Time- and Space-Order Effects in Timed Discrimination of

Brightness and Size of Paired Visual Stimuli

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

Stockholm University

Author Note

Geoffrey R. Patching, Mats P. Englund, and Åke Hellström, Department of Psychology, Stockholm University, Sweden.

Geoffrey R. Patching is now at the Department of Psychology, Lund University, Lund, Sweden.

Correspondence regarding this work may be addressed to G. R. Patching, E-mail: Geoffrey.Patching@psychology.lu.se, at the Department of Psychology, Lund University, SE-221 00 Lund, Sweden.

Abstract

Despite the importance of both response probability and response time for testing models of choice there is a dearth of chronometric studies examining systematic asymmetries that occur over time- and space-orders in the method of paired comparisons. In this study, systematic asymmetries in discriminating the magnitude of paired visual stimuli are examined by way of log-odds ratios of binary responses as well as by signed response speed. Hierarchical Bayesian modeling is used to map response probabilities and response speed onto constituent psychological process, and processing capacity is also assessed using response time distribution hazard functions. The findings include characteristic order effects that change systematically in magnitude and direction with changes in the magnitude and separation of the stimuli. After Hellström (1979, 2000), Sensation Weighting (SW) model analyses show that such order effects are reflected in the weighted accumulation of noisy information about the difference between stimulus values over time, and interindividual differences in weightings asymmetries are related to the relative processing capacity of participants. An account of sensation weighting based on the use of reference level information and maximization of signal-to-noise ratios is posited, which finds support from theoretically driven analyses of behavioral data.

Keywords: Time-order effects, space-order effects, adaptive perception, wave theory, diffusion model, processing capacity.

Time- and Space-Order Effects in Timed Discrimination of

Brightness and Size of Paired Visual Stimuli

Experimental work examining human discrimination of paired visual stimuli is

fundamental to development of an understanding of how we compare stimuli and how we

perceive differences between them. In the present study the focus is on paired comparisons of

visual magnitude (i.e., either size or brightness or both). Situations are considered where

people are required to compare visual magnitudes by making a binary choice, and where

choices are timed. An objective is to examine those perceptual-cognitive mechanisms

involved in the comparison and discrimination of visual stimulus magnitudes, where sensation

weighting operates to facilitate discrimination of small changes in the magnitude difference

between paired stimuli. One major advance is that systematic asymmetries in choosing

between paired stimuli are jointly examined by way of both choice probabilities and response

times, enabling theoretical assessment of the relative speed of perceptual accumulation

processes in the comparison and discrimination of paired stimuli. A further advance is that

relations between interindividual differences in the magnitude and direction of such

asymmetries and participants' relative efficiency of stimulus processing are assessed, and

discussed in regard to discrimination optimization.

Since Fechner (1860) it has been known that when two stimuli are presented for

comparison, one stimulus is usually over- or underestimated relative to the other. Moreover, it

is known that the magnitude and direction of this asymmetry varies systematically with

changes in the physical magnitude of the stimuli and with changes in the temporal or spatial

separation between stimuli (see Guilford, 1954, chap. 12; Hellström, 1985; Peak, 1940, for

comprehensive reviews of this research). The term *time-order effect* (TOE) is used to refer to

systematic asymmetries in paired comparisons of stimuli separated by a time interval, and the

term *space-order effect* (SOE) to systematic asymmetries in paired comparisons of stimuli

separated spatially.

On each trial of a paired comparison task, two stimuli separated by a temporal or spatial interval are presented for comparison by an experimental participant. In the simplest situation, the participant is requested to choose one of the paired stimuli on the basis of some given attribute. Over the course of one or more experimental sessions a predefined range of stimulus values is presented, so that over this range the probability of choosing one stimulus over the other rises monotonically from zero to one. The function described is known as the *psychometric function*, and it is central to the study of psychophysics (Fechner, 1860; Guilford, 1954; Klein, 2001). Classically, the physical magnitude on the continuum of the variable attribute of one stimulus that has a probability of 0.5 of being chosen when compared to the other is termed the *point of subjective equality* (PSE) and, where the given attribute of the first or left stimulus, the Standard (St), is held constant, subtracting the magnitude of the St from the PSE yields the constant error (CE = PSE - St). [1] The TOE and the SOE are special cases of the CE, and using these measures extends assessment of order effects independently of the psychophysical scale (Guilford, 1954; Hellström, 1979, 2000).

In the present study, the TOE and SOE are assessed in terms of the logit (log-odds ratio) of $P,$ the proportion of 'first greater' or 'left greater' responses: logit $P = \log_e [P / (1 - P)]$. The logistic function is favored by statisticians for analysis of binary data (Collett, 2002) and has been employed widely by researchers in development of sequential sampling and diffusion models of the psychometric function (Link, 1975, 1978, 1992; Link & Heath, 1975; Palmer, Huk, & Shadlen, 2005). By convention, all measures of the TOE and SOE are such that a positive effect is taken to refer to an overestimation of the first (or left) stimulus as compared to the second (or right) stimulus, and a negative effect to an underestimation of the first (or left) as compared to the second (or right) stimulus.

Focusing on response probability, TOEs are known to arise in comparative judgments

of auditory loudness (Hellström, 1979; John, 1975), auditory pitch (Tresselt, 1948), auditory duration (Woodrow & Stott, 1936), visual duration (Jamieson & Petrusic, 1975b), duration of empty intervals (Woodrow, 1935), heaviness of weights (Michels & Helson, 1954), line length (Tresselt, 1944a), the visual area of squares, triangles and circles (Inomata, 1959; Nachmias, 2006; Postman, 1947; Tresselt, 1944a, 1994b), and visual brightness (Bentley, 1899; Holzman, 1954; Kreezer, 1938; Maeda, 1959, 1960; Ono, 1949). Generally speaking, TOEs have been found to be positive when paired stimuli are separated by a short temporal interstimulus interval (ISI) and have low levels of stimulus intensity, and negative for longer ISIs and high levels of stimulus intensity (Bartlett, 1939; Needham, 1934, 1935; Woodrow, 1933). However, the precise magnitude and direction of TOEs is known to depend on the modality and type of stimulus judgment (Postman, 1946), as well as on the duration of the stimuli and the ISI (Hellström, 1979, 2003) and may be estimated, using different logistic parameters for the two presentation orders, to be larger than otherwise thought (Lapid, Ulrich, & Rammsayer, 2008; Ulrich & Vorberg, 2009; Yeshurun, Carraso, & Maloney, 2008).

In discrimination of brightness predominantly negative TOEs have been obtained (Bentley, 1899; Holzman, 1954; Kreezer, 1938), although there is evidence of positive TOEs when the two visual stimuli have been separated by short temporal intervals (Maeda, 1959). In discrimination of visual area (here termed, size discrimination), TOEs have been found to be negative for small stimuli (Nachmias, 2006; Postman 1947; Tresselt, 1944a) and positive for large stimuli (Tresselt, 1944a).

Somewhat fewer psychophysical studies have examined SOEs, but they are known to arise in human comparative judgments of the heaviness of lifted weights (Woodruff, Jennings, & Rico, 1975), line length (Brown, 1953; Hellström, 2003; Vidotto, Vicario, & Tomat, 1996), visual brightness (Kellogg, 1931), and stimulus size (Charles, Sahraie, & McGeorge, 2007). For instance, Kellogg (1931) reported a negative asymmetry in split-disk brightness

discrimination, in that most participants chose the luminance on the right-half of the disc as being brighter more frequently than the left despite equally balanced changes in the luminance difference between the two, left and right, halves of the disc. In regard to size discrimination, Charles et al. (2007) found that in paired comparisons of the horizontal width of an ellipse to a standard circle size, participants tended to overestimate the width of the ellipse when presented to the left, as compared to the right, of the standard circle. However, this obtained only when the stimuli were separated spatially by 10^0. When the same stimuli were spatially separated by 5^0 or 20^0 the effect was not statistically reliable. In sum, there is less evidence of SOEs than TOEs in comparisons of brightness and size, and in both cases studies have yet to determine the precise extent to which such order effects change systematically in magnitude and direction with changes in stimulus magnitude and with changes in the spatial and temporal separation between paired stimuli.

Researchers all too often dismiss TOEs and SOEs as methodological artifacts (Stevens, 1957) or assume that such effects arise merely as the result of bias (Green & Swets, 1966). Response bias may arise as a result of prejudiced decision criteria (Allan, 1977), or as a result of a tendency to implicitly generate a dichotomous verbal response depending on the absolute magnitude of the stimuli (John, 1975). Focusing on negative TOEs, others envisage some kind of retention loss (Köhler, 1923; Lauenstein, 1933; Link, 1992) such that the activation inspired by one stimulus is compared to a weaker mental representation of the other. In regard to SOEs, similar appeal has been made by reference to known functional asymmetries in neural anatomy (Mattingley, Bradshaw, Nettleton, & Bradshaw, 1994), noncentral fixation, and spatial scanning effects (Masin & Agostini, 1991).

On these grounds TOEs and SOEs are regularly ignored except by the practice of averaging over the two presentation orders. For instance, one widely applied psychometric model for analyzing paired comparison data is the Bradley-Terry-Luce (BTL) model (Bradley

& Terry, 1952; Luce, 1959). The BTL model predicts the probability, $p(A > B)$, of choosing

stimulus A over stimulus B by

$$p(A > B) = \frac{\psi(A)}{\psi(A) + \psi(B)},$$ 1

where $\psi(\bullet)$ is a ratio scale (Luce, 1959) representing the strength or magnitude of the stimuli.

In its original form the BTL model contains no parameters by which to account for systematic

asymmetries in paired comparisons, yet this model is still in common use (Courcoux,

Chaunier, DellaValle, Lourdin, & Semenou, 2005; Wickelmaier & Schmid, 2004; see also

Batchelder, Hu, & Smith, 2009).

However, the classic BTL model has been extended to account for systematic

asymmetries in paired stimulus comparisons by assuming such effects arise as a result of bias.

The extended BTL model (Davidson & Beaver, 1977) comprises the order effect in the

parameter $w > 0$. Let $p(A > B \mid A, B)$ denote the probability that A is chosen over B, given

that A was presented first (or to the left). Then the choice probabilities are formally defined as

$$p(A > B \mid A, B) = \frac{\psi(A)}{\psi(A) + w \cdot \psi(B)} \qquad p(B > A \mid A, B) = \frac{w \cdot \psi(B)}{\psi(A) + w \cdot \psi(B)}.$$ 2

If w is smaller (greater) than one, the bias favors the first / left (second / right) choice interval;

there is no order effect if $w = 1$. In terms of the extended BTL model, the order effect is

additive in $\log_e \psi$: $\mathrm{logit}[p(A > B \mid A, B)] = \log_e \psi(A) - \log_e \psi(B) - \log_e w$. Additive effects

models are most usually associated with the notion that systematic asymmetries observed in

the method of paired comparisons arise as a result of bias. On this account, bias may reflect a

general response preference which may vary randomly from trial to trial and with changes in

stimulus response assignment. Alternatively, bias may reflect response readiness which may

additionally depend on the temporal or spatial separation between paired stimuli but, as an

additive constant, bias is not assumed to arise as a result of perceptual processes involved in

the sensory comparison of stimulus magnitudes per se.

Alternatively, perhaps the most comprehensive model of TOEs and SOEs to date is the

Sensation Weighting (SW) model as developed by Hellström (1979, 2000, 2003). Hellström's

SW model posits a weighting of activation inspired by each stimulus event and a reference

level (ReL) based on previous experience and on an averaging of incoming stimulation (after

Helson, 1964; Michels & Helson, 1954). Formally the SW model is expressed

$$d_{12} = k\{[s_1\psi_1 + (1 - s_1)\psi_{r1}] - [s_2\psi_2 + (1 - s_2)\psi_{r2}]\} + b, \qquad\qquad 3$$

where d_{12} 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. The constant term b allows

for a possible effect of bias on d_{12} that operates independently of the weighting mechanisms.

Essentially, Hellström's SW model is based on the notion that a primary task of

perception is to represent important changes in stimuli and stimulus relations as faithfully as

possible. This leads to the principle of *adaptive perception* (Hellström, 1986, 1989; cf. Hake,

Rodwan, & Weintraub, 1966), whereby stimulus specific and reference level information are

weighted together in the comparison process to optimize stimulus discrimination. According

to this notion, and the formal discrimination optimization model described by Hellström

(1986, 1989), the relative magnitude and direction of TOEs and SOEs are determined by the

relative dispersion of sensation magnitudes and reference levels, as well as their

intercorrelations. This approach bears comparison to advances in the design of modern

astronomical telescopes (see Beckers, 1993, for a review) in which adaptive optics (Babcock,

1953) operates to counteract atmospheric turbulence: The image of a nearby reference star,

subject to a distortion which is highly correlated over time with that of the target, is used to

continuously adjust the target image and thereby increase its sharpness. The approach may be

compared to the Bayesian inference approach of, for example, Ashourian and Loewenstein

(2011), in which the dynamic aspect is, however, missing.

Table 1 about here

Table 1 shows the optimization of s values for a few representative cases as computed

on the basis of Hellström's (1986, 1989) discrimination optimization model, as detailed in the

Appendix of the present paper. In this table three series are shown. In each series, the values

of the intercorrelations between ψ_1, ψ_2, ψ_{r1}, and ψ_{r2}, and standard deviations of ψ_1 and ψ_2 are

assumed to be fixed, whereas the standard deviations of ψ_{r1} and ψ_{r2} are assumed to be equal

and varied over each series. Table 1 shows that the optimal ratio s_1 / s_2 increases with the ratio

of standard deviations $\sigma(\psi_2) / \sigma(\psi_1)$. This means that, in order to maximize signal to noise

ratios, stimulus weighting is inversely related and the use of reference level information

directly related to stimulus uncertainty. The attainable gain in S / N ($G\%$) is highest for low

values of $\sigma(\psi_r)$.

Yet, despite the importance of both choice probabilities and response times (RTs) for

testing models of choice, there is a dearth of studies in which both measures have been scaled

in similar fashion to examine mechanisms underlying the TOE and SOE. However, RT is

known to be regularly related to stimulus difference in discrimination tasks (Audley & Wallis,

1964; Cartwright, 1941; Cartwright & Festinger, 1943; Crossman, 1953, 1955; Henmon,

1906, 1911; Hick, 1952; Hyman, 1953; Jamieson & Petrusic, 1975a; Kellogg, 1931; Laming,

1968; Link, 1975; Ratcliff, 2002). Studies indicate that the more similar the alternatives

between which experimental participants are requested to choose, the longer and more

variable their RT (Cartwright, 1941; Crossman, 1953, 1955; Henmon, 1906; Kellogg, 1931).

