

TIME-ORDER EFFECTS IN TIMED BRIGHTNESS AND SIZE DISCRIMINATION OF PAIRED VISUAL STIMULI

Geoffrey R. Patching, Mats P. Englund and Åke Hellström.

Department of Psychology, Stockholm University, SE-106 91 Stockholm, Sweden.

Email: grp@psychology.su.se

Abstract

An aim of the present paper is to examine changes in the magnitude and direction of time-order effects with changes in the physical magnitude of the stimuli, overall levels of performance, and with changes in load processing capacity. Increments and decrements in performance are gauged by means of the steepness of psychometric and chronometric functions. Load processing capacity is assessed by dividing integrated hazard functions for stimuli varying congruently along the dimensions of both size and brightness by the sum of those for stimuli varying along either dimension alone. Taken altogether the findings suggest that time-order effects in timed brightness and size discrimination are perceptual, and the relative sensory weighting of stimulus magnitudes related to the alertness of participants and their capacity to process the stimuli. Implications of these findings for common random walk and diffusion models of sensory discrimination are discussed.

Since the days of Gustav Fechner (1801-1887) a vast number of studies have shown that when two stimuli are presented successively for magnitude comparison one stimulus is usually over- or underestimated relative to the other; termed, time-order effect (TOE). Moreover, the magnitude and direction of such effects are known to vary systematically with changes in the physical magnitude and configuration of the stimuli (for a review see Hellström, 1985). The precise nature of processes underlying the comparison and discrimination of stimuli remains a matter of debate, although there is little doubt that TOEs reflect an important role for attention in the comparison process.

Numerous theories have been proposed to account for systematic asymmetries in paired stimulus comparisons, but still the most comprehensive account to date is Hellström's (1979) sensation weighting (SW) model. Hellström's SW model posits a weighting of activation inspired by each stimulus event and a reference level (ReL) based on generic information. Formally the SW model is defined as

$$d = k \{ [s_1 \psi_1 + (1 - s_1) \psi_{r1}] - [s_2 \psi_2 + (1 - s_2) \psi_{r2}] \}, \quad 1$$

where d is the subjective difference between the compared stimuli, k is a scale constant, s_1 and s_2 are weighting coefficients (1 and 2 indicate the temporal order or spatial position of the stimuli; i.e., first or left=1, second or right=2), ψ_1 and ψ_2 are the sensation magnitudes of the stimuli and ψ_{r1} and ψ_{r2} are the sensation magnitudes of the ReLs.

Hellström (2007) further generalized his SW model to encompass paired multidimensional stimuli. In this case, with P dimensions and Minkowski constant m (=2 for Euclidean space), and assuming $\psi_{r1i} = \psi_{r2i} = \psi_{ri}$, Equation 1 becomes

$$d_{12} = \left\{ \sum_{i=1}^P [s_1(\psi_{1i} - \psi_{ri}) - s_2(\psi_{2i} - \psi_{ri})]^m \right\}^{1/m}. \quad 2$$

Essentially, the SW model is based on the notion that stimulus specific information and generic information are weighted together in the comparison process to optimize signal-to-noise ratios and maximize stimulus discrimination. This process is thought to give rise to systematic asymmetries in paired stimulus comparisons but generally improve the discriminability of stimuli by reducing the variability of the estimates and so is considered to be evolutionarily advantageous. To date, however, TOE studies have focused primarily on the dependent measure of response probability, although response time is an additional dependent variable of considerable interest in development of an understanding of time-order effects (Jamieson & Petrusic, 1975).

Common sequential sampling (Link & Heath, 1975) and diffusion models (Ratcliff, 1978) directly tie choice proportions to response times through modelling the choice process in terms of a random walk. According to these models the process of comparison consists of the accumulation of noisy information about the difference between stimulus values over time, until either of two boundaries (A or $-A$) is reached. Discrimination time is defined as the time from the start of the process until one boundary is reached and response probability determined by the likelihood of crossing either boundary.

Of relevance is that Vandekerckhove and Tuerlinckx (2008) extend applicability of the Ratcliff Diffusion Model (RDM) by adopting a design matrix approach, similar to standard techniques of general linear modelling, to impose restrictions on parameters across conditions. According to this approach each free parameter (generically denoted θ) of the RDM can be decomposed into a weighted linear combination of N known predictor values

$$\theta_{(i)} = \sum_{j=1}^N \Phi_{(ji)} \theta_{(j)}^* \tag{3}$$

where $\Phi_{(ji)}$ is the value of the j th predictor in condition i (note the asterisk denotes a design parameter). Consequently, substantial restrictions can be placed on parameters to examine the role of attention and that of attention capacity in the comparison and discrimination of stimuli.

