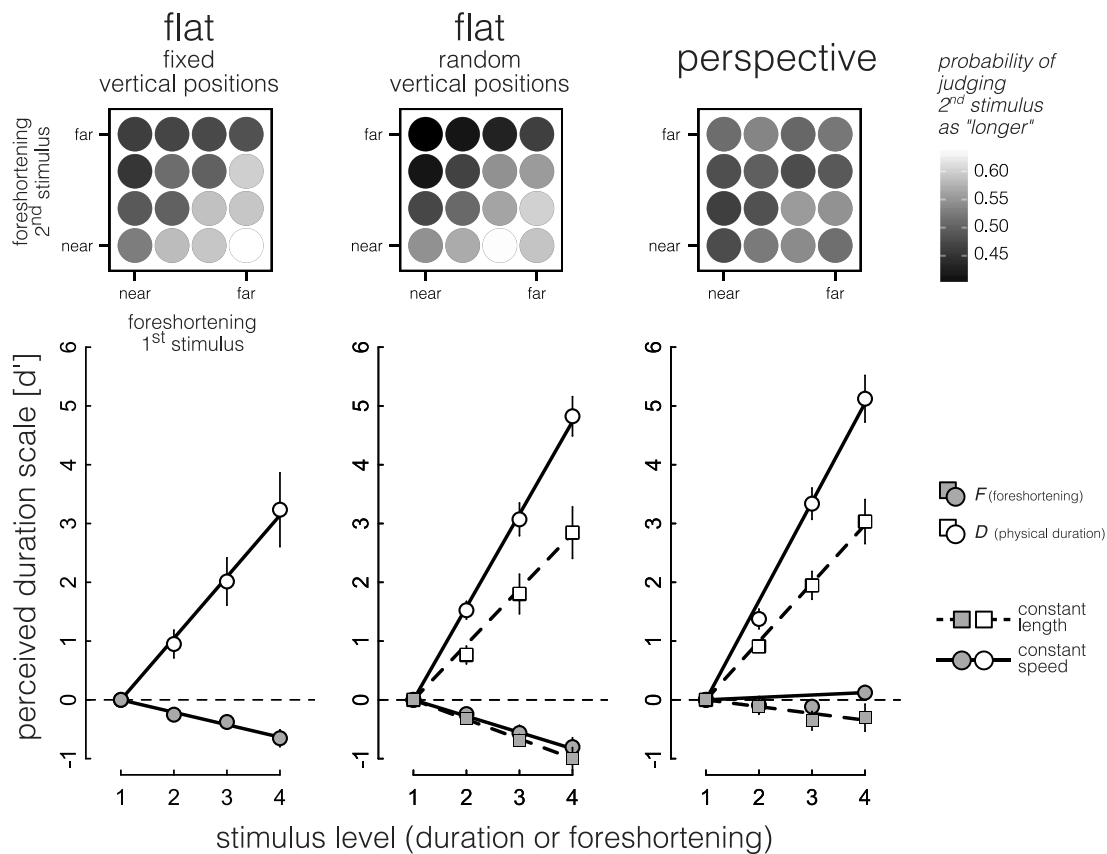
Figure 3. Results. (Top) The marginal effect of foreshortening (four levels, from near to far; see Figure 1) on the duration judgments for each of the three main experimental conditions (flat conditions with fixed and random vertical positions, and perspective condition); the shading of the discs indicates the proportion of responses "second stimulus longer" averaged over all durations and observers. Only in the flat conditions did observers tend to judge near stimuli to last longer than far stimuli (for equal physical durations). (Bottom) The perceived duration scales for the same three conditions as in the top panels, plotted as a function of the four stimulus levels (either physical duration or foreshortening; see main text for details). Points represent the estimated scale values for the additive model averaged across observers (error bars show ± 1 *SEM*), while the lines show the additive-linear model fit to the data (for details, see Methods). Note that the first level along both dimensions (physical duration and foreshortening) is fixed at 0 for model identifiability (see main text), and is used as a baseline for the remaining scale values; the latter thus represent the difference in perceived duration relative to the baseline. The overall effect of foreshortening was evaluated with likelihood-ratio tests that were significant only for the flat conditions (see main text).

information criterion (Akaike, 1974), a measure of the relative quality of statistical models. We did the comparison separately for each experiment and condition: Differences in the Akaike information criterion ranged between 18 and 36, and in all cases favored the additive-linear model. Thus the assumption that all observers have linear perceptual scales allows for a model that uses fewer parameters while fitting the data equally well as the additive model.