Moreover, numerous researchers have argued in favor of models by which to account for

variation in choice probabilities as well as RTs (Audley, 1960; Link, 1992; Palmer et al.,

2005; Ratcliff, 1978, 2002). For instance, Audley (1960) was among the first to suggest that

choice RTs in paired stimulus discrimination tasks might be scaled using the BTL model (Equation 1) in a similar manner as response probability. More recently, Link (1992) has made similar appeal on the basis of a theory in which the noisy information about the difference between paired stimuli is accumulated by sequential sampling over time. Link's relative judgment theory of two choice RT (Link, 1975) and, by extension, Ratcliff's (1978) diffusion model directly tie choice proportions to RTs by modeling the discrimination process in terms of a random walk.

Random walk models invoke a noisy accumulation process to explain patterns of RT and response probability in timed discrimination tasks. 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 probabilities are determined by the likelihood of crossing either boundary.

To account for asymmetries in stimulus discrimination Link (1975, 1978, 1992) assumed that in paired stimulus comparisons a subjective referent is compared to a value evoked by the stimulus, each defined on commensurate psychological (or psychophysical) continua. After Link (1978) the predicted relationship between response probability and RT is defined formally as

$$ERT_i = \frac{A}{m} \left[(2P_{Ai} - 1) / (S_i - S_r) \right] + K, \qquad\qquad 4$$

where ERT_i is expected response time, A is the distance from the unbiased starting point of the accumulation process to the boundary, m is a constant, K is mean nondecision time, S_i is stimulus magnitude, and S_r the value of the mental referent defined in units of the stimulus. The probability P_{Ai} of the random walk exceeding one boundary (A) given stimulus S_i is determined by the equation of the logistic distribution function of $\theta_i A$, which is the logit of P_{Ai}, such that $P_{Ai} = \exp(\theta_i A) / [1 + \exp(\theta_i A)]$.

In Link's analysis, drift rate μ_i is estimated by $m(S_i - S_r)$, where m is a constant. Subjective stimulus difference is not measured directly by logit P but represented by $\theta_i A$, which needs only to be monotonically related to μ_i. Ignoring the nondecision component of the RT yields an alternative way of estimating μ_i / A in terms of *signed response speed* (*SRS*), defined as $SRS = 1 / RT$ for first (or left) greater responses and $SRS = -1 / RT$ for second (or right) greater responses, separately for each RT. One advantage of using *SRS,* as compared to raw or signed RT, is that *SRS* can be scaled in a similar and directly comparable manner as response probability (Hellström, 2008; Patching, Englund, & Hellström, 2008). A further advantage is that the use of *SRS* to estimate μ_i / A is not limited to the sequential sampling model of Link, (1975, 1978, 1992). Link's model can be thought of as a discrete-time version of the Wiener process (Townsend & Wenger, 2004), which provides the basis for the Ratcliff (1978) diffusion model.

The Ratcliff diffusion model incorporates intertrial variability in the start position and drift rate of the diffusion process, as well as non-decision times, to predict the entire distribution of RTs to choice alternatives. On this basis, it has been successfully applied to data from a wide range of two-alternative forced choice tasks in many different fields (see Wagenmakers, 2009, for a review). Yet, despite its success no attempt has previously been made to apply this model to an understanding of TOEs and SOEs.

Estimating the parameters of the Ratcliff diffusion model is known to be notoriously difficult (Diederich & Busemeyer, 2003; Vandekerckhove & Tuerlinckx, 2007), and in paired comparison tasks sufficient data points are required for both alternate responses. Most recently, though, a realistic solution to these problems has emerged (Vandekerckhove, Tuerlinckx, & Lee, 2011). In particular, Vandekerckhove and colleagues detail a hierarchical Bayesian method of fitting the Ratcliff diffusion model which incorporates analysis of the data from all participants while allowing for differences between participants to emerge (see

Lee, 2011, for an overview). This approach provides for the possibility of modeling interindividual differences in the comparison process and further allows for the possibility of explaining variability in the parameters of the diffusion model by regression of relevant parameters on defined predictors with unknown coefficients. Once fit, it predicts the entire RT distribution individually for each participant, which can then be compared to that obtained empirically by way of the integrated RT distribution hazard function, H(t).

In the present context, the RT distribution hazard function, H(t), is of particular interest because it is known to speak directly to notions of processing capacity, construed as the amount of work done during some unit of time (Townsend & Ashby, 1978, 1983) and, so, processes of adaptive perception as posited by Hellström (1986, 1989; cf. Ashourian & Loewenstein, 2011). In particular, Townsend and Wenger (2004; Wenger & Gibson, 2004) delineate capacity issues in two situations. Wenger and Gibson (2004) emphasize issues of so called 'constant-load processing capacity' where capacity is reflected in changes of RT under conditions in which the number of targets (or stimulus dimensions) to be processed is held constant, whereas Townsend and Wenger (2004; Townsend & Nozawa, 1975; 1977) emphasize issues of 'load processing capacity' where capacity is reflected in changes of RT as a function of variation in the number of targets (or stimulus dimensions) between trials. In the present study, the first of these issues is addressed by way of computing an integrated hazard ratio (*HR*) measure of *constant-load processing capacity* by dividing the value of the integrated hazard function, H(t), as computed for each participant, by its group mean; i.e., $HR(t) = \text{H}(t) / \overline{\text{H}(t)}$. Relating the integrated hazard ratio, *HR*(t), to the asymmetry of the weightings in terms of s_1 and s_2 (Equation 3) then promises to shed new light on the precise extent to which relative interindividual differences in the efficiency of stimulus processing reflect interindividual differences in the relative processing of one or the other stimulus magnitudes in each stimulus pair, as distinguished from discrimination optimization in the

comparison process.

In regard to load processing capacity, Townsend and Nozawa (1995; Townsend &

Eidels, 2011) have assayed workload efficiency in terms of the capacity coefficient $C(t)$,

which is derived by dividing the integrated hazard function for timed responses to a particular

type of two-dimensional stimulus $H_{AB}(t)$ by the sum of the integrated hazard functions $[H_A(t)$

and $H_B(t)]$ for timed responses to each dimension alone,

$$C(t) = \frac{H_{AB}(t)}{H_A(t) + H_B(t)} . \qquad\qquad 5$$

Subsequently, Townsend and colleagues (Townsend & Nozawa, 1995; Townsend & Wenger,

2004; Townsend & Eidels, 2011; Wenger & Townsend, 2000) have related values of $C(t)$ to

stochastic inhibitory and facilitatory dependencies, and independence between the target

dimensions in the processing of two- as compared to one-dimensional stimuli.

To date, derivation of the capacity coefficient $C(t)$ has been primarily undertaken on the

basis of simple and disjunctive (go / no-go) RT tasks in which participants are instructed to

respond whenever a defined target stimulus is presented in one location, or in another

location, or simultaneously in both locations, but withhold responding otherwise (see

Townsend & Eidels, 2011; Townsend & Wenger, 2004, for reviews of this research). Yet,

systematic congruent variation of both the brightness and size of paired visual stimuli, over a

predefined range of magnitude values, and comparable systematic manipulation of either

dimension alone, permits coincident assessment of both $C(t)$ and order effects using the

method of paired comparisons. Relating interindividual differences in load processing

capacity, as assessed on the basis of $C(t)$, to the relative weighting of stimulus magnitudes

then promises to shed additional light on individual differences in the relative efficiency of

processing two-dimensional stimuli, as set apart from mechanisms of discrimination

optimization.

Central to the present article is the debate about perceptual processes underlying TOEs

and SOEs. Although this might appear to be a rather narrow issue derived from a particular experimental paradigm, it nevertheless provides a useful starting point to examine some fundamental issues about how we compare stimuli and how we perceive a difference between them. Indeed, a central concern is the degree to which TOEs and SOEs reflect adaptive perceptual processes that express themselves in terms of those perceptual-cognitive mechanisms brought to bear on perception of small differences between paired sensory magnitudes.

EXPERIMENTS 1 AND 2: TIME-ORDER EFFECTS (TOEs)

Experiments 1 and 2 were initially conducted to examine TOEs in timed brightness discrimination (Experiment 1) and separately in timed size discrimination (Experiment 2) of paired visual stimuli. In the first instance, it was of interest to examine *SRS* in relation to logit *P*, and determine precisely the extent to which TOEs change in magnitude and direction with changes in stimulus intensity and with changes in the temporal separation between stimuli. On the basis of Hellström's SW model, it was of further interest to determine whether the systematic weighting of stimulus values maps onto accumulation rates of noisy information about the difference between stimulus values as assessed by fitting the Ratcliff diffusion model. In addition, relations between the magnitude and direction of weightings asymmetries and derived integrated hazard ratio, *HR*, scores were investigated to examine interindividual differences in our capacity to discriminate between paired stimuli.

Method

Participants.

Forty participants, 27 women and 13 men between the ages 20 and 45 (mean 28 yrs), took part in Experiment 1, and 40 different participants, 31 women and nine men between the ages 19 and 44 (mean 26 yrs), took part in Experiment 2. All participants were recruited from Stockholm University's student population and received course credits for taking part. All

reported normal or corrected to normal vision and all claimed to be right handed.

Apparatus.

A microcomputer (Dell Precision PWS370, Dell Inc, Round Rock, Texas 78682, USA) running MATLAB (The MathWorks, Inc.) controlled the experiments. Stimulus presentation and timing were controlled using a Cambridge Research Systems Bits++ digital video processor (Cambridge Research Systems Ltd., Rochester, Kent, U.K.) along with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). RTs were collected by way of two keys connected to the microcomputer via an ActiveWire USB device (ActiveWire Inc., Palo Alto, California 94306, USA). Timing tests of the experimental set-up, conducted using the Black Box Toolkit (Plant & Hammond, 2002), verified the consistency of the timings requested by the experiment script (see Englund & Patching, 2009, for details).

Visual stimuli were presented on a gamma corrected 21" (40.5 cm by 30.5 cm viewable area) ViewSonic G220f video monitor (ViewSonic Corporation, 381 Brea Canyon Road, Walnut, CA 91789, USA). Luminance was measured using a ColorCal optical photometer (Cambridge Research Systems Ltd.). A chin rest was fixed at a distance of 57 cm from the screen of the video monitor to ensure a constant viewing distance. Participants responded by way of two response keys positioned 15 cm apart, placed centrally and horizontally in regard to the vertical midline of the video monitor at a convenient distance from participants.

All participants were tested individually in a quiet and darkened testing room. The microcomputer used to control the experiments was placed outside and adjacent to the testing room. All other equipment was concealed behind black poster board whereby only the screen of the monitor was visible to participants. The pixel resolution of the video monitor was 1024 by 768 with a refresh rate of 100 Hz. Background illumination of the monitor was held constant at 0.1 cd/m^2.

Stimuli.

The stimuli were successively presented paired circular spots of light. In Experiment 1, each light spot was presented at nine different luminance levels from 3.5 to 5.9 cd/m^2 in eight steps of 0.3 cd/m^2. The size of each spot was held constant at 5 mm diameter. In Experiment 2, the diameter of each spot was systematically varied from 5.1 to 6.7 mm in eight steps of 0.2 mm. [2] The luminance of each spot was held constant at 4.7 cd/m^2. In both Experiments 1 and 2, the temporal separation (ISI) between the paired spots was 400, 800, 1600, or 3200 ms.

Design.

Figure 1 about here

Figure 1 illustrates the factorial combination of the luminance of light spots used in Experiment 1. After Hellström (1978) the nine separate luminance levels of each spot were combined factorially in mean luminance and difference of luminance to create 25 different stimulus pairs in a 'diamond'-shaped arrangement. This design has proved to be efficient for studying discrimination of paired stimuli and avoids problems of collinearity in later statistical analysis (Hellström, 1979, 2003). In Experiment 1, the intrapair differences in luminance were ±1.2, ±0.6, and 0 cd/m^2 (Figure 1, left to right diagonal), and the mean luminances of the two light spots in each stimulus pair were 4.1, 4.4, 4.7, 5.0, and 5.3 cd/m^2 (Figure 1, right to left diagonal). Each stimulus pair was presented with an ISI of 400, 800, 1600, or 3200 ms, yielding 100 unique stimulus pairs. The same design was used in Experiment 2, only in this experiment the intrapair size differences in diameter were ±0.8, ±0.4 and 0 mm, and the mean diameter of the two light spots in each stimulus pair was 5.1, 5.5, 5.9, 6.3, and 6.7 mm.

In each experiment participants took part in four sessions, one on each of four different days. Each session was divided into two sections (practice, experimental), with no obvious

transition from one section to the next. Within each session, each stimulus pair was shown in

pseudorandomly constructed cycles of 100 items. The first 25 trials of each session comprised

a random selection of the stimulus pairs and were designated as practice trials. After

presentation of the 25 practice trials, each participant was required to complete a further 400

trials in the session. New pseudorandom orders were used for each participant and each

session.

In each experiment 20 of the participants were instructed to press the left response key

with the index finger of their left hand if they perceived the first light spot to be brighter

(Experiment 1) or larger (Experiment 2) than the second, and the right response key with the

index finger of their right hand if they perceived the second light spot to be brighter (or

larger) than the first. The other 20 participants in each experiment were instructed to press the

left response key with the index finger of their left hand if they perceived the second light

spot to be brighter (Experiment 1) or larger (Experiment 2) than the first, and the right

response key with the index finger of their right hand if they perceived the first light spot to

be brighter (or larger) than the second.

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. Participants were required to press one of the response keys to indicate that they

had understood the instructions and to start the experimental session. On each trial, each light

spot was presented for 200 ms. RTs were measured consistently from the onset of the

secondly presented light spot of each pair. The intertrial interval was set to 3000 ms. On the

average, participants took 40 minutes to complete each experimental session.

Data analysis.