Various notions of processing capacity exist but what is perhaps one of the most explicit conceptualizations comes from Townsend and colleagues (cf. Townsend & Wenger, 2004). Central to their approach is the integrated hazard function which provides a measure of the total amount of energy needed to complete processing by some critical time t . Of particular interest, in the present context, is their definition of processing capacity as given by the *capacity coefficient* $C(t)$. $C(t)$ is derived by dividing the integrated hazard function for responses to a particular type of multidimensional signal, by the sum of the integrated hazard functions for responses to each dimension presented alone. In brief, the basic idea is that each dimension inspires a stream of activation, that may be processed as efficiently [$C(t) = 1$], less efficiently [$C(t) < 1$] or more efficiently [$C(t) > 1$], when presented as part of a two-dimensional stimulus than when presented alone, and that accumulate information over time in favour of one response over the other, until a negative or positive response criterion is met that determines the choice response.

Random walk, drift diffusion, accumulation and recruitment models assume that stimulus comparison merely involves the simple subtraction of noisy activation inspired by each choice alternative. They are based on notions of symmetry belying systematic asymmetries in the human comparison of stimuli, but relating participant's capacity to process the stimuli to the sensory weighting of stimulus magnitudes promises to shed new light on the role of attention in paired stimulus comparison. On these grounds, the present experiment extended the standard method of paired comparisons by measurement of both response probability and response time and by inclusion of both unidimensional paired stimuli varying uniquely along the single dimension of luminance or size and paired bidimensional stimuli

varying congruently along the dimensions of both luminance and size, whereby participants could make their choice response based on the relative brightness OR size of each stimulus.

Experimental method

Participants. 40 participants took part in the experiment of whom 37 were women and 3 men between the ages 19 to 44 (mean 22 yrs). All but one claimed to be right handed.

Stimuli/design. The stimuli were successively presented paired round-shaped visual stimuli. For visual stimuli varying along the single dimension of luminance, 5 luminance levels were combined to create 6 paired stimuli of, 4.1 - 4.7, 4.4 - 5.0, 4.7 - 5.3, 4.7 - 4.1, 5.0 - 4.4 5.3 - 4.7 cd/m^2 , with a constant intrapair difference of $\pm 0.6 \text{ cd/m}^2$ and 3 mean average magnitude levels of 4.4, 4.7, and 5.0 cd/m^2 . For each of these 6 stimuli the size of the stimuli was held constant at one of 5 levels from 5.5 to 6.3 mm diameter, in steps of 0.2 mm.

For visual stimuli varying along the single dimension of size, 5 size levels were similarly combined to create 6 paired stimuli of, 5.5 - 5.9, 5.7 - 6.1, 5.9 - 6.3, 5.9 - 5.5, 6.1 - 5.7 6.3 - 5.9 mm diameter, with constant intrapair difference of $\pm 0.4 \text{ mm}$ and 3 mean average magnitude levels of, 5.7, 5.9, and 6.1 mm diameter. For each of these 6 stimuli the luminance of the stimuli was held constant at one of 5 levels from 4.1 to 5.3 cd/m^2 in steps of 0.3 cd/m^2 .

For the stimuli varying along both dimensions of luminance and size a constant intrapair difference of 0 was also used. In this respect, the values of luminance and size were combined factorially to create 9 stimuli of, 4.1/5.5 - 4.7/5.9, 4.4/5.7 - 5.0/6.1, 4.7/5.9 - 5.3/6.3, 4.4/5.7 - 4.4/5.7, 4.7/5.9 - 4.7/5.9, 5.0/6.1 - 5.0/6.1, 4.7/5.9 - 4.1/5.5, 5.0/6.1 - 4.4/5.7, 5.3/6.3 - 4.7/5.9 $\text{cd/m}^2 / \text{mm}$ diameter respectively, with constant intrapair differences of $\pm 0.6 \text{ cd/m}^2$, 0, and $\pm 0.4 \text{ mm}$ diameter, and 3 mean average magnitude levels of, 4.4/5.7, 4.7/5.9, and 5.0/6.1 $\text{cd/m}^2 / \text{mm}$ diameter respectively. Here incongruent combinations of difference in luminance and in size were avoided, and in every case the ISI was held constant at 800 msec.