We statistically tested the effect of foreshortening by comparing the additive-linear model with the independence model (for details, see the Conjoint measurement models subsection). The likelihood-ratio test revealed that the slope coding for the influence of foreshortening was significantly different from 0 in both constant-speed, $\chi^2(1) = 11.81$, $p < 0.001$, and constant-length, $\chi^2(1) = 11.89$, $p < 0.001$, experiments, confirming that

foreshortening induces in these conditions a significant contraction of the perceived duration. Additionally, the effect of foreshortening was significant in the control experiment with a flat background and fixed vertical positions, $\chi^2(1) = 7.84$, $p = 0.005$, where stimuli appeared at the same screen locations as in the perspective condition. Although this condition contains some information on the 3-D layout, it appears that it is not sufficient to cancel the effect of foreshortening. Moreover, this result rules out any potential confound due to the balls' location differences between the two flat conditions (with fixed and random trajectory locations). In contrast, the same test (additive vs. independence model) performed on the results of the perspective conditions (i.e., with the perspective background) did not reach significance for either the constant-speed, $\chi^2(1) = 0.59$, $p > 0.25$, or the constant-
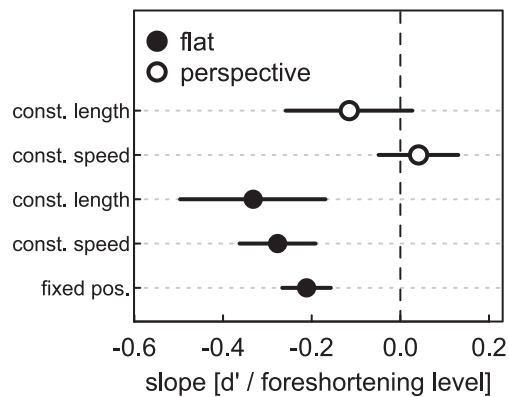
Figure 4. Effects of foreshortening on perceived duration. The linear slopes of the component scales coding for the effect of foreshortening on perceived duration are represented for each experiment and condition, along with bootstrapped 95% confidence intervals. The magnitude of the slope indicates by how much (in units of d') the perceived duration of a stimulus contracts after a 30% increase of the viewing distance. It is only in the flat conditions (filled black dots) that the confidence intervals did not include 0.

length, $\chi^2(1) = 2.28$, $p = 0.13$, experiment. The effect of foreshortening on the apparent duration, as summarized by the slopes of the component scales, is represented in Figure 4 for all experiments and conditions. Overall, for the stimuli tested, foreshortening made a contribution of about 23% to perceived duration in the flat condition (ratio of the slopes for foreshortening and physical duration) and only 2% in the perspective condition.

The perspective and flat conditions of the constant-length and constant-speed experiments were performed by the same groups of observers, so we tested whether the differences in slope between these conditions resulted in significant within-subject interactions (for details, see Methods). We found a significant change in the slope for the contribution of foreshortening in both constant-speed, $\chi^2(1) = 53.33$, $p < 0.001$, and constant-length, $\chi^2(1) = 5.91$, $p = 0.01$, experiments, revealing that the contribution of foreshortening to perceived duration depended on the presence or absence of perspective cues. Conversely, no significant change in slope between flat and perspective conditions was found for the contribution of physical duration, in either the constant-speed, $\chi^2(1) = 2.84$, $p = 0.09$, or constant-length, $\chi^2(1) = 0.36$, $p > 0.25$, experiments. To sum up, in agreement with our hypothesis of time constancy, we find that duration judgments are unaffected by changes in the proximal aspects of the stimulus (foreshortening) only when those changes are made in conditions where information about viewing distance is present (linear perspective cues).

The comparison with the saturated model (see the Conjoint measurement models subsection) had non-
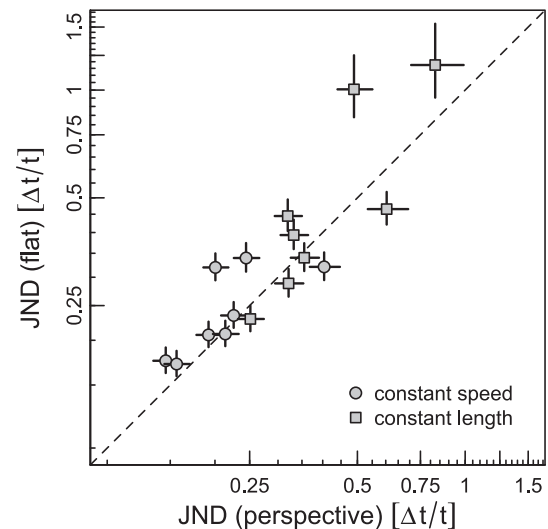