Data analysis was conducted in 4 parts. First, relations between logit P and mean SRS

were examined, providing support for the view that choice probabilities and RTs are closely

tied in paired magnitude discrimination. Subsequently, logit P and mean SRS were

summarized over stimulus magnitude, which shows the magnitude and direction of TOEs to

depend on stimulus magnitude and ISI. Second, the data were analyzed in terms of

Hellström's SW model, by estimation of weightings W_1 and W_2, where $W_1 \propto s_1$ and $W_2 \propto s_2$

(Equation 3). This analysis revealed a differential weighting of stimulus magnitudes, and by

further analysis of $(W_1 - W_2)$ and $(W_1 + W_2)$ we show that changes in the magnitude and

direction of TOEs with changes in paired stimulus magnitude fall naturally out of the relation

$W_1 \neq W_2$. Third, the Ratcliff diffusion model was fit to the data, and W_1 and W_2 estimated from

changes in diffusion drift rates over stimulus pairs at each ISI, providing further support for

the view that subjective stimulus magnitudes are systematically weighted in the comparison

process. Fourth, weightings differential percentages were related to the hazard ratio (HR) and

separately to the correlation between the physical paired stimulus differences and empirically

assessed subjective differences, d_{12} (Equation 3). These analyses provide an initial indication

that increased efficiency of stimulus processing, in terms of decreased RT, is quite different

from discrimination optimization and achieved by way of decreased relative weighting of one

of the two stimulus magnitudes in each stimulus pair.

Results

In analysis of the data all responses faster than 200 ms were defined as premature

responses and all responses slower than 2000 ms defined as misses and discarded. This

resulted in the removal of 1.78% of responses in Experiment 1 and 0.69% of responses in

Experiment 2.

On the basis of arguments made in the introduction, SRS was calculated as 1 / RT (in

seconds) for first brighter / larger responses, and -1 / RT (in seconds) for second brighter /

larger responses separately for each RT. In addition, the relative frequency of binary choice

responses to the stimuli was expressed in terms of logit P. Figure 2 shows the relationship

between mean SRS and logit P as computed on the basis of all the choice responses, minus

premature responses and misses, as made by participants to each of the paired stimuli at each

ISI, and as fit by standard linear regression separately over the 25 different stimulus pairs at

each ISI.

Figure 2 about here

In the present case, logit P and SRS show close linear correspondence, all values of

Adjusted R^2 were found to be equal to .99, and so mean SRS was computed over trials for each

of the five average magnitude levels of the paired stimuli (see Figure 1, left to right diagonal)

and for each participant separately. In addition, log odds ratios (logit P) were computed for

each magnitude level. Summaries of these data are shown graphically in Figure 3.

Figure 3 about here

The data summarized in Figure 3 were submitted, separately for Experiments 1 and 2, to

an analysis of variance (ANOVA) with one between-participants' factor (response

assignment) and three within-participant factors [dependent measure (logit P, SRS), ISI, and

stimulus magnitude]. For Experiment 1 this analysis revealed a main effect of stimulus

magnitude, $F(4, 152) = 12.36$, $p < .001$, $\eta_p^2 = .25$, along with an interaction between

dependent measure and stimulus magnitude, $F(4, 152) = 26.90$, $p < .001$, $\eta_p^2 = .41$. For

Experiment 2 statistically reliable interactions were found between dependent measure and

stimulus magnitude, $F(4, 152) = 39.37$, $p < .001$, $\eta_p^2 = .51$, between ISI and stimulus

magnitude, $F(12, 458) = 3.23$, $p < .001$, $\eta_p^2 = .08$, and between dependent measure, ISI, and

stimulus magnitude, $F(12, 456) = 4.72$, $p < .001$, $\eta_p^2 = .11$. No other main effects or

interactions were found to be statistically reliable (all ps > .05).

Sensation weighting. To examine TOEs in more detail the data were subsequently

analyzed in terms of Hellström's SW model (Equation 3). After Hellström (1979, 2003) we

first fit $\qquad\qquad\qquad d_{12} = W_1\psi_1 - W_2\psi_2 + U,$ $\qquad\qquad\qquad$ 6

to empirical scale values (logit P, mean SRS) of stimulus difference, d_{12}, where k is a scale

constant and $W_1 = ks_1$ and $W_2 = ks_2$. Weights W_1 and W_2 were estimated by logistic regression

of binary responses and by linear regression of SRS on the standardized log values of the

luminance [i.e., $z_{\log_e(cd/m^2)}$], and diameter [i.e., $z_{\log_e(mm)}$], of the firstly and secondly presented

stimuli of each stimulus pair. [3], [4]

To assess systematic asymmetries in the comparison and discrimination of paired

stimuli, weightings differential percentages ($WD\%$) were then calculated for each participant

and each ISI, by dividing the difference between the weightings, $W_1 - W_2$, by their mean;

$WD\% = 200(W_1 - W_2) / (W_1 + W_2)$. 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 this analysis are shown in Table

2. Further analyses of the influence of response assignment, by way of (two-tailed) two-

sample t-tests, on the magnitude and direction of $WD\%$ failed to reveal any statistically

reliable effects (all ps > .05).

Table 2 about here

In line with the procedures described by Hellström (1985, 2003), additional tests of the

SW and extended BTL models were conducted by logistic regression of response probability

and separately standard regression of *SRS* on $(\Phi_1 - \Phi_2)$ and $(\Phi_1 + \Phi_2)$, where Φ_1 and Φ_2 are the standardized log values of the luminance [$z_{\log_e(cd/m^2)}$; Experiment 1] and size [$z_{\log_e(mm\ diameter)}$; Experiment 2] of the firstly and secondly presented stimuli in each stimulus pair. Fit individually to each participant's data, these analyses provide a simple yet powerful statistical method of a) verifying that the speed and direction of participants' choice responses are based on the differences between the luminance or size of the stimuli in each pair and, b) examining the influence of stimulus magnitude on the logistic probability and signed speed of said choice responses.

For clarification, and in line with the procedures as discussed by Hellström (1979, 2003), Equation 6 may be rewritten equivalently as

$$d_{12} = (W_1 + W_2)\,(\Phi_1 - \Phi_2)\,/\,2 + (W_1 - W_2)\,(\Phi_1 + \Phi_2)\,/\,2 + U. \qquad 7$$

On this basis, statistically reliable coefficients for $(\Phi_1 - \Phi_2)$ indicate that a significant proportion of the variance in d_{12} (i.e., the subjective difference between the compared stimuli) can be attributed to the defined physical differences between the paired stimuli. Most importantly, if statistically reliable coefficients obtain additionally for $(\Phi_1 + \Phi_2)$ then a significant proportion of otherwise unaccounted for variance in d_{12} can be attributed to changes in the overall magnitude of paired stimuli which, by all account, indicates that $(W_1 - W_2) \neq 0$.

Testing the distribution of obtained coefficients against zero by way of one-sample (two-tailed) *t*-tests, these analyses revealed all coefficients for $(\Phi_1 - \Phi_2)$ to be statistically reliable (all *p*s < .05), indicating that the speed and proportions of participants' alternate responses were based on the perceived differences between the paired stimuli depending on their physical difference, and in both Experiments 1 and 2 all coefficients for $(\Phi_1 + \Phi_2)$ were found to be statistically significant (all *p*s < .05), implying $W_1 \neq W_2$. Consequently, the present findings indicate that TOEs in discrimination of the brightness and in discrimination of the

size of visual stimuli do not arise merely as result of additive bias (cf. Davidson & Beaver, 1977) and, as the BTL model implies $W_1 = W_2$, this model cannot account for the results.

Diffusion model analyses. To investigate further the underlying reason why TOEs arise, the proportions of the alternate responses as well as the RTs were subject to diffusion model analyses. These analyses were conducted following the procedures described by Vandekerckhove et al. (2011) who adopt a hierarchical Bayesian approach, as implemented in the statistical software package WinBUGS (Lunn, Thomas, Best, & Spiegelhalter, 2000), to fit the Ratcliff diffusion model. According to this approach the joint density function y of hitting one boundary (A) at time t is determined by a Wiener process characterized by four main parameters, $y \sim \text{Wiener } [\alpha, \beta, \tau, \delta]$.

The standard interpretation of each of these parameters is: α = boundary separation, the difference between the two boundaries A and $-A$; β = bias, the trial to trial start position of the diffusion process (its distance from $-A$, divided by $2A$); δ = drift rate, the trial to trial speed and direction of the accumulation process, and τ = nondecision time, residual RT not accounted for by the diffusion process. Following the procedures described by Vandekerckhove et al. (2011) boundary separation (α) was allowed to vary over participants, but held constant over stimulus conditions, bias (β) was allowed to vary over participants, over ISIs, and within reasonable limits, randomly over trials, while drift rate (δ) and nondecision time (τ) were allowed to vary over participants and randomly over trials.

Most importantly, mean drift rate (μ) was regressed on Φ_1 and Φ_2:

$$\mu_{(pis)} = W_{1(pi)}\Phi_{1(pis)} - W_{2(pi)}\Phi_{2(pis)} + U_{(pi)}, \hspace{3cm} 8$$

where $W_{1(pi)}$, $W_{2(pi)}$, and $U_{(pi)}$ were drawn from standard non-informative priors, with indices p for participants ($p = 1,\ldots, 40$); i for ISI ($i = 1,\ldots, 4$), and s for stimulus pairs ($s = 1,\ldots, 25$). After Vandekerckhove et al. all estimated parameters of the fit model are based on 2000 samples following a 'burn-in' of 4000 samples, which revealed good convergence by visual

inspection of 'trace plots'. Table 3 shows mean W_1, W_2, and $WD\%$, as determined by regression of mean drift rates on Φ_1 and Φ_2, and the relation $WD\% \neq 0$ as tested by (two-tailed) one sample t-tests.

Table 3 about here

To evaluate absolute model fit, a new data set was simulated for each participant as predicted by the estimated parameters of the fit diffusion model. Subsequently, logit P was calculated for each participant and each stimulus condition in exactly the same manner as described for the analysis of empirically obtained response probabilities. For the purpose of forthcoming capacity analysis, RT distribution integrated hazard functions were also computed on the basis of the empirically obtained RTs, and on the basis of the simulated data, following the procedures laid down by Wenger and Townsend (2000). Figure 4 shows logit P for the five average magnitude levels of the paired stimuli, along with the RT distribution integrated hazard functions ($H_B(t)$ for brightness discrimination and $H_S(t)$ for size discrimination), collapsed over ISIs. As shown in Figure 4, both the mean and spread of the empirically determined psychometric and chronometric functions are reasonably well captured by the fit diffusion model.

Figure 4 about here

Processing capacity. To examine relations between constant-load processing capacity and weightings differential percentages, a hazard ratio score [i.e., $HR(t)$] of relative constant-load processing capacity was computed for each participant as described in the introduction of the present paper. The resultant ratio scores were then averaged over time bins (t) to yield an

overall integrated hazard ratio score, *HR*, of relative constant-load processing capacity for each

participant. Stepwise polynomial 3rd-degree regressions were subsequently used to investigate

relations between *HR* and weightings differential percentages, *WD*%, as determined on the

basis of logit *P* and as determined on the basis of *SRS*, averaged over ISIs. [5] For Experiment 1

both the linear ($p < .001$) and the quadratic term ($p < .05$) proved to be statistically reliable, as

determined on the basis of logit *P*, and as determined on the basis of *SRS*. For Experiment 2,

the quadratic ($p < .05$) and cubic term ($p < .001$) were found to be statistically significant as

determined on the basis of logit *P*, and the cubic term ($p < .001$) proved statistically reliable as

determined on the basis of *SRS*. No other linear, quadratic or cubic terms were found to be

statistically reliable (all ps > .05). Figure 5 (panels a and c) shows the relations between *HR* and

WD%, as fit by 2nd-degree polynomials to the data averaged over ISIs for each participant.

Figure 5 about here

To address issues concerning the efficiency (speed) versus optimality (accuracy) of

discriminatory responses the correlation (Pearson's *r*), between the physical stimulus

difference ($\Phi_1 - \Phi_2$) and d_{12} as empirically assessed (i.e., logit *P* and *SRS*), was regressed on

WD%, using 3rd-degree polynomials in the same manner as described above. On the basis of

the discrimination optimization model presented in the Appendix, the highest correlation

between the physical stimulus difference ($\Phi_1 - \Phi_2$) and response measures (logit *P* or *SRS*)

should be found for a *WD*% close to its optimal value. For Experiment 1 this analysis failed to

reveal any statistically reliable linear, quadratic or cubic terms (all ps > .05). For Experiment

2, this analysis revealed statistically reliable quadratic ($p < .01$) and cubic terms ($p < .05$), as

determined on the basis of logit *P*, and a statistically reliable cubic term ($p < .001$) as

determined on the basis of *SRS*. All other terms were found to be statistically unreliable (ps >

.05). Figure 5 (panels b and d) shows the relations between the correlation, r, of the reported and physical difference between the paired stimuli, and weightings differential percentages, *WD%*, as fit by 2nd-degree polynomials to the data averaged over ISIs for each participant.

Discussion

The findings of Experiments 1 and 2 show characteristic time-order effects in timed brightness discrimination and separately size discrimination of paired visual stimuli (Figure 3). For Experiment 1 (brightness discrimination), the magnitude and direction of TOEs was found to vary with changes in ISI, and go from positive to negative with increased stimulus magnitude. For Experiment 2 (size discrimination), a similar but less pronounced pattern of TOEs obtained. In this respect, the present data agree with previous studies of TOEs (Bartlett, 1939; Hellström, 2003; Needham, 1934; Postman, 1946; Woodrow, 1933), which indicate that the precise magnitude and direction of TOEs depends on the type of stimulus judgment, but tend to be positive when paired stimuli are separated by short temporal interstimulus intervals (ISI) and have low levels of stimulus intensity, and negative for longer ISIs and high levels of stimulus intensity.

The present findings show TOEs to occur for both response probability and RT. In particular, a close correspondence obtained between response probability scaled in terms of logit P and RTs scaled in terms of *SRS*. Moreover, both psychometric and chronometric results revealed TOEs, and similar patterns of weightings differentials as determined on the basis of Hellström's SW model. Therefore, it appears that both logit P and *SRS* are sensitive to systematic asymmetries in the method of paired comparisons and can be used in similar fashion to scale performance as a function of the time-order of stimulus presentation.

In showing a close correspondence between logit P and *SRS*, the present findings agree with Link's sequential sampling and Ratcliff's diffusion model, which directly tie choice proportions to RTs (Link, 1975, 1978, 1992; Link & Heath, 1975; Ratcliff, 1978, 2002). In

regard to the Ratcliff diffusion model, a similar pattern of weightings and weightings differential percentages was obtained by regression of the standardized stimulus values on drift rate, as found by regression of the standardized stimulus values on logit *P* and *SRS*. Drift rate change is thought to reflect the influence of stimulus encoding on both RTs and choice proportions (Palmer et al., 2005; Vandekerckhove & Tuerlinckx, 2007). Consequently, it appears that the noisy activation evoked by the luminance or size of the stimuli is systematically weighted in human paired comparisons of these stimulus attributes, providing support for Hellström's SW model which specifies a role for reference level information in discrimination of paired stimulus magnitudes.