Each pair of stimuli differing along the single dimension of luminance or size was presented 16 times while the magnitude of the alternate dimension was held constant at one of five levels giving 80 judgments for each stimulus difference, and each pair of stimuli differing congruently along the dimensions of luminance and size were likewise presented 80 times.

In each experiment participants took part in four sessions, one on each of 4 days. Within each session, each stimulus pair was shown in pseudo-randomly constructed cycles of 105 items, comprising 60 stimuli varying along the dimension of either luminance or size, 30 stimuli varying congruently along the dimensions of both luminance and size, and 15 stimuli in which the first and second stimuli shared the same stimulus values. The first 25 trials were chosen at random and designated as practice trials. After presentation of the 25 practice trials, each participant was required to complete a further 420 trials in each session. New random orders were used for each participant and each session.

Half of the participants were instructed to press a left response key if they perceived the first stimulus to be larger or brighter than the second, or a right key if they perceived the second stimulus to be larger or brighter than the first. The other 20 participants were instructed to use the reverse stimulus-response assignment.

Procedure. At the beginning of each session participants were presented with written instructions on the video monitor. The importance of responding as quickly and as accurately as possible was stressed, and all participants were informed that on no occasion would a stimulus pair comprise a brighter but smaller, or dimmer but larger stimulus difference. The inter-trial interval was 3000 msec. On the average, participants took 40 minutes to complete each session.

Results

All premature responses less than 200 msec and all misses greater than 2000 msec were discarded from further analysis. This resulted in the removal of just 0.68% of responses.

Table 1. Mean log-odds ratios (logit p) and mean signed response speed ($SRS \times 10^{-2}$) for the mean average magnitude of the 1st and 2nd stimuli in each stimulus pair.

Imperative dimension(s)								
Brightness			Size			Brightness / Size		
Mean luminance (cd/m ²)	Logit p	$SRS \times 10^{-2}$	Mean size (mm)	Logit p	$SRS \times 10^{-2}$	Mean luminance / size	Logit p	$SRS \times 10^{-2}$
4.4	-0.103	-1.55	5.7	-0.122	-1.92	4.4 / 5.7	-0.024	-1.53
4.7	-0.181	-1.87	5.9	-0.171	-1.69	4.7 / 5.9	-0.184	-2.07
5.0	-0.386	-3.04	6.1	-0.174	-1.44	5.0 / 6.1	-0.304	-2.12

On the basis of arguments made by Patching, Englund, & Hellström, (2008) systematic asymmetries in brightness and size discrimination are examined by way of binary response probability scaled in terms of log-odds ratios (logit p), as well as by signed response speed (SRS ; i.e., the inverse of response time with the sign of the judged difference). Examination of the relationship between logit p and SRS by way of linear regression revealed group averaged values of adjusted R-square $> .94$, indicating that they are linearly related. SRS was then averaged over trials for each of the three pair averaged magnitude levels and for each participant separately. In addition, mean log odds ratios were computed for each pair averaged magnitude level. Summaries of these data are presented in Table 1. By convention, a negative effect refers to an underestimation of the 1st as compared to 2nd stimulus.

Sensation weighing. Following the procedures described by Hellström (1979) whereby the following preliminary SW model, $d = W_1\psi_1 - W_2\psi_2 + U$, was first fit to empirical scale values of stimulus difference, where k is a scale constant and $W_1 = ks_1$ and $W_2 = ks_2$, weights W_1 , W_2 were estimated by logistic regression of binary responses and by standard regression of SRS on the standardized log values (generically denoted Φ) of first and secondly presented unidimensional stimuli. For the multidimensional stimuli the procedures as discussed by Hellström (2007) were followed, whereby logit p and SRS were regressed on the sum of the standardized log values of luminance and size, where $\Phi_1 = [\Phi_{1luminance} + \Phi_{1size}]$ and $\Phi_2 = [\Phi_{2luminance} + \Phi_{2size}]$. Subsequently, a weightings differential ($WD\%$) was calculated separately for each participant for the judged dimensions by dividing the difference between the weightings by their mean, $WD\% = 200(W_1 - W_2)/(W_1 + W_2)$, which permits comparison of relative weighting asymmetries between participants and between relevant conditions.