Figure 5. Discrimination thresholds expressed as Weber fractions (just-noticeable differences) in the perspective versus flat conditions. Each point represents one observer; only the flat conditions (with random vertical positions) are represented (the group of observers that performed the flat fixed condition did not perform any perspective condition). Vertical and horizontal bars are ±1 standard error. Although there are relatively more points above than below the equality line (indicating a tendency to lower discrimination performances in the flat condition), this effect was not significant (see main text).

significant results (all ps > 0.5) for all experiments and conditions. We therefore conclude that modeling the interaction between temporal duration and foreshortening as a simple additive contamination of the former by the latter provides an adequate description of observers' responses.

Finally, to allow for comparison of the performance in our task with that of other studies on time perception, we derived for each condition each observer's discrimination threshold. Thresholds were inferred from a probit analysis combining stimuli with all levels of foreshortening, and were defined as the proportion of duration increase that yielded a change from 0.5 to 0.75 in the probability of the second stimulus being judged as having a longer duration than the first. Figure 5 presents the individual thresholds in the flat (random vertical positions) condition as a function of the thresholds measured in the perspective condition. The average discrimination thresholds were 0.24 ($SD = 0.07$) and 0.22 ($SD = 0.08$) in the flat and perspective conditions of the constant-speed experiment, respectively; 0.54 ($SD = 0.35$) and 0.44 ($SD = 0.19$) in the flat and perspective conditions of the constant-length experiment; and 0.55 ($SD = 0.51$) in the experiment with a flat background and fixed positions. For the observers who performed both the flat and perspective conditions, we did not find any significant difference in the discrimination thresh-

olds—constant-speed: $t(7) = 1.11$, $p = 0.30$; constant length: $t(7) = 1.35$, $p = 0.22$.

Discussion

It is known that the perceived duration of a visual event in the second-to-subsecond range depends largely on the spatiotemporal properties (e.g., size, speed) of the stimulation. While traditionally these modulators of perceived duration have been studied in isolation, in this study we examined their joint effects on the perceived duration of an object moving at different viewing distances within a realistic perspective rendering of a 3-D environment. We find that when the fluctuations in the proximal (i.e., retinal) spatial properties of the moving object are arranged so that they respect realistic perspective rules, and when contextual information about viewing distance is present (perspective conditions; Figure 1b and e), their respective effects on the perceived duration of the moving object seem to cancel out, so that duration judgments are unaffected by changes in viewing distance. Conversely, when the same exact proximal stimuli are presented in the absence of additional distance cues (flat condition; Figure 1d), these effects result in a contraction of the perceived duration that is proportional to the foreshortening of the moving object. Taken together, our results indicate that when information about viewing distance is available, the brain compensates for the changes in the proximal properties of the visual events so as to preserve a constant perceived duration independent of the events' location in depth. This time-constancy phenomenon is the temporal equivalent of the well-known size constancy: Both phenomena refer to the fact that our perception (of the real size of objects and of their physical duration) remains constant regardless of the changes in the retinal input caused by variations in the viewing distance.

It has been proposed that judgments about space and time rely on a generalized magnitude processing mechanism (Cai & Connell, 2015; Walsh, 2003), an idea that is in agreement with the numerous reports of interaction between space and time in perception. However, this theory is not constrained by any specific timekeeping mechanism (Bueti & Walsh, 2009), and the brain mechanisms subserving time estimation remain largely debated (Finnerty, Shadlen, Jazayeri, Nobre, & Buonomano, 2015; Merchant, Harrington, & Meck, 2013). One view is that there are modality-specific mechanisms for perceived duration of sensory events (e.g., visual and auditory; Burr & Morrone, 2006; Burr, Tozzi, & Morrone, 2007; Johnston, Arnold, & Nishida, 2006; Yuasa & Yotsumoto, 2015) that could be based, for example, on mechanisms for coding speed (Gorea & Kim, 2015; Kaneko & Murakami, 2009) or temporal