In addition, the present findings indicate that participants' relative capacity to process the stimuli, as assessed in terms of *HR*, is related to the magnitude and sign of weightings differential percentages (Figure 5, panels a and b). One possibility, here, is that participants increased their efficiency of stimulus processing by increased reliance on the magnitude of the firstly presented stimulus of each stimulus pair. For Experiments 1 and 2, inverted '*U*' shaped relations obtained between the correlation, *r*, of the reported and physical difference between the paired stimuli, and weightings differential percentages, *WD%* (Figure 5, panels b and d). Consequently, the findings of Experiments 1 and 2 suggest that increased processing efficiency, in terms of decreased RT, is achieved by basing the choice response on the perceived magnitude of the firstly presented stimulus of each stimulus pair. The failure to find statistically reliable relations in Experiment 1 could simply be due to the necessarily restricted range of magnitude levels used in the present experiments.

EXPERIMENTS 3 AND 4: SPACE-ORDER EFFECTS (SOEs)

It remains to be seen whether SOEs will arise in a similar manner as the TOE in Experiments 1 and 2 when the paired visual stimuli are presented simultaneously and separated by a spatial interval. One account of the TOE is that it reflects the waxing and

waning of the two sensory memory traces over time, as inspired by the firstly as compared to secondly presented stimulus (Köhler, 1923; Lauenstein, 1933; Link, 1992). Assuming that over time the memory trace loses information content rather than physical strength as Köhler thought, the SW model for TOEs describes the integration of information at the moment of comparison, substituting reference level information for lost or degraded stimulus information (Equation 3). On this basis, therefore, the lesser evidence of SOEs than of TOEs in the literature could be due to the greater memorial demands placed on participants when paired stimuli are presented successively and separated temporally than when stimuli are presented simultaneously and separated by a spatial interval. Consequently, it was of further interest to determine whether SOEs would arise in paired comparisons of the brightness and separately size of the visual stimuli as used in Experiments 1 and 2, and also whether such order effects would similarly map onto diffusion drift rates and relate to relative constant-load processing capacity, as found in present analyses of TOEs.

Participants.

40 participants, 31 women and nine men, between the ages 19 and 62 (mean 30 yrs) who had not previously taken part in Experiments 1 and 2 took part in Experiment 3, and a further 40 different participants took part in Experiment 4, comprising 28 women and 12 men between the ages 18 and 44 (mean 27 yrs). They were recruited from Stockholm University's student population and received course credits for taking part. All reported normal or corrected to normal vision and all claimed to be right handed.

Stimuli.

In line with Experiments 1 and 2 the visual stimuli were circular spots of light. However, pilot work showed it somewhat harder to discriminate between the brightness and separately size of these visual stimuli when presented simultaneously and separated spatially than when presented successively and separated by a temporal interval. Consequently, the

number of practice trials was increased from 25 to 100 in Experiment 3, and the range of

magnitude levels, and hence physical difference between paired stimuli, was increased in both

Experiments 3 and 4. In Experiment 3, the luminance of the light spots ranged from 1.5 to 7.9

cd/m^2 in eight steps of 0.8 cd/m^2, and their diameter was held constant at 5 mm. In

Experiment 4, the diameter of the light spots ranged from 4.3 to 7.5 mm in eight steps of 0.4

mm, and their luminance was held constant at 4.7 cd/m^2.

Procedure.

On each trial, the two light spots were presented simultaneously for 200 ms with a

spatial separation between the spots of 10, 20, 40, or 80 mm. In Experiment 3, participants

were instructed to respond by pressing the left response key if they perceived the left spot to

be brighter than the right spot and the right key if they perceived the right spot to be brighter

than the left spot. No attempt was made to manipulate response assignment in Experiment 3,

because of difficulties reported by participants, in an earlier pilot study, when attempting to

follow instructions to respond by pressing the left response key if they perceived the right

light spot to be brighter than the left. This issue was addressed in Experiment 4. In

Experiment 4, 20 participants were instructed to press the left key if they perceived the left

light spot to be larger than the right and the right key if they perceived the right light spot to

be larger than the left. The other 20 participants were instructed to press the left key if they

perceived the left light spot to be smaller than the right and the right key if they perceived the

right light spot to be smaller than the left. All remaining aspects of Experiments 3 and 4 were

the same as in Experiments 1 and 2.

Data analysis.

Data analysis was conducted in the same manner as for Experiments 1 and 2. First,

relations between logit *P* and mean *SRS* were examined, and logit *P* and mean *SRS* were

summarized over paired stimulus magnitudes revealing SOEs which were found to depend on

stimulus magnitude and ISI. Second, the data were analyzed in terms of Hellström's SW model by estimation of weightings W_1 and W_2, providing further support for the view that the dependence of SOEs on paired stimulus magnitude falls naturally out of the relation $W_1 \neq W_2$. Third, the Ratcliff diffusion model was fit to the data, and weightings W_1 and W_2 estimated from diffusion drift rates, providing additional support for the view that subjective stimulus magnitudes are systematically weighted in the comparison process. Fourth, weightings differential percentages were related to the hazard ratio (*HR*) and separately to the correlation between the physical paired stimulus differences and empirical assessment of subjective differences, d_{12} (Equation 3). This analysis again indicates that increased efficiency of stimulus processing comes with the cost of decreased fidelity of one or the other stimulus magnitudes in each stimulus pair. In sum, it appears that SOEs fall out of similar mechanisms of discrimination optimization as TOEs, although weightings differentials were found to be attenuated in analysis of SOEs as compared to those obtained in analysis of TOEs (Experiments 1 and 2).

Results

The data were analyzed in the same manner as in Experiments 1 and 2. All responses faster than 200 ms were considered premature responses and all responses slower than 2000 ms considered as misses and removed. This resulted in the removal of 0.53% of responses from Experiment 3 and 0.13% of responses from Experiment 4.

Figure 6 about here

As in Experiments 1 and 2, *SRS* was calculated as 1 / RT (in seconds) for left brighter / larger responses, and -1 / RT (in seconds) for right brighter / larger responses, and the relative frequency of binary choice responses to the stimuli expressed in terms of logit *P*. Figure 6

shows the relationship between mean *SRS* and logit *P,* as computed on the basis of all choice

responses over participants for each stimulus pair, and as fit by linear regression over the 25

stimulus pairs at each spatial separation. As detailed in Figure 6 all values of Adjusted R^2 were

found to be equal to .99. Subsequently, *SRS* and logit *P* were computed over trials for each of

the five stimulus magnitude levels and for each participant separately. Summaries of these data

are shown graphically in Figure 7.

Figure 7 about here

For Experiment 3 the data summarized in Figure 7 were submitted to an ANOVA with

three within-participant factors [dependent measure (logit *P, SRS*), spatial separation, and

stimulus magnitude]. For Experiment 4 the between-participants factor of judged magnitude

direction (i.e., 'smaller than' versus 'larger than' judgments) was also included. Analysis of the

data obtained in Experiment 3 revealed a main effect of spatial separation, $F(3, 117) = 3.59$, *p*

$< .05$, $\eta_p^2 = .08$, and an interaction between spatial separation and stimulus magnitude, $F(12,$

$468) = 2.00$, $p < .05$, $\eta_p^2 = .05$. For Experiment 4 a main effect of stimulus magnitude, $F(4,$

$152) = 2.60$, $p < .05$, $\eta_p^2 = .06$, obtained. No other main effects or interactions were found to

be statistically reliable.

Sensation weighing. Following the procedures adopted in Experiments 1 and 2, SOEs

were examined further in terms of Hellström's SW model. Weightings differential

percentages were subsequently calculated for each participant by changes in the spatial

separation between paired stimuli, and the relation *WD%* $\neq 0$ tested by way of (two-tailed)

one-sample *t*-tests. Analysis of the influence of the judged direction in Experiment 4 (i.e.,

'lesser than' versus 'greater than' judgments) on the magnitude and direction of *WD%* failed to

reveal any statistically reliable effects (all *p*s $> .05$), and so Table 4 details the weightings and

weightings differentials collapsed over this factor.

Table 4 about here

Further analysis of the BTL and SW models by logistic regression of binary responses, and by standard regression of *SRS*, on $(\Phi_1 - \Phi_2)$ and $(\Phi_1 + \Phi_2)$ revealed significant coefficients for $(\Phi_1 - \Phi_2)$ in both Experiments 3 and 4 (all *p*s < .01), indicating that the speed and direction of participants' choice responses were based on the perceived differences between the paired stimuli, which in turn depended on their physical difference. In addition, testing of the distribution of obtained coefficients against zero by way of one-sample (two-tailed) *t*-tests, this analysis revealed all coefficients for $(\Phi_1 + \Phi_2)$ to be statistically reliable (all *p*s < .05) indicating that the speed and direction of participants' responses was reliably influenced by changes in the mean physical magnitude of the stimuli, ruling out the BTL and extended BTL models which assume $W_1 = W_2$.

Diffusion model analysis. The empirical response probability and RT data were subsequently fit to the Ratcliff diffusion model in exactly the same manner as for Experiments 1 and 2. Table 5 shows mean W_1, W_2, and *WD*%, as determined by regression of drift rates on Φ_1 and Φ_2, together with the *p*-values for the relation $WD\% \neq 0$ as tested by single-sample (two-tailed) *t*-tests.

Table 5 about here

Figure 8 shows a graphical illustration of absolute diffusion model fit, as formulated in terms of logit *P* over the five average magnitude levels of the paired stimuli, and over integrated hazard functions of RT, collapsed over spatial separations between the stimuli.

Figure 8 again shows a reasonable fit of the Ratcliff diffusion model to the mean and spread of psychometric and chronometric functions.

Figure 8 about here

Processing capacity. Following the procedures adopted in Experiments 1 and 2, relations between $HR(t)$ [where $HR(t) = H(t) / \overline{H(t)}$] and $WD\%$ were examined by averaging $HR(t)$ over time bins (t) and by using 3rd-degree polynomial regressions of HR on $WD\%$. For Experiment 3, this analysis failed to reveal any significant terms for regression of HR on $WD\%$ (all ps > .05). For Experiment 4, a statistically reliable cubic term obtained, in regard to $WD\%$ as determined on the basis of SRS ($p < .05$). Fit by 2nd-degree polynomials to the data averaged over ISIs for each participant a 'U' shaped relation obtained between HR and $WD\%$, and no other terms proved to be statistically reliable (all ps > .05).

To examine further issues concerning the relative efficiency versus optimality of participants' discriminatory responses, Pearson's r, as obtained by correlating the response measures (logit P, SRS) with the physical difference between stimulus values (Φ_1 - Φ_2) for each participant, were regressed on $WD\%$ in the same manner as described in Experiments 1 and 2. For Experiment 3, this analysis failed to reveal any statistically reliable coefficients (all ps >.05). For Experiment 4 a statistically significant cubic term obtained, for $WD\%$ as determined on the basis of logit P, and as determined on the basis of SRS (both ps < .001). As fit by 2nd-degree polynomials to the data averaged over ISIs for each participant, a statistically reliable inverted 'U' shaped relation obtained between the correlation r and $WD\%$; for $WD\%$ as determined on the basis of logit P, Adjusted R^2 = .42, $p < .001$, and for $WD\%$ as determined on the basis of SRS, Adjusted R^2 = .46, $p < .001$.

Discussion

Both Experiments 3 and 4 revealed SOEs (Figure 7). For timed brightness discrimination, Experiment 3 shows SOEs to go from negative to positive with increased spatial separation between the paired stimuli, and with increased stimulus brightness. For timed size discrimination, Experiment 4 shows SOEs to change in magnitude and direction with changes in pair average stimulus size. On this basis, the present findings conform to Kellogg's (1931), finding of a negative SOE in split disc brightness discrimination, where the paired stimuli were spatially abutting and, in line with the findings of Charles et al. (2007), Experiment 4 shows SOEs in timed size discrimination. Taken together with Experiments 1 and 2, therefore, it now appears that similar patterns of order effects arise in timed brightness and size discrimination of paired visual stimuli, regardless of whether paired stimuli are presented successively and separated by a time interval or presented simultaneously and separated spatially.

In regard to the sequential sampling model of Link (1975, 1992) and the diffusion model of Ratcliff (1978, 2002) the findings of Experiments 3 and 4 again show a close correspondence between response probability scaled in terms of logit P and RTs scaled in terms of SRS. As analyzed in terms of Hellström's SW model, Experiment 4 revealed positive weightings differentials for simultaneously presented stimuli separated spatially by 80 mm. Moreover, a similar pattern of weightings differentials obtained as analyzed by way of the Ratcliff diffusion model. However, in timed brightness and size discrimination of paired stimuli presented simultaneously and separated spatially, weightings asymmetries are attenuated as compared to those revealed in the current investigation of TOEs.

One interpretation of the relative attenuation of weightings asymmetries found in analysis of comparisons of simultaneous stimuli, as compared to those obtained in analysis of comparisons of successive stimuli, is that weightings asymmetries reflect the memorial demand of the task. In Experiments 3 and 4 the memorial demands of the discrimination task

were reduced by presenting the paired stimuli simultaneously, as compared to Experiments 1 and 2 where the stimuli were presented consecutively and separated by a temporal interval. In addition, the findings of Experiment 4 revealed a '*U*' shaped relation between *WD%* and processing efficiency as assessed in terms of *HR*. Moreover, regression of the correlation, *r*, of the reported and physical difference between the paired stimuli, on weightings differential percentages, *WD%*, revealed an inverted '*U*' shaped relation to the faithfulness of the representation of paired stimulus differences. Consequently, the findings of Experiment 4 again suggest that increased processing efficiency, in terms of decreased RT, comes with the cost of decreased faithfulness of the representation of one or the other stimulus magnitudes in each stimulus pair. As in Experiment 1, the failure to find similar relations in brightness discrimination (Experiment 3) could simply be due to the necessarily restricted range of stimulus values used in the present experiments.

EXPERIMENTS 5 AND 6:

THE TOE AND SOE IN TIMED BRIGHTNESS OR SIZE DISCRIMINATION

Experiments 5 and 6 set out to examine TOEs and SOEs in timed brightness or size discrimination under conditions in which the paired stimuli were varied physically along the dimensions of either luminance or size, or both luminance and size. To this end, three subsets of trials were intermixed in each experiment: a) the luminance difference between the paired stimuli was varied systematically while their size was held constant; b) the size difference between the paired stimuli was varied while their luminance was held constant; c) both attributes of luminance and size were varied congruently. In both Experiments 5 and 6, incongruent combinations of difference in luminance and in size were avoided, and participants were simply instructed to make their choice response based on which of the two paired stimuli was either larger or brighter than the other. In Experiment 5 the paired stimuli were presented successively and separated by a time interval. In Experiment 6 the same

stimuli were presented simultaneously and separated by a spatial interval.