On the basis of the SW model $WD\%$ indicates the direction of asymmetry in terms of the weighting of each stimulus, and the relation $WD\% \neq 0$ was tested by (two tailed) one sample t -tests. The results of testing the relation $WD\% \neq 0$ are detailed in Table 2. Further analysis of the influence of response assignment on the magnitude and direction of $WD\%$ failed to reveal any statistically reliable effects (all $ps > .05$).

In support of the SW model regression of logit p and SRS on $[\Phi_1 - \Phi_2]$ and $[\Phi_1 + \Phi_2]$ revealed significant coefficients for $[\Phi_1 - \Phi_2]$; all $ps < .01$. In addition, regression of binary responses and SRS on $[\Phi_1 + \Phi_2]$ revealed statistically significant coefficients for stimuli differing only in brightness, as determined on the basis of logit p , $t(474) = 1.61$, $p = .054$, $\beta = 0.18$, and for stimuli differing along the dimensions of both brightness and size, as determined on the basis of logit p , $t(474) = 2.54$, $p = .006$, $\beta = 0.12$, and as determined on the basis of SRS ,

Table 2. Mean weightings W_1 / W_2 along with the weightings differential $WD\%$.

Judged dimension	Logit p			SRS		
	Brightness	Size	Brightness / Size	Brightness	Size	Brightness / Size
W_1 / W_2	0.598 / 0.917	1.38 / 1.45	0.931 / 1.13	0.0435 / 0.0555	0.0867 / 0.0831	0.0522 / 0.0538
$WD\%$	-42.1	-4.95	-19.31	-24.24	4.24	-3.02
$WD\% \neq 0$	$p < .001$	n.s.	$p < .001$	$p < .001$	n.s.	n.s.

Note: n.s., not significant ($p > .05$).

$t(474) = 2.07, p = .02, \beta = .004$. The implication is, therefore, that (since the TOE changes with stimulus magnitude) the weights W_1 and W_2 are different.

It was of further interest to examine changes in the direction and magnitude of $WD\%$ with changes in the speed of each participant's responses. To this end each participant's data were subdivided into RT quartiles from fastest to slowest, and weights W_1 and W_2 estimated for each quartile. By and large, performance was found to decrease and the magnitude of $WD\%$ increase, becoming more negative, as the speed of responses decreased. The results of this analysis are detailed in Table 3.

Diffusion model analysis. Diffusion model analysis was conducted using the Diffusion Model Analysis Toolbox (DMAT; Vandkerckhove & Tuerlinckx, 2008). Six nested models of increasing complexity were fit to the data: 1) All parameters constrained to be constant across all stimulus conditions. 2) Regression of drift rates (v) on weighted stimulus values [i.e., $v_{(i)} = v_{(1)}^* + W_1\Phi_{1i}v_{(2)}^* - W_2\Phi_{2i}v_{(3)}^*$]. 3) Regression of drift rates on weighted stimulus values, and start position of the diffusion process free to vary over conditions. 4) Drift rates free to vary and start position free to vary over conditions. 5) Drift rates, start position and variability of the start position free to vary. 6) All 7 parameters free to vary over conditions. On the average, this analysis revealed no significant improvement in fit beyond Model 4.

The capacity coefficient $C(t)$. Following the procedures laid down by Wenger and Townsend (2000), values of $C(t)$ were computed for each participant in 10 ms RT bins, for first greatest responses, $C(t)_1$, and separately for second greatest responses, $C(t)_2$, excluding response times to paired stimuli sharing the same physical values of luminance and size. In the main, values of $C(t)$ were found to range between 1 and 0.5, indicative of the independent processing of the dimensions of brightness and size (Townsend & Wenger, 2004). However, for fast responses values of $C(t)$ exceeded 1, and in some case fell below 0.25. Specifically, for first greatest responses, 27 of the 40 participants tested produced values of $C(t)_1 > 1$, and 12 produced values of $C(t)_1 < .25$, three of whom produced values of both $C(t)_1 > 1$ and $C(t)_1 < .25$. For second greatest responses 31 of the 40 participants tested produced values of $C(t)_2 > 1$, and 5 produced values of $C(t)_2 < .25$, two of whom produced values of both $C(t)_2 > 1$ and $C(t)_2 < .25$. In this respect, there is evidence for both positive and negative dependencies in timed brightness and size discrimination of paired visual stimuli.