frequency (Kanai et al., 2006; Linares & Gorea, 2015). A common characteristic of these studies is that they relate perceived duration to the strength of the neural response evoked by the stimulus (Eagleman & Pariyadath, 2009; Pariyadath & Eagleman, 2007). The present results, however, demonstrate that, provided that information about viewing distance is available, the brain can correctly compare durations of similar events even when these events are placed at different viewing distances and therefore evoke very different neural responses (i.e., they result in different proximal stimuli). Although it has been shown that viewing distance (in the form of linear perspective cues) can rescale the spatial extent of neural activity as early as the primary visual cortex (He, Mo, Wang, & Fang, 2015; Murray, Boyaci, & Kersten, 2006; Ni, Murray, & Horwitz, 2014), we argue that this modulation cannot fully account for the present results: While the degree of rescaling seems limited to a fraction of the objects' sizes, presumably due to feed-forward inputs to V1 (Murray et al., 2006), the spatial extent of the visual events that were perceived as having equal durations in the perspective (but not flat) conditions of our experiments could vary up to a factor of 5.5 (ratio between the farthest and nearest trajectory lengths and ball sizes). Our results therefore call for a more general mechanism (or brain system) for time perception that not only collects information from modality-specific brain areas to assess the duration of an event (Merchant et al., 2013) but also weights these inputs according to ecological contextual cues, such as viewing distance.

In sum, we have presented the first demonstration to our knowledge of perceptual constancy in the time domain. It suggests that the brain not only uses spatial aspects of the visual stimulation to time visual events in the subsecond-to-second range (Abe, 1935; Kanai et al., 2006; Kaneko & Murakami, 2009) but also adjusts the weights of these cues to normalize perceptual time across the viewing distance. In this view, one could speculate that time distortions induced by the physical features (such as speed and size) of stimuli whose duration is to be judged have in fact evolved phylogenetically so as to ensure time constancy in a 3-D world.

It should be pointed out that the present results lend themselves to alternative, though unlikely, interpretations. It is possible that even though observers were specifically instructed to base their duration judgments on duration itself and exclude any other stimulation cues, they nonetheless used such cues (namely speed in the constant-length experiment or trajectory length in the constant-speed experiment). If so, judgments based either on the trajectory length or on the speed of moving objects should have been more accurate in the presence of spatial references (the checkerboard plane in the perspective background; e.g., Bonnet, 1984; Gogel & McNulty, 1983). Such increased discrimination accuracy should have translated into a steeper

slope of the linear function representing the contribution of physical duration in the perspective condition (Figure 3, bottom right panel) with respect to the flat condition (Figure 3, bottom middle panel), reflecting larger perceptual intervals between the four physical durations. This was not the case: The respective slopes were not significantly different, revealing that discrimination accuracy (for comparisons made within the same level of foreshortening) was not influenced by the experimental condition (flat vs. perspective). Yet another alternative interpretation of the present results could be that the observed time constancy is but a consequence of speed constancy (Brown, 1931; Mckee & Smallman, 1998; Tozawa, 2008) or size constancy (Gregory, 1963), which equalized the perceived features of the stimulus across the different levels of foreshortening. Because we did not measure perceived speed, size, or traveled distance, this possibility cannot be excluded. Further research is needed to test this hypothesis and determine whether time constancy generalizes over more complex visual environments including factors such as self-motion (e.g., Combe & Wexler, 2010) and stereopsis.

The constancy of perceived time is an aspect of human perception that we take for granted, as it is hard to conceptualize how otherwise we could share the sense of time in a 3-D visual world. The present results point to a link between time and size/speed constancies. They suggest a common process that infers physical (or distal) properties of the world and drives perception of both space and time. A failure of such a shared process might be at the origin of deficits in both interval timing (Bonnot et al., 2011; Carroll, O'Donnell, Shekhar, & Hetrick, 2009) and size constancy (Kidd, 1964; Macdorman, Rivoire, Gallagher, & MacDorman, 1964; Weckowicz, Sommer, & Hall, 1958) reported in individuals with schizophrenia.

Keywords: time perception, perceptual constancy, visual perception, linear perspective, Ponzo illusion

Acknowledgments

Corresponding author: Matteo Lisi.
Email: matteo.lisi@parisdescartes.fr.
Address: Université Paris Descartes, Sorbonne Paris Cité, Paris, France.

Footnote

[1] We point out that because time (duration) does not exist by itself (i.e., it is always associated with an event), the notion of *proximal stimulus* is undefined in the time domain. For the current purposes we call proximal the *ensemble* of physical features that reach the senses and define the event. Accordingly, with respect to our experiments, the term *constancy* indicates that two different proximal events (different sizes and speeds but identical durations) are judged to be perceptually identical in the presence of 3-D cues but different in their absence. Hence, constancy according to this definition refers to the fact that *different proximal stimuli in a 2-D world* become *perceptually equivalent in a 3-D world.*

References

Abe, S. (1935). Experimental study of the corelation between time and space. *Tohoku Psychologica Folia, 3,* 53–68.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723, doi:10.1109/TAC.1974.1100705.