On the grounds that stimulus weighting is inversely related and the use of reference level information directly related to stimulus uncertainty, it was of interest to determine whether participants would base their choice responses more on the size than brightness of the paired stimuli, or vice versa. To this end, systematic asymmetries in paired stimulus comparisons are examined by extension of Hellström's (1979) SW model to include a V-weighting of the dimensions of brightness and size, and the precise extent to which weightings map onto diffusion drift rates examined by fitting the Ratcliff diffusion model. In Experiments 1 through 4, TOEs (Experiments 1 and 2; Figure 3) and SOEs (Experiments 3 and 4; Figure 7) have been found to be of greater magnitude in discrimination of the brightness than size of the paired stimuli. Consequently, on the basis that the physically defined size differences between the paired stimuli are more discriminable than defined brightness differences, evidence of a greater V-weighting on the size than brightness of the paired stimuli would provide further evidence in support of Hellström's (1986, 1989) discrimination optimization model.

In addition, the design of Experiments 5 and 6 allows for intraindividual assessment of load processing capacity, as distinct from constant load processing capacity $HR(t)$, by computation of the capacity coefficient $C(t)$, as defined in Equation 5. Relating interindividual changes in load processing capacity, as assessed on the basis of $C(t)$, to the relative weighting of stimulus magnitudes then promises to shed new light on the precise extent to which load processing capacity relates to systematic asymmetries known to arise in the method of paired comparisons.

Method

Participants.

Forty participants, 37 women and three men, between the ages 19 and 50 (mean 25 yrs),

took part in Experiment 5, and a further 40 different participants took part in Experiment 6, comprising 29 women and 11 men between the ages 18 and 46 (mean 26 yrs). None of the participants had previously taken part in Experiments 1 to 4. All reported normal or corrected to normal vision, and all but three participants (one in Experiment 5 and two in Experiment 6) claimed to be right handed.

Stimuli / design.

The visual stimuli were circular spots of light as used in Experiments 1 to 4. For the stimuli differing along the single physical dimensions of luminance or size, five different luminance levels and separately five different sizes were combined to create 6 stimulus pairs for each dimension. For each of these six stimulus pairs the alternate dimension was held constant over the five magnitude levels of that dimension. For the stimuli differing congruently along both physical dimensions of luminance and size a constant intrapair difference of 0 was also used. In this respect, the five different values of luminance and size were combined factorially in mean luminance, mean size, and difference of luminance and of size to create 9 stimulus pairs in which variations in luminance and size were positively correlated.

In Experiment 5, for the paired visual stimuli differing along the single dimension of luminance, the luminance levels of the stimuli in the six stimulus pairs were 4.1 - 4.7, 4.4 - 5.0, 4.7 - 5.3, 4.7 - 4.1, 5.0 - 4.4, and 5.3 - 4.7 cd/m^2, with a constant intrapair difference of ± 0.6 cd/m^2, and mean average luminance of the two stimuli in each pair of 4.4, 4.7, and 5.0 cd/m^2. For each of these six stimulus pairs the size of the stimuli was held constant at one of five levels from 5.5 to 6.3 mm diameter, in steps of 0.2 mm.

For the visual stimuli differing along the single dimension of size, the five size levels were similarly combined to create six stimulus pairs of 5.5 - 5.9, 5.7 - 6.1, 5.9 - 6.3, 5.9 - 5.5, 6.1 - 5.7, and 6.3 - 5.9 mm diameter, with a constant intrapair difference of ± 0.4 mm and three

mean average size levels of 5.7, 5.9, and 6.1 mm diameter. For each of these six pairs the luminance of the stimuli was held constant at one of five levels from 4.1 to 5.3 cd/m^2 in steps of 0.3 cd/m^2.

In regard to the positively correlated stimuli, differing along both dimensions of luminance and size, the values of luminance and size (cd/m^2 / mm diameter, respectively) were combined factorially to create nine stimulus pairs: (4.1 / 5.5 and 4.7 / 5.9); (4.4 / 5.7 and 5.0 / 6.1); (4.7 / 5.9 and 5.3 / 6.3); (4.4 / 5.7 and 4.4 / 5.7); (4.7 / 5.9 and 4.7 / 5.9); (5.0 / 6.1 and 5.0 / 6.1); (4.7 / 5.9 and 4.1 / 5.5); (5.0 / 6.1 and 4.4 / 5.7); (5.3 / 6.3 - 4.7 / 5.9), with constant intrapair differences of ±0.6 cd/m^2 / ±0.4 mm diameter, and 0 cd/m^2 / 0 mm diameter respectively, along with the three mean average magnitude levels of (4.4 / 5.7); (4.7 / 5.9); (5.0 / 6.1) cd/m^2 / mm diameter. In Experiment 5 the ISI between the successively presented paired stimuli was held constant at 800 ms.

The same design was used in Experiment 6, but on the basis of results from Experiments 1 to 4, the physical range of stimulus values was increased in Experiment 6 as compared to that used in Experiment 5. In Experiment 6 the luminance values of the individual stimuli ranged from 3.1 to 6.3 cd/m^2 in four steps of 0.8 cd/m^2 and the size values ranged from 5.1 to 6.7 mm diameter in four steps of 0.4 mm; the spatial separation between the simultaneously presented paired stimuli was held constant at 40 mm.

Procedure.

In Experiments 5 and 6 each pair of stimuli differing along the single dimension of either luminance or size was presented 16 times, at each of the five magnitude levels of the alternate dimension, giving 80 binary judgments of each stimulus pair. Likewise, each stimulus pair differing congruently along the dimensions of both luminance and size was presented 80 times.

In each experiment participants took part in four sessions, one on each of four days.

Within each session, each stimulus pair was shown in pseudorandomly constructed blocks of 105 items, comprising 60 pairs of stimuli differing along the dimension of either luminance or size, 30 pairs of stimuli differing congruently along the dimensions of both luminance and size, and 15 pairs of stimuli in which there was no physical difference between them. 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. The 105 stimuli within each block were randomly intermixed and new pseudorandom orders were used for each participant and each session.

In Experiment 5, the response keys were placed horizontally on the midline of the video monitor and participants were instructed to use the index finger of their left hand to press the button on their left, and the index finger of their right hand to press the button on their right. In Experiment 6, the response keys were aligned vertically to the midline of the video monitor, and participants were instructed to use the index finger of their left hand to press the button nearest them, and the index finger of their right hand for button presses furthest from them. [6] In each experiment, 20 of the participants were instructed to use the index finger of their left hand if they perceived the first (or left) stimulus to be larger or brighter than the second (or right), or the index finger of their right hand if they perceived the second (or right) stimulus to be larger or brighter than the first (or left). The other 20 participants in each experiment were instructed to use the index finger of their right hand if they perceived the first (or left) stimulus to be larger or brighter than the second (or right), or the index finger of their left hand if they perceived the second (or right) stimulus to be larger or brighter than the first (or left).

All participants were informed explicitly that on no occasion would one visual stimulus be brighter but smaller, or darker but larger than the other. All remaining aspects of Experiments 5 and 6 were the same as Experiments 1 through 4.

Data analysis.

Data analysis was again conducted in 4 parts. First relations between logit P and mean SRS were examined, and logit P and SRS were summarized over pair average stimulus magnitudes, which shows TOEs (Experiment 5) and SOEs (Experiment 6) to arise in magnitude discrimination of the paired stimuli. Second, Hellström's SW model was extended to include a V-weighting of the dimensions of brightness and size and the weightings W_1 and W_2 estimated by regression analysis, providing further support for processes of adaptive perception as posited by Hellström (1986, 1989). Third, the Ratcliff diffusion model was fit to the data, and the weightings W_1 and W_2, along with the V-weightings, estimated from diffusion drift rates. The results agreed with the view that subjective stimulus magnitudes are systematically weighted in the comparison process. Fourth, interindividual differences in the magnitude and direction of TOEs and SOEs were investigated by examining relations between constant load processing capacities as assessed in terms of HR and weightings differential percentages, and intraindividual differences in load processing capacity assessed in term of the capacity coefficient $C(t)$. These analysis provide further evidence by which to suggest that increased efficiency of stimulus processing comes with the cost of decreased fidelity of one or the other stimulus magnitudes in each stimulus pair, whereas no relations were found between weightings differential percentages and $C(t)$.

Results

The data were initially analyzed in the same manner as in Experiments 1 through 4. All responses faster than 200 ms were considered premature responses and all responses slower than 2000 ms considered as misses and removed. This resulted in the removal of 0.71% of responses from Experiment 5 and 0.29% of responses from Experiment 6.

Figure 9 about here

Figure 9 shows the relationship between mean *SRS* and logit *P*, as fit by linear

regression over each stimulus pair. For both Experiments 5 and 6 Adjusted R^2 was found to be

equal to .99. Subsequently, group averaged values of *SRS* and logit *P* were computed for each

of the three average magnitude levels of the paired stimuli. Summaries of these data are

shown graphically in Figure 10.

Figure 10 about here

The data summarized graphically in Figure 10 were submitted, separately for

Experiments 5 and 6, to ANOVAs with one between-participants factor (response

assignment) and three within-participant factors [dependent measure (logit *P*, *SRS*), stimulus

attribute (luminance, size), and stimulus magnitude]. For Experiment 5 this analysis revealed

main effects of stimulus attribute, $F(1, 38) = 5.05$, $p < .05$, $\eta_p^2 = .117$, and stimulus

magnitude, $F(2, 76) = 9.97$, $p < .001$, $\eta_p^2 = .21$. In addition, this analysis revealed statistically

reliable interactions between dependent measure and stimulus attribute, $F(1, 38) = 6.84$, $p <$

.05, $\eta_p^2 = .15$, dependent measure and stimulus magnitude, $F(2, 76) = 183.99$, $p < .001$, $\eta_p^2 =$

.83, and between stimulus attribute and stimulus magnitude, $F(2, 76) = 14.51$, $p < .001$, $\eta_p^2 =$

.27. The main effect of response assignment was not found to be statistically reliable ($p >$

.05), and no other interactions proved statistically significant (all ps > .05).

For Experiment 6 this analysis revealed a main effect of stimulus magnitude, $F(2, 76) =$

7.18, $p < .001$, $\eta_p^2 = .16$, along with an interaction between stimulus attribute and stimulus

magnitude, $F(2, 76) = 4.60$, $p < .05$, $\eta_p^2 = .11$. In addition this analysis revealed an interaction

between dependent measure and response assignment, $F(1, 38) = 10.31$, $p < .01$, $\eta_p^2 = .21$.

Tukey's HSD tests were conducted to examine further the dependent measure by response

assignment interaction. This analysis revealed a statistically reliable decrease in mean *SRS*

(0.08 vs. 0.24) for right hand, left brightest / left largest responses as compared to left hand,

left brightest / left largest responses. No statistically reliable difference in logit *P* (0.10 vs.

0.19, *p* > .05) was found with changes in stimulus-response assignment, and no other main

effects or interactions proved statistically reliable (all *p*s > .05).

Sensation weighting. To examine TOEs and SOEs further, Hellström's SW model

(Equation 3) was generalized to encompass paired two-dimensional stimuli. In this case,

stimuli are conceived as composed of two dimensions that may have different scale factors,

reflecting their importance,

$$d_{12} = W_1 [V_B\Phi_{B1} + V_S\Phi_{S1}] - W_2[V_B\Phi_{B2} + V_S\Phi_{S2}] + U. \qquad\qquad 9$$

In Equation 9, V_B and V_S are the dimensional weightings for brightness and size, respectively,

and their sum must equal one; i.e., $V_B + V_S = 1$. Φ_{B1} and Φ_{B2} are the standardized log values of

the physical luminance of the stimuli and Φ_{S1} and Φ_{S2} are the standardized log values of the

physical size of the stimuli, for the two presentation orders in each experiment (indexed: first

or left = 1, and second or right = 2).

To estimate the parameters of Equation 9, individually for each participant, the

dimensional weightings, V_B and V_S, were first determined by letting $V_B = B_B / (B_B + B_S)$ and V_S

$= B_S / (B_B + B_S)$, where the unstandardized weighting coefficients, B_B and B_S, were obtained

by linear regression of *SRS,* and separately logistic regression of binary responses, for each of

the stimulus pairs differing physically along both dimensions of luminance and size, on *SRS*

and choice responses, respectively, for the corresponding stimulus pairs differing physically

along either dimension alone. The weightings W_1 and W_2 were subsequently calculated for

each participant by logistic regression of choice responses and by linear regression of *SRS* on

the *V*-weighted standardized log values of luminance and size. This allowed calculation of

weightings differential percentages (i.e., *WD%*) in the same manner as in Experiments 1

through 4, and the results of this analysis, along with the results of testing the relation $WD\% \neq$

0 by (two-tailed) one-sample *t*-tests, are detailed in Table 6. Further analyses of the influence

of response assignment, by way of (two-tailed) two-sample *t*-tests, on the magnitude and

direction of *WD*% failed to reveal any statistically reliable effects (all *p*s > .05).

Additional tests to verify that participants' responses were based on perceived

differences between the stimuli depending on the physical values of the stimuli, and for

stimulus magnitude effects on said choice responses were conducted by regression of logit *P*

and *SRS* on $(\Phi_1 - \Phi_2)$ and $(\Phi_1 + \Phi_2)$, using *V*-weighted sums of the dimensions for the two-

dimensional stimuli. For both Experiments 5 and 6 this analysis revealed all coefficients for

$(\Phi_1 - \Phi_2)$ to be statistically reliable (all *p*s < .05), and in both Experiments 5 and 6, testing of

the distribution of obtained coefficients against zero by way of one-sample (two-tailed) *t*-tests,

all coefficients for $(\Phi_1 + \Phi_2)$ were also found to be statistically significant (all *p*s < .05),

providing a further indication that for brightness and size discrimination, $W_1 \neq W_2$. In line

with Experiments 1 through 4, therefore, the indication is that TOEs and SOEs in timed

discrimination of the brightness and size of paired visual stimuli do not arise merely as a

result of additive bias (cf. Hellström, 2003).