Load processing capacity / sensory weighting of stimulus magnitudes. Pearson correlation coefficients (two tailed) were determined to investigate relations between $C(t)_1$ and $C(t)_2$ and weightings W_1, W_2 . For this analysis all values of $C(t) > 3$ were removed, since in a small number of cases values of $C(t)$ were overinflated due to there being fast RTs to the bidimensional stimuli but no correspondingly fast RTS to the unidimensional stimuli.

Table 3. Mean weightings W_1/W_2 and the weightings differential $WD\%$ for each RT Quartile.

RT Quartile	Logit p				SRS				
	Fast	2nd	3rd	Slow	Fast	2nd	3rd	Slow	
Brightness					Brightness				
W_1	1.39	0.790	0.563	0.377	0.0747	0.0476	0.0324	0.0176	
W_2	1.95	1.50	1.03	0.589	0.0749	0.0650	0.0525	0.0270	
$WD\%$	-33.54	-62.01	-58.63	-43.89	-0.2674	-30.91	-47.35	-42.15	
Size					Size				
W_1	3.19	2.10	1.45	0.917	0.130	0.0915	0.0706	0.0397	
W_2	3.30	2.24	1.53	1.10	0.119	0.0961	0.0725	0.0456	
$WD\%$	-3.390	-6.452	-5.369	-18.15	8.835	-4.904	-2.655	-13.84	
Brightness / size					Brightness / size				
W_1	1.92	1.23	0.945	0.592	0.0761	0.0525	0.0376	0.0235	
W_2	1.77	1.49	1.31	0.815	0.0634	0.0595	0.0494	0.0319	
$WD\%$	8.130	-19.12	-32.37	-31.70	18.21	-12.50	-27.13	-30.32	

$C(t)_1$ and $C(t)_2$ were then averaged over time bins, t , for each participant separately. The resultant mean values of C_1 were then correlated with W_1 , and separately C_2 with W_2 .

This analysis revealed a significant correlation between C_1 and W_1 as determined by logit p , $r = 0.56$, $p < .001$, and significant correlations between C_2 and W_2 as determined by logit p , $r = 0.67$, $p < .001$, and as determined by SRS, $r = 0.52$, $p < .001$. These findings indicate that changes in load processing capacity between participants are moderately related to the sensory weighting of stimulus magnitudes.

Conclusion

The findings show characteristic time order effects in timed brightness and size discrimination of visual stimuli. In the main, the magnitude of time-order effects was found to increase with increased visual magnitude and, in gross summary, performance was found to decrease with increased asymmetry in the differential weighting of stimulus magnitudes (cf. Hellström, 1979). Moreover, for first and second greatest responses the magnitude of sensory weightings was found to be moderately related to load processing capacity, as assessed in terms of the capacity coefficient $C(t)$ for the two choice alternatives. Taken together these findings suggest that time-order effects in timed brightness and size discrimination are perceptual, and the differential sensory weighting of stimulus magnitudes related to the alertness of participants and their capacity to process the stimuli. Finally, joint analysis of response probability and response times using the RDM supports the view that systematic asymmetries in the method of paired comparisons arise as a result of changes in the rate of drift diffusion and prior beliefs about the stimuli, which implies a role for both attention and generic information in the comparison process.

Acknowledgments

This research was supported by a grant from the Swedish Research Council.

References

- Hellström, Å. (1979). Time errors and differential sensation weighting. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 460-477.
- Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, 97, 35-61.
- Hellström, Å. (2007). *Temporal asymmetry and "magnet effect" in similarity and discrimination of prototypical and nonprototypical stimuli: Consequences of differential sensation weighing*. Fechner Day, 2007. Tokyo, Japan.
- Jamieson, D.G., & Petrusic, W.M. (1975). Presentation order effects in duration discrimination. *Perception & Psychophysics*, 17, 197-202.
- Link, S.W., & Heath, R.A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77-105.
- Patching, G.R., Englund, M.P., Hellström, Å. (2008). *Time- and space-order effects in timed brightness discrimination of paired visual stimuli*. Fechner Day, 2008. Toronto, Canada.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Townsend, J. T., & Wenger. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003-1035.
- Vandkerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, 40, 61-72.
- Wenger, J.W. & Townsend, J.T. (2000). Basic response time tools for studying general processing capacity in attention, perception and cognition. *The Journal of General Psychology*, 127, 67-99.