Aks, D. J., & Enns, J. T. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1467–1481, doi:10.1037/0096-1523.22.6.1467.

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

Bates, D. M. (in preparation). *lme4: Mixed-effects modeling with R.* New York: Springer. Retrieved from http://lme4.r-forge.r-project.org/lMMwR/lrgprt.pdf

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (R package version 1.1-7) [Computer software]. Retrieved from http://cran.r-project.org/package=lme4

Bonnet, C. (1984). Discrimination of velocities and mechanisms of motion perception. *Perception, 13*(3), 275–282, doi:10.1068/p130275.

Bonnot, O., de Montalembert, M., Kermarrec, S., Botbol, M., Walter, M., & Coulon, N. (2011). Are impairments of time perception in schizophrenia a neglected phenomenon? *Journal of Physiology*

Paris, *105*(4–6), 164–169, doi:10.1016/j.jphysparis. 2011.07.006.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century.

Brown, J. F. (1931). The visual perception of velocity. *Psychologische Forschung*, *14*(1), 199–232, doi:10. 1007/BF00403873.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1831–1840, doi:10.1098/rstb.2009.0028.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6*, 755–765, doi:10.1038/nrn1764.

Burr, D. C., & Morrone, M. C. (2006). Time perception: Space–time in the brain. *Current Biology*, *16*(5), R169–R171, doi:10.1016/j.cub.2006. 02.042.

Burr, D., Tozzi, A., & Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature Neuroscience, 10*(4), 423–425, doi:10.1038/nn1874.

Cai, Z. G., & Connell, L. (2015). Space-time interdependence: Evidence against asymmetric mapping between time and space. *Cognition, 136*, 268–281.

Carroll, C. A., O'Donnell, B. F., Shekhar, A., & Hetrick, W. P. (2009). Timing dysfunctions in schizophrenia span from millisecond to several-second durations. *Brain and Cognition*, *70*(2), 181–190, doi:10.1016/j.bandc.2009.02.001.

Combe, E., & Wexler, M. (2010). Observer movement and size constancy. *Psychological Science*, *21*(5), 667–675, doi:10.1177/0956797610367753.

Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology, 18*, 131–136, doi:10.1016/j.conb.2008.06.002.

Eagleman, D. M., & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1841–1851, doi:10. 1098/rstb.2009.0026.

Fineman, M. B. (1981). Complexity of context and orientation of figure in the corridor illusion. *Perceptual and Motor Skills*, *53*(1), 11–14, doi:10. 2466/pms.1981.53.1.11.

Finnerty, G. T., Shadlen, M. N., Jazayeri, M., Nobre, A. C., & Buonomano, D. V. (2015). Time in cortical circuits. *The Journal of Neuroscience*, *35*(41), 13912–13916, doi:10.1523/JNEUROSCI. 2654-15.2015.

Fraisse, P. (1963). *The psychology of time*. New York: Harper and Row.

Gerardin, P., Devinck, F., Dojat, M., & Knoblauch, K. (2014). Contributions of contour frequency, amplitude, and luminance to the watercolor effect estimated by conjoint measurement. *Journal of Vision*, *14*(4):9, 1–14, doi:10.1167/14.4.9. [PubMed] [Article]

Gogel, W. C., & McNulty, P. (1983). Perceived velocity as a function of reference mark density. *Scandinavian Journal of Psychology*, *24*(4), 257–265, doi:10. 1111/j.1467-9450.1983.tb00499.x.

Gorea, A. (2011). Ticks per thought or thoughts per tick? A selective review of time perception with hints on future research. *Journal of Physiology Paris*, *105*(4–6), 153–163, doi:10.1016/j.jphysparis. 2011.09.008.

Gorea, A., & Hau, J. (2013). Time in perspective. *Psychological Science*, *24*(8), 1477–1486, doi:10. 1177/0956797612473486.

Gorea, A., & Kim, J. (2015). Time dilates more with apparent than with physical speed. *Journal of Vision*, *15*(1):7, 1–11, doi:10.1167/15.1.7. [PubMed] [Article]

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: John Wiley.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, *199*(4894), 678–680.

Hass, J., & Durstewitz, D. (2014). Neurocomputational models of time perception. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (pp. 49–71). New York: Springer, doi:10.1007/ 978-1-4939-1782-2_4.