Diffusion model analysis. The same hierarchical Bayesian approach was used to fit the

Ratcliff diffusion model as detailed in analysis of Experiments 1 and 2, with the exception

that mean drift rates (μ) were now determined by way of Equation 9, using *V*-weighted sums

of the standardized log values of luminance and size. Table 7 shows mean average estimates

of V_B and V_S, together with W_1, W_2, and *WD*%, along with the *p*-values for the relation *WD*%

$\neq$ 0 as tested by single-sample (two-tailed) *t*-tests.

Table 7 about here

Absolute model fit was assessed in the same manner as in Experiments 1 through 4. Figure 11 shows a graphical illustration of diffusion model fit in terms of logit P as computed over the three average magnitude levels of the paired stimuli, and over RT distribution integrated hazard functions, as computed individually for each participant. Again, Figure 11 shows a reasonable fit of the diffusion model to the empirically obtained response probability and RT data.

Figure 11 about here

Processing capacity. As described in analysis of Experiments 1 and 2, an integrated hazard ratio (*HR*) of relative constant-load processing capacity was computed for each participant by dividing the integrated hazard functions as determined for each participant by the group mean [$HR(t) = \mathrm{H}(t) / \overline{\mathrm{H}(t)}$]. In addition, values of the capacity coefficient $C(t)$ were computed, for each participant in 10-ms time bins, by dividing the integrated hazard function for timed responses to the paired stimuli differing physically along both dimensions of luminance and size [$\mathrm{H}_{BS}(t)$] by the sum of the integrated hazard functions [$\mathrm{H}_B(t)$ and $\mathrm{H}_S(t)$] for those stimuli differing physically along either dimension alone, as detailed in Equation 5. Both ratio measures of processing capacity were then averaged over time bins (t) to yield a measure of constant-load processing capacity (*HR*) and a single measure of load processing capacity (*C*) for each participant.

First, relations between *WD*% and *HR* were examined by way of stepwise 3rd-degree polynomial regressions in the same manner as Experiments 1 through 4. For Experiment 5, this analysis revealed statistically reliable linear, quadratic, and cubic terms (all *ps* < .05) for *WD*% as determined on the basis of logit *P*, and statistically reliable linear and cubic terms

(both *ps* < .05) for *WD*% as determined on the basis of *SRS*. Figure 12a shows the relations

between *WD*% and *HR*, as found for Experiment 5 and as fit by 2nd-degree polynomials over

each participant. For Experiment 6 it made no sense to plot relations between *WD*% and *HR*,

because none of regression terms were found to be statistically reliable (all *ps* > .05).

Figure 12 about here

Second, relations between *WD*% and load processing capacity, *C*, as defined by

Townsend and colleagues (Equation 5), were examined by way of stepwise 3rd degree

polynomial regressions in the same manner as described above. This analysis failed to reveal

any statistically reliable coefficients (all *ps* > .05) for either Experiment 5 or 6. Consequently,

there is no evidence to suggest that changes in the magnitude and direction of TOEs and

SOEs relate to interindividual differences in load processing capacity, as distinct from

constant-load processing capacity, with trial-by-trial changes in the physical differences

between the paired stimulus (either brightness, size, or both).

Finally, to address issues concerning the efficiency versus discriminatory optimality of

participants' choice responses, the correlation *r*, [obtained by correlating the response

measures (logit *P*, *SRS*) with the difference between the now *V*-weighted sums of the two

physical stimulus values for each participant], was regressed on *WD*% in the same manner as

in Experiments 1 through 4. For Experiment 5, this analysis revealed statistically significant

quadratic and cubic terms as determined on the basis of logit *P* (both *ps* < .01), and as

determined on the basis of *SRS* (both *ps* < .05). Figure 12b shows the relations between *WD*%

and *r*, as found for Experiment 5 and as fit by 2nd-degree polynomials over participants. For

Experiment 6, no statistically significant regression terms obtained (all *ps* >.05).

Discussion

Experiments 5 and 6 extend the classic method of paired comparisons by incorporating paired stimuli differing along both dimensions of both brightness and size. Yet again, logit P and SRS were found to be linearly related, and both measures show similar patterns of order effects (Figures 9 and 10). As with the corresponding stimulus pairs in Experiments 1 and 2, the present data show negative TOEs which were found to become increasingly negative with increased stimulus magnitude. Moreover, a similar but conversely signed pattern of SOEs obtained (Figure 10). On this basis, the present findings agree with those of Experiments 1 through 4 and again show that order effects arise systematically in paired comparisons of visual magnitude regardless of whether the paired stimuli are presented successively and separated by a time interval or presented simultaneously and separated spatially.

To examine the underlying reasons why such time- and space-order effects arise, Hellström's SW model was extended to encompass paired two-dimensional stimuli by way of V-weighted sums of the dimensions of perceived brightness and size. This analysis revealed a greater V-weighting on size than brightness in analysis of both logit P and SRS and, in agreement, a greater V-weighting on size than brightness obtained as assessed by way of fitting the Ratcliff diffusion model to the empirical response probability and RT data. One interpretation of the greater V-weighting on size than on brightness is that the physically defined size differences between the paired stimuli were more discriminable, and so relied on more in the discrimination of the paired stimuli, than the physically defined brightness differences. However, further research is required in which the relative discriminability of the size and brightness of paired visual stimuli is systematically manipulated to examine this issue in detail. For now it is sufficient to note that in the present experiments participants' choice responses were made more on the basis of the standardized log values of size than of brightness of the paired stimuli.

As in Experiments 1 to 4, the present analysis of response probabilities and RTs again

revealed differential weighting of sensory magnitudes. In regard to Hellström's SW model, analysis of Experiment 5 further revealed a greater weighting of the secondly as compared to firstly presented stimuli, while Experiment 6 revealed a greater weighting on the left- as compared to rightwardly presented stimuli. Moreover, the relative weighting of stimulus magnitudes over time- and space-order were found to map onto diffusion drift rates, which revealed a similar pattern of weightings to those obtained on the basis of logit P and SRS. On this basis, the present findings agree with those of Experiments 1 through 4 and, taken together, provide further converging evidence of a systematic differential weighting of stimulus dimensions, which operates in the perceptual comparison and discrimination process.

For SOEs, it again appears that weightings asymmetries are attenuated as compared to those revealed in the current investigation of TOEs. This could be due to the decreased memorial demands, when stimuli are presented simultaneously and separated spatially than when stimuli are presented sequentially and separated by a time interval. As in Experiments 1 to 4, the present experiments went on to use RT distribution hazard functions to assess participants' processing capacity in relation to the magnitude and direction of TOEs and SOEs. In this regard, Experiment 5 shows further evidence of a skewed 'U' shaped curvilinear relation between the magnitude and direction of weightings differential percentages and participants' relative constant-load processing capacity as assessed in terms of HR. Moreover, inverted 'U' shaped relations obtained between the correlation, r, of the reported and now V-weighted sums of the two physical stimulus values for each participant, and weightings differential percentages. On this basis, the present experimental work provides further evidence by which to suggest that increased efficiency of stimulus processing, in terms of decreased RT, is achieved by basing the choice response on the perceived magnitude of the firstly presented stimulus of each stimulus pair.

For SOEs, relations between weightings differentials and relative constant-load

processing capacity failed to obtain. Moreover, no evidence obtained of any relations between the magnitude and direction of weightings differential percentages and load processing capacity as assessed in terms of the capacity coefficient, C (Equation 5), in either Experiment 5 (TOEs) or Experiment 6 (SOEs). Consequently, while the relative attenuation of weightings asymmetries underlying SOEs, as compared to those obtained in analysis of TOEs, may be partly due to the decreased memorial demands of the discrimination task, Experiments 5 and 6 provide no evidence to suggest that interindividual changes in participants' load processing capacity, as distinct from constant-load processing capacity, arise as a result of the differential weighting of sensory events over time- and space-orders. [7] In this respect, further detailed examination of perceptual-cognitive mechanisms involved in the timed brightness and size discrimination of paired stimuli might rewardingly form a part of future investigations. Certain clues as to how one might proceed are laid out in the General Discussion.

General Discussion

The present experiments were conducted to examine systematic asymmetries that arise in magnitude discrimination of paired visual stimuli, using measures based on both response probabilities and RTs. A further aim has been to determine the extent to which changes in the magnitude and direction of TOEs and SOEs are related to participants' capacity to process the stimuli. Such experimental work follows naturally on from arguments (Hellström, 1979, 1989, 2003) that TOEs and SOEs reflect the operation of processes designed to maximize signal-to-noise ratios and optimize stimulus discrimination in paired stimulus comparison, and from the arguments of common sequential sampling and diffusion models (Link, 1975, 1992; Ratcliff, 1978, 2002) that directly tie choice proportions to RTs.

To this end both response probability and RTs to paired visual stimuli differing along the dimensions of brightness and size have been examined under conditions in which the stimuli were presented sequentially and separated by a time interval, and under conditions in

which the same visual stimuli were presented simultaneously and separated spatially. On the

basis of Link's sequential sampling model (Link, 1975, 1978, 1992; Link & Heath, 1975)

response probability was scaled in terms of logit *P,* and RTs in terms of *SRS*. These measures

were subsequently found to be linearly related, providing support for random walk models

which posit the accumulation of noisy information about the difference between stimulus

values over time (Link, 1975, 1992; Link & Heath, 1975; Ratcliff, 1978, 2002). Moreover,

both measures of logit *P* and *SRS* were found to yield TOEs and SOEs and, analyzed in terms

of Hellström's SW model, gave rise to similar patterns of weightings and weightings

differential percentages, providing support for the view that sensation magnitudes are

systematically weighted in the comparison process. In addition, participants' relative

constant-load capacity to process the stimuli was found to be related to their individual

magnitude and direction of weightings differential percentages. Yet, current assessment of the

faithfulness of stimulus representations, by correlating the response measures (logit *P*, *SRS*)

with the physical difference between paired stimuli, in relation to the magnitude and direction

of weightings asymmetries, suggests that interindividual differences in the magnitude and

direction of TOEs and SOEs may, at least in part, represent a trade-off between accuracy, in

terms of optimality of discrimination, and efficiency of stimulus processing, in terms of faster

RTs. Consequently, in the comparison and discrimination of paired stimuli it appears that

paired stimulus magnitudes, presented over time- and space-orders, are systematically

weighted by inclusion of differential reference level information in the comparison process

for optimal discriminatory performance which, nonetheless, may be traded for increased

efficiency of stimulus processing by increased reliance on one or the other stimulus

magnitudes in each stimulus pair. The message is that processing efficiency, in terms of faster

RT, is quite different from discrimination optimization, and in timed discrimination of paired

stimuli participants may choose to attend more to the magnitude of the first stimulus in each

stimulus pair for increased efficiency of stimulus processing.

In the present study, the primary focus has been on sensation weighting as a function of temporal and spatial order relations. However, Hake et al. (1966) provide some clues as to how one might proceed in determining maximal signal to noise ratios through optimized *V*-weighting of multidimensional stimuli. In particular, Hake et al. discuss discrimination of two-dimensional stimuli with optimized weighting of changes in the two dimensions for noise reduction and optimal discrimination but, in contrast to the present approach, weighting according to temporal and spatial position is missing. Further research is required to examine multidimensional stimulus discrimination in detail, but the basic idea remains that changes in processing capacity arise with changes in the systematic weighting of sensory events, which depends on perceived variability and intercorrelations between stimuli, and on increased utilization of reference level information in the comparison process.

For paired stimuli presented successively and separated by a time interval, reducing the weight of the degraded memory trace of the firstly presented stimulus, and instead increasing the weight of its reference level, may increase the signal to noise ratio in the perception of changes in the difference between the stimuli. In regard to SOEs, scanning effects and hemispheric localization of function may similarly operate to increase the signal to noise ratio associated with changes in the perceived stimulus difference. According to the discrimination optimization model described, the optimized weighting, and thereby the relative magnitude and direction of TOEs and SOEs are determined by the relative dispersion of sensation magnitudes and reference levels, as well as their intercorrelations.

On the basis of theoretically inferred processes of adaptive perception, weightings analyses have considerable implications for behavioral diagnosis of cerebral dysfunction (Hellström & Almkvist, 1997; Hellström, Forssell, & Fernaeus, 1989; McIntosh, Schindler, Birchall, & Milner, 2005). For instance, Hellström and Almkvist (1997) applied a paired

comparisons task to patients with senile dementia, who showed a greatly reduced capability to discriminate tone durations separated temporally by ISIs of 2 and 4 seconds as compared to when the same stimuli were temporally separated by ISIs of 0.5 and 1 second; indicative of accelerated forgetting due to cognitive deterioration. More recently, McIntosh et al. (2005) describe an end-point weightings analysis of systematic asymmetries in horizontal line bisection, based on the regression of each participant's lateral response on the left and right endpoint locations of each line, which in cases of brain damage were found to be differentially weighted yielding a sensitive index of hemispatial neglect. [8] On this basis and on the basis of our current theoretical analysis, therefore, systematic asymmetries in paired line length comparisons may be linked directly to person-specific parameters such as the variability of stimulus-generated sensations, perceived intercorrelations between stimuli, and use of reference level information in the comparison process.

In conclusion, six experiments have been reported examining TOEs and SOEs in timed brightness and size discrimination of paired visual stimuli. We have shown consistently that logit *P* and *SRS* are linearly related, providing support for common random walk models that directly tie choice proportions to choice reaction times. Moreover, characteristic patterns of TOEs and SOEs were found to occur for both response probability and response speed in paired comparisons of the brightness and size of visual stimuli, and have been consistently found to map onto diffusion drift rates, by way of fitting the Ratcliff diffusion model. In addition, interindividual variability in participants' capacity to processes the stimuli was found to be associated with the direction and magnitude of weightings differentials, as determined on the basis of Hellström's (1979) SW model. On these grounds, processes of adaptive perception and discrimination optimization, as posited by Hellström (1986, 1989), find support from current theoretically driven analyses of both response probabilities and RTs in comparison and discrimination of the brightness and size of paired stimuli.

References

Allan, L. G. (1977). The time-order error in judgments of duration. *Canadian Journal of*

   Psychology, 31, 24-31.

Ashourian, P., & Loewenstein, Y. (2011). Bayesian inference underlies the contraction bias in

   delayed comparison tasks. *PLoS ONE, 6(5), e19551. doi:10.1371/*

   journal.pone.0019551.

Audley, R. J. (1960). A stochastic model for individual choice behavior. *Psychological*

   Review, 67, 1-15.

Audley, R. J. & Wallis, C. P. (1964). Response instructions and the speed of relative

   judgements: I. Some experiments on brightness discrimination. *British Journal of*

   Psychology, 55, 59-73.

Babcock, H. W. (1953). The possibility of compensating astronomical seeing. *Publications of*

   the Astronomical Society of the Pacific, 65, 229-236.

Barrett, A. M., Buxbaum, L. J., Coslett, H. B., Edwards, E., Heilman, K. M., Hillis, A. E.,

   Milberg, W.P., & Robertson, I.H. (2006). Cognitive rehabilitation interventions for

   neglect and related disorders: Moving from bench to bedside in stroke patients. *Journal*

   of Cognitive Neuroscience, 18, 1223-1236.

Bartlett, R. J. (1939). Measurement in psychology. *Advancement of Science, 1*, 422-441.

Batchelder, W. H., Hu, X., & Smith, J. B. (2009). Multinomial processing tree models for

   discrete choice. *Zeitschrift für Psychologie / Journal of Psychology, 217*, 149 – 158.

Beckers, J. M. (1993). Adaptive optics for astronomy: principles, performance, and

   applications. *Annual Review of Astronomy and Astrophysics, 31*, 13–62.

Bentley, I. M. (1899). The memory image and its qualitative fidelity. *American Journal of*

   Psychology, 11, 1-48.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The

method of paired comparisons. *Biometrika, 39*, 324-345.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Brown, D. R. (1953). Stimulus similarity and the anchoring of subjective scales. *American Journal of Psychology, 66*, 199-214.

Cartwright, D. (1941). Relation of decision time to the categories of response. *American Journal of Psychology, 54*, 174-196.

Cartwright, D., & Festinger, L. (1943). A quantitative theory of decision. *Psychological Review, 56*, 595-621.

Charles, J., Sahraie, A., & McGeorge, P. (2007). Hemispatial asymmetries in judgment of stimulus size. *Perception & Psychophysics, 69*, 687-698.

Collett, D. (2002). *Modeling binary data*. Chapman & Hall, London.

Courcoux, P., Chaunier, L., DellaValle, G., Lourdin, D., & Semenou, M. (2005). Paired comparisons for the evaluation of crispness of cereal flakes by untrained assessors: correlation with descriptive analysis and acoustic measurements. *Journal of Chemometrics, 19*, 129-137.

Crossman, E. R. F. W. (1953). Entropy of choice time: the effect of frequency unbalance on choice response. *Quarterly Journal of Experimental Psychology, 5*, 41-51.

Crossman, E. R. F. W. (1955). The measurement of discriminability. *The Quarterly Journal of Experimental Psychology, 7*, 176-195.

Davidson, R. R., & Beaver, R. J. (1977). On extending the Bradley-Terry model to incorporate within-pair order effects. *Biometrics, 33*, 693-702.

Diederich, A., & Busemeyer, J. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology, 47*, 304-322.

Englund, M. P., & Patching, G. R. (2009). An inexpensive and accurate method of measuring

the force of responses in reaction time research. *Behavior Research Methods, 41*, 1254-1261.

Fechner, G. T. (1860). *Elemente der Psychophysik* [Elements of psychophysics]. Leipzig, Germany: Breitkopf & Härtel.

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

Guilford, J. P. (1954). *Psychometric methods* (2nd edition). New York: McGraw-Hill.

Gulliksen, H. (1958). Comparatal dispersion, a measurement of accuracy of judgment. *Psychometrika, 23*, 137-150.

Hake, H. W., Rodwan, A., & Weintraub, D. (1966). Noise reduction in perception. In K. R. Hammond (Ed.), *The psychology of Egon Brunswik* (pp. 277-316). New York: Holt.

Hellström, Å. (1978). Factors producing and factors not producing time errors: An experiment with loudness comparisons. *Perception & Psychophysics, 23*, 433-444.

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. (1986). Sensation weighting in comparing: A tool for optimizing discrimination. In B. Berglund, U. Berglund, & R. Teghtsoonian (Eds.), *Fechner Day 1986: Proceedings of the Second Annual Meeting of the International Society for Psychophysics* (pp. 89-94). Stockholm: International Society for Psychophysics.

Hellström, Å. (1989). What happens when we compare two stimuli? In G. Ljunggren & S. Dornic (Eds.) *Psychophysics in Action* (pp. 25-39). Berlin: Springer.

Hellström, Å. (2000). Sensation weighting in comparison and discrimination of heaviness. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 6-17.

Hellström, Å. (2003). Comparison is not just subtraction: Effects of time- and space-order on subjective stimulus difference. *Perception & Psychophysics, 65*, 1161-1177.

Hellström, Å. (2008). Time- and space-order effects in stimulus comparison in the light of response-time data. In Schneider, B. A., Ben-David, B. M., Parker, S., & Wong, W. (Eds.). *Fechner Day 2008. Proceedings of the 24th Annual Meeting of the International Society for Psychophysics* (pp. 139-144). Toronto: International Society for Psychophysics.

Hellström, Å., & Almkvist, O. (1997). Tone duration discrimination in demented, memory-impaired, and healthy elderly. *Dementia and Geriatric Cognitive Disorders, 8*, 49-54.

Hellström, Å., Forssell, L. G., & Fernaeus, S. E. (1989). Early stages of late onset Alzheimer's disease. V. Psychometric evaluation of perceptual/cognitive processes. *Acta Neurologica, 79* (Suppl. 21), 87-92.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Henmon, V. A. C. (1906). The time of perception as a measure of differences in sensations. *Archives of Philosophy, Psychology and Scientific Methods, 8*, 3–75.

Henmom, V. A. C. (1911). The relation of the time of a judgment to accuracy. *American Journal of Psychology, 18*, 86-210.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11–26.

Holzman, P. S. (1954). The relation of assimilation tendencies in visual, auditory, and kinesthetic time-errors to cognitive attitudes of leveling and sharpening. *Journal of Personality, 22*, 375-394.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*, 188-196.

Inomata, S. (1959). The analysis of conditions behind the time-error in the successive

comparison of visual length. *Japanese Psychological Research, 7*, 36-54.

Jamieson, D. G., & Petrusic, W. M. (1975a). Presentation order effects in duration

discrimination. *Perception & Psychophysics, 17*, 197-202.

Jamieson, D. G., & Petrusic, W. M. (1975b). Pairing effects and time-order errors in duration

discrimination. *Perception & Psychophysics, 18*, 107-113.

John, I. D. (1975). A common mechanism mediating the time-order error and the cross-over

effect in comparative judgments of loudness. *Australian Journal of Psychology, 27*, 51-

60.

Kellogg, W. M. (1931). The time of judgment in psychometric measures. *American Journal

of Psychology, 241*, 1-52.

Köhler, W. (1923). Zur Analyse des Sukzessivvergleichs und der Zeitfehler [On the analysis

of successive comparison and time-order error]. *Psychologische Forschung, 4*, 115-175.

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A

commentary. *Perception & Psychophysics, 63*, 1421-1455.

Kreezer, G. (1938). The neurological level of the factors underlying time-errors. *American

Journal of Psychology, 51*, 18-43.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. Academic Press,

London.

Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in

duration discrimination tasks: A comparison of the 2AFC and the reminder task.

Perception & Psychophysics, 70, 291-305.

Lauenstein, G. (1933). Ansats zu einer physiologischen Theorie des Vergleichs und der

Zeitfehler [Attempt at a physiological theory of comparison and time-order errors].

Psychologische Forschung, 17, 130-177.

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models.

Journal of Mathematical Psychology, 55, 1-7.

Link, S. W. (1975). The relative judgment theory of two choice response time. *Journal of Mathematical Psychology, 12*, 114-115.

Link, S. W. (1978). The relative judgment theory of the psychometric function. In J. Requin (Ed.), *Attention & Performance VII*. (pp. 619-630). Hillsdale, NJ, USA: Erlbaum.

Link, S. W. (1992). *The wave theory of difference and similarity*. Hillsdale, NJ: Erlbaum.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*, 77-105.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS -- A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing, 10*, 325-337.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

Maeda, H. (1959). On the inhibitory effects of extrapolated stimuli in the successive comparison of brightness (I). *Japanese Journal of Psychology, 30*, 3-20.

Maeda, H. (1960). On the inhibitory effects of extrapolated stimuli in the successive comparison of brightness (II). *Japanese Journal of Psychology, 31*, 9-20.

Masin, S. C., & Agostini, A. (1991). Attentional scanning and space errors. *Perception & Psychophysics, 50*, 285-289.

Mattingley, J. B., Bradshaw, J. L., Nettleton, N. C., & Bradshaw, J. A. (1994). Can task specific perceptual bias be distinguished from unilateral neglect? *Neuropsychologia, 32*, 805-817.

McIntosh, R. D., Schindler, I., Birchall, D., & Milner, A. D. (2005). Weights and measures: a new look at line bisection behaviour in neglect. *Cognitive Brain Research, 25*, 833-850.

Michels, W. C., & Helson, H. (1954). A quantitative theory of time order effects. *American Journal of Psychology, 67*, 327-334.

Nachmias, J. (2006). The role of virtual standards in visual discrimination. *Vision Research, 46*, 2456-2464.

Needham, J. G. (1934). The time-error as a function of continued experimentation. *American Journal of Psychology, 67*, 558-567.

Needham, J. G. (1935). The effect of the time interval upon the time-error at different intensive levels. *Journal of Experimental Psychology, 18*, 539-543.

Ono, S. (1949). On the positive time error in the successive comparison of brightness. *Japanese Journal of Psychology, 20*, 6-15.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision, 5*, 376-404.

Patching, G. R., Englund, M. P., & Hellström, Å. (2008). Time- and space-order effects in timed brightness discrimination of paired visual stimuli. In B. A. Schneider, B. M. Ben-David, S. Parker, & W. Wong, (Eds.). *Fechner Day 2008. Proceedings of the 24th Annual Meeting of the International Society for Psychophysics* (pp. 283-288). Toronto: The International Society for Psychophysics.

Peak, H. (1940). The time order error in successive judgments and in reflexes. III. Time error theories. *Psychological Review, 47*, 1-20.

Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 433-436.

Plant, R. R., & Hammond, N. V. (2002). Towards an experimental timing standards lab: Benchmarking precision in the real world. *Behavior Research Methods, Instruments, & Computers, 34*, 218-226.

Postman, L. (1946). The time error in auditory perception. *American Journal of Psychology, 59*, 193-219.

Postman, L. (1947). Retroactive inhibition and psychophysical judgment. *American Journal*

of Psychology, 10, 367-377.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59-108.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness

discrimination task: Fitting real data and failing to fit fake but plausible data.

Psychonomic Bulletin & Review, 9, 278-291.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-

choice decision tasks. *Neural Computation, 20*, 873-992.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64*, 153-181.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing

systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200–239).

Hillsdale, NJ: Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological

processes*. Cambridge, England: Cambridge University Press.

Townsend, J. T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for

response time measures of efficiency as workload is varied. *Psychonomic Bulletin &

Review, 18*, 659-681.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception:

An investigation of parallel, serial, and coactive theories. *Journal of Mathematical

Psychology, 39*, 321-359.

Townsend, J. T., & Nozawa, G. (1997). Serial exhaustive models can violate the race model

inequality: Implications for architecture and capacity. *Psychological Review, 104*, 595-

602.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New

capacity measures and predictions for a response time inequality series. *Psychological

Review, 111*, 1003-1035.

Tresselt, M. R. (1944a). The time-errors in visual extent and areas. *Journal of Psychology, 17*,

21-30.

Tresselt, M. R. (1944b). Time errors in successive comparison of simple visual objects.

American Journal of Psychology, 57, 555-558.

Tresselt, M. R. (1948). Time errors in successive comparison of tonal pitch. *American

Journal of Psychology, 61*, 335-342.

Ulrich, R., & Vorberg, D. (2009). Estimating the difference limen in 2AFC tasks: Pitfalls and

improved estimators. *Attention Perception & Psychophysics, 71*, 1219-1227.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to

experimental data. *Psychonomic Bulletin & Review, 14*, 1011-1026.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for

two-choice response times. *Psychological Methods, 16*, 44-62.

Vidotto, G., Vicario, G. B., & Tomat, L. (1996). Length estimation of evolving lines. In

*Fechner Day 1996. Proceedings of the 12th Annual Meeting of the International Society

for Psychophysics* (pp. 411-415). Padua: The International Society for Psychophysics.

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff

diffusion model of response times and accuracy. *The European Journal of Cognitive

Psychology, 21*, 641-671.

Wenger, J. M., & Gibson, B. (2004). Using hazard functions to assess changes in processing

capacity in an attentional cuing paradigm. *Journal of Experimental Psychology: Human

Perception and Performance, 30*, 708-719.

Wenger, J. M., & Townsend, J. T. (2000). Basic response time tools for studying general

processing capacity in attention, perception and cognition. *Journal of General

Psychology, 127*, 67-99.

Wickelmaier, F., & Schmid, C. (2004). A Matlab function to estimate choice model

parameters from paired-comparison data. *Behavior Research Methods, Instruments &*
Computers, 36, 29-40.

Woodrow, H. (1933). Weight discrimination with a varying standard. *American Journal of*
Psychology, 45, 391-416.

Woodrow, H. (1935). The effect of practice upon time-order errors in the comparison of
temporal intervals. *Psychological Review, 42*, 127-153.

Woodrow, H., & Stott, L. H. (1936). The effect of practice on positive time-order errors.
Journal of Experimental Psychology, 19, 694-705.

Woodruff, B., Jennings, D. L., & Rico, N. L. (1975). Time error in lifted weights as affected
by presentation order and judgment mode. *Perception & Psychophysics, 18*, 98-104.

Yeshurun, Y., Carrasco, M., & Maloney, L. T. (2008). Bias and sensitivity in two-interval
forced choice procedures: Tests of the difference model. *Vision Research, 48*, 1837-
1851.

Footnotes

[1]Note that when the second or right stimulus is held constant the constant error (CE) is calculated as CE = St - PSE.

[2]For ease of exposition all measurements of the size of the visual stimuli are stated in mm. In every case participants viewed the video monitor from 57 cm, so 10 mm corresponds to 1 degree of visual angle.

[3]In the present analyses U is treated merely as an additive constant required to render the mean of d_{12} independent of the magnitudes of W_1 and W_2, but see Hellström (1979, 2003) for more detailed treatment of U. Note that U may bear relation to the zero point of drift rate as discussed by Ratcliff and McKoon (2008), but as an additive parameter the zero point for drift rate is not considered equivalent to subjective reference levels, ψ_{r1}, and ψ_{r2}, because changes in the zero point of drift rate fails to optimize discriminatory performance.

[4]The logged physical values of the stimuli were standardized for comparative purposes by calculation of $z = (x - \mu) / \sigma$, where z is the standardized value, x the natural logarithm of the stimulus value in physical units, μ the arithmetic mean of the logged stimulus series, and σ the standard deviation.

[5]Given the exploratory nature of this analysis it made no sense to examine relations between $WD\%$ and derived hazard ratios separately for the stimuli at each ISI.

[6]In both Experiments 5 and 6, the response keys were secured firmly in position to

avoid participants inadvertently realigning the response keys with the stimulus display.

[7]Further analyses of relations between the V-weightings and constant-load processing capacity, as assessed in terms of HR, and load processing capacity, as assessed in terms of the capacity coefficient C, by way of 2nd- and 3rd-order polynomial regressions failed to reveal any statistically reliable relations (all ps > .05).

[8]Hemispatial neglect is a debilitating disorder that arises as a result of damage to the right cerebral hemisphere, although it can also result from damage to the left cerebral hemisphere. It is defined by asymmetric perception and action that cannot be attributed to primary motor or sensory dysfunction (Barrett, Buxbaum, Coslett, et al., 2006).

Figure captions

Figure 1. Semi-factorial combination of stimuli used in Experiment 1. The black squares show the stimulus pairings; see the text for details. The same design was used in Experiment 2, only in this case stimulus luminance was held constant and the size of the stimuli systematically varied.

Figure 2. Logit *P* versus mean *SRS* by ISI, for panel (a) paired discrimination of brightness (Experiment 1) and panel (b) paired discrimination of size (Experiment 2). To show each linear fit clearly the data obtained for stimuli separated by an ISI of 800 ms are displaced by 0.5 logit steps; those separated temporally by an ISI 1600 ms are displaced by 1.0 logit steps, and for paired stimuli separated temporally by an ISI of 3200 ms the data are displaced by 1.5 logit steps.

Figure 3. Logit *P* and mean *SRS* for the five average magnitude levels of the 25 pairs of visual stimuli presented successively and separated by a time interval. Panels (a) and (b) show logit *P* and mean *SRS* for discrimination of brightness (Experiment 1). Panels (c) and (d) show logit *P* and mean *SRS* for discrimination of size (Experiment 2). The error bars show standard error of the mean.

Figure 4. Absolute fit of the diffusion model. Panels (a) and (c) show logit *p* for the five average magnitude levels of the 25 paired visual stimuli presented successively and separated by a time interval collapsed over ISIs, for Experiments 1 and 2, respectively. Panels (b) and (d) show the RT distribution integrated hazard functions $H_B(t)$ for discrimination of brightness (Experiment 1) and $H_S(t)$ for discrimination of size (Experiment 2), respectively, again

collapsed over ISIs. In each panel, the dashed black line shows the simulated data as computed over participants with the vertical bars to illustrate the minimum and maximum predicted for participants. The solid black line shows the empirically derived data averaged over participants, with logit P and $H_B(t)$ (Experiment 1) or $H_S(t)$ (Experiment 2) for each participant shown as grey lines.

Figure 5. Relations between $WD\%$ and (a and c) constant-load processing capacity and (b and d) faithful stimulus representation, as determined for each participant, on the basis of the empirical data obtained in Experiments 1 and 2. Panels (a) and (c): integrated hazard ratio (HR) of relative constant-load processing capacity versus weightings differential percentage ($WD\%$), for Experiment 1 (panel a) and for Experiment 2 (panel c). Panels (b) and (d): the correlation (Pearson's r), of the reported (logit P, SRS) and physical difference between the paired stimuli versus weightings differential percentages ($WD\%$), for Experiment 1 (panel b) and for Experiment 2 (panel d).

Figure 6. Logit P versus mean SRS by spatial separation, for panel (a) paired discrimination of brightness (Experiment 3) and panel (b) paired discrimination of size (Experiment 4). To show each linear fit clearly the data obtained for stimuli separated by a spatial separation of 20 mm are displaced by 0.5 logit steps; those separated spatially by 40 mm are displaced by 1.0 logit steps, and for paired stimuli separated spatially by 80 mm the data are displaced by 1.5 logit steps.

Figure 7. Logit P and mean SRS for the five average magnitude levels of the 25 paired visual stimuli presented simultaneously and separated spatially. Panels (a) and (b) show logit P and mean SRS for discrimination of brightness (Experiment 3). Panels (c) and (d) show logit P

and mean *SRS* for discrimination of size (Experiment 4). The error bars show standard error

of the mean.

Figure 8. Absolute fit of the diffusion model. Panels (a) and (c) show logit *P* for the five

average magnitude levels of the 25 pairs of visual stimuli presented simultaneously and

separated spatially collapsed over spatial separations between stimuli, for Experiments 3 and

4, respectively. Panels (b) and (d) show the RT distribution integrated hazard functions, $H_B(t)$,

for discrimination of brightness (Experiment 3) and $H_S(t)$ for discrimination of size

(Experiment 4), again collapsed over spatial separations between the stimuli. In each panel,

the dashed black line shows the simulated data as computed over participants with the vertical

bars to illustrate the minimum and maximum predicted for participants. The solid black line

shows the empirically derived data computed over participants, with logit *P* and $H_B(t)$

(Experiment 4) or $H_S(t)$ (Experiment 5) for each participant shown as grey lines.

Figure 9. Logit *P* versus mean *SRS* for the 21 stimulus pairs differing along the physical

dimension of either luminance (6 stimulus pairs) or size (6 stimulus pairs). To show each

linear fit clearly the data obtained for Experiment 6 are displaced by 1 logit step.

Figure 10. Logit *P* and mean *SRS* for the three average physical magnitude levels of the

stimuli used in Experiments 5 and 6. Panels (a) and (b) show logit *P* and mean *SRS* for the

paired stimuli presented successively and separated by a time interval (Experiment 5). Panels

(c) and (d) show logit *P* and mean *SRS* for the paired stimuli presented simultaneously and

separated by a spatial interval (Experiment 6). The error bars show standard error of the mean.

Figure 11. Absolute fit of the diffusion model. Panels (a) and (c) show logit *P* for the three

average magnitude levels of the stimuli used in Experiments 5 and 6, respectively. Panels (b) and (d), respectively, show the RT distribution integrated hazard functions, H(t), for the successively presented stimuli separated by a time interval (Experiment 5), and for the simultaneously presented stimuli separated by a spatial interval (Experiment 6). In each panel, the dashed black line shows the simulated data as computed over participants with the vertical bars to illustrate the minimum and maximum predicted for participants. The solid black line shows the empirically derived data averaged over participants, with logit P and H(t) for each participant shown as grey lines.

Figure 12. Relations between $WD\%$ and (a) constant-load processing capacity and (b) faithful stimulus representation, as determined for each participant, on the basis of the empirical data obtained in Experiment 5. Panel (a): integrated hazard ratio (HR) of relative constant-load processing capacity versus weightings differential percentage ($WD\%$). Panel (b): the correlation (Pearson's r), between reported stimulus difference (logit P, SRS) and difference between the V-weighted sums of the two physical stimulus values, versus weightings differential percentage ($WD\%$).

Figure 1

Figure 2

Figure 3

(a)

(b)

(c)

(d)
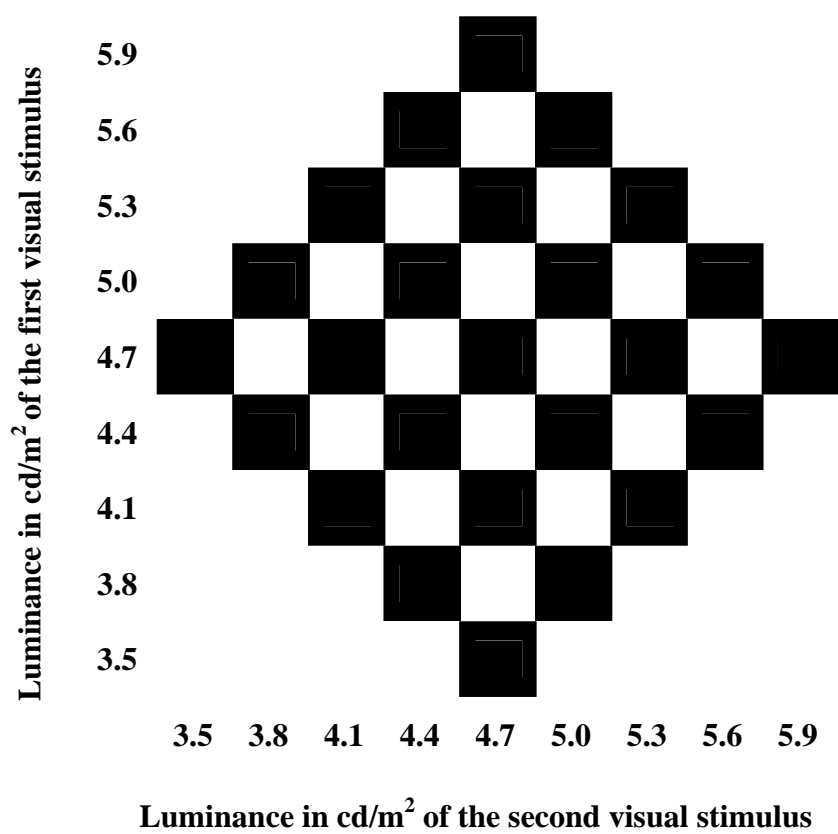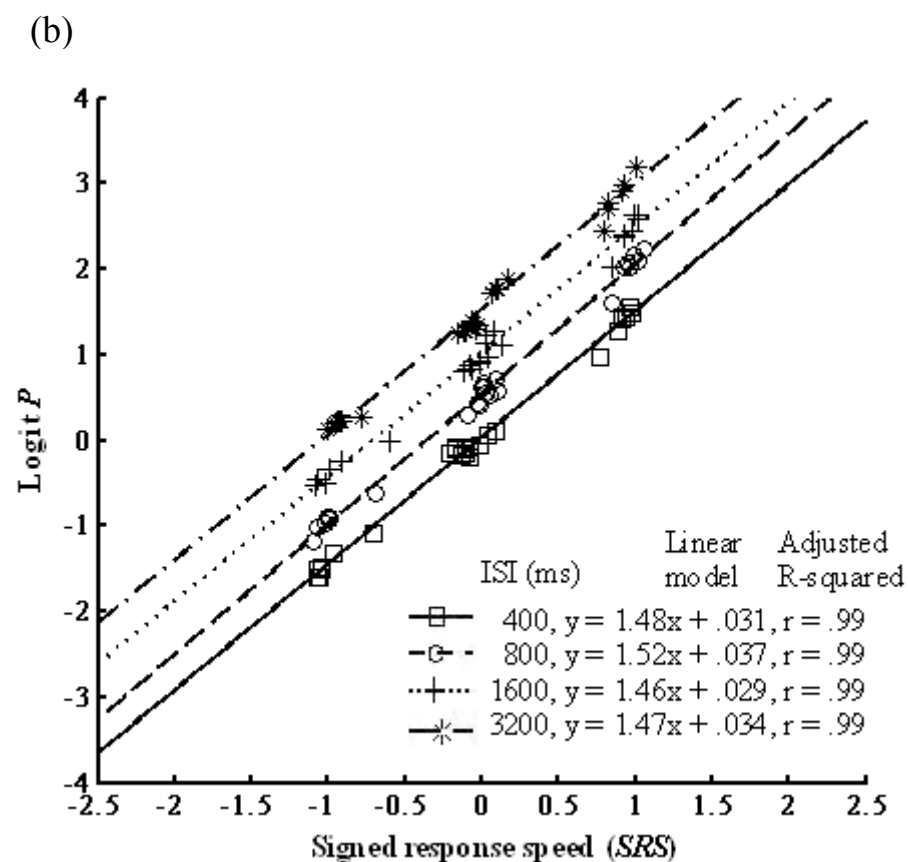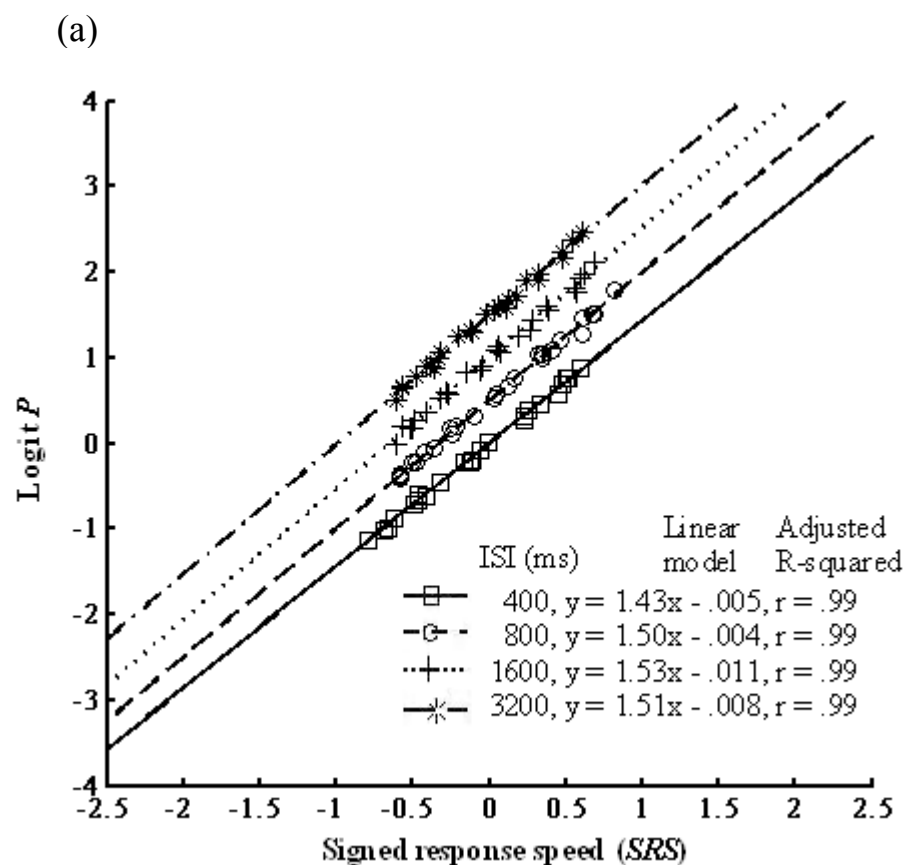