He, D., Mo, C., Wang, Y., & Fang, F. (2015). Position shifts of fMRI-based population receptive fields in human visual cortex induced by Ponzo illusion. *Experimental Brain Research*, *233*(12), 3535–3541, doi:10.1007/s00221-015-4425-3.

Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science*, *19*(2), 196–204, doi:10.1111/ j.1467-9280.2008.02067.x.

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *The American Journal of Psychology*, *54*(1), 21–37, doi: 10.2307/1417790.

James, W. (1890). The perception of time. In *The*

principles of psychology (pp. PP–PP). New York: H. Holt & Co.

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

Kanai, R., Paffen, C. L. E., Hogendoorn, H., & Verstraten, F. A. J. (2006). Time dilation in dynamic visual display. *Journal of Vision*, *6*(12):8, 1421–1430, doi:10.1167/6.12.8. [PubMed] [Article]

Kaneko, S., & Murakami, I. (2009). Perceived duration of visual motion increases with speed. *Journal of Vision*, *9*(7):14, 1–12, doi:10.1167/9.7.14. [PubMed] [Article]

Kidd, A. H. (1964). Monocular distance perception in schizophrenics. *Journal of Abnormal and Social Psychology*, *68*(1), 100–103, doi:10.1037/h0045150.

Knoblauch, K., & Maloney, L. (2012). *Modeling psychophysical data in R*. New York: Springer, doi:10.1007/978-1-4614-4475-6.

Linares, D., & Gorea, A. (2015). Temporal frequency of events rather than speed dilates perceived duration of moving objects. *Scientific Reports, 5*, 1–9, doi:10.1038/srep08825.

Luce, D. R., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, *1*(1), 1–27, doi:10.1016/0022-2496(64)90015-X.

Macdorman, C. F., Rivoire, J. L., Gallagher, P. J., & MacDorman, C. F. (1964). Size constancy of adolescent schizophrenics. *Journal of Abnormal and Social Psychology*, *69*(3), 258–263, doi:10.1037/h0042347.

Mcgraw, P. V., Roach, N. W., Badcock, D. R., & Whitaker, D. (2012). Size-induced distortions in perceptual maps of visual space. *Journal of Vision*, *12*(4):8, 1–14, doi:10.1167/12.4.8. [PubMed] [Article]

Mckee, S. P., & Smallman, H. S. (1998). Size and speed constancy. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy* (pp. 373–408). Cambridge, UK: Cambridge University Press.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336, doi:10.1146/annurev-neuro-062012-170349.

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, *9*(3), 429–434, doi:10.1038/nn1641.

Ni, A. M., Murray, S. O., & Horwitz, G. D. (2014).

Object-centered shifts of receptive field positions in monkey primary visual cortex. *Current Biology*, *24*(14), 1653–1658, doi:10.1016/j.cub.2014.06.003.

Pariyadath, V., & Eagleman, D. M. (2007). The effect of predictability on subjective duration. *PLoS One*, *2*(11), e1264, doi:10.1371/journal.pone.0001264.

Ponzo, M. (1911). Intorno ad alcune illusioni nel campo delle sensazioni tattili sull'illusione di Aristotele e fenomeni analoghi [On some tactile illusions, Aristotle's illusion, and similar phenomena]. *Archive für die Gesamte Psychologie*, *16*, 307–345.

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, 1–10, doi:10.1167/14.5.17. [PubMed] [Article]

Rogers, M., Knoblauch, K., & Franklin, A. (2016). Maximum likelihood conjoint measurement of lightness and chroma. *Journal of the Optical Society of America A*, *33*(3), A184–A193, doi:10.1364/JOSAA.33.00A184.

Tozawa, J. (2008). Speed constancy and the perception of distance. *Perception*, *37*(1), 3–21, doi:10.1068/p5585.

van Wassenhove, V. (2009). Minding time in an amodal representational space. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1815–1830, doi:10.1098/rstb.2009.0023.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488, doi:10.1016/j.tics.2003.09.002.

Weckowicz, T. E., Sommer, R., & Hall, R. (1958). Distance constancy in schizophrenic patients. *The British Journal of Psychiatry*, *104*(437), 1174–1182, doi:10.1192/bjp.104.437.1174.

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

Yuasa, K., & Yotsumoto, Y. (2015). Opposite distortions in interval timing perception for visual and auditory stimuli with temporal modulations. *PLoS ONE*, *10*(8), e0135646, doi:10.1371/journal.pone.0135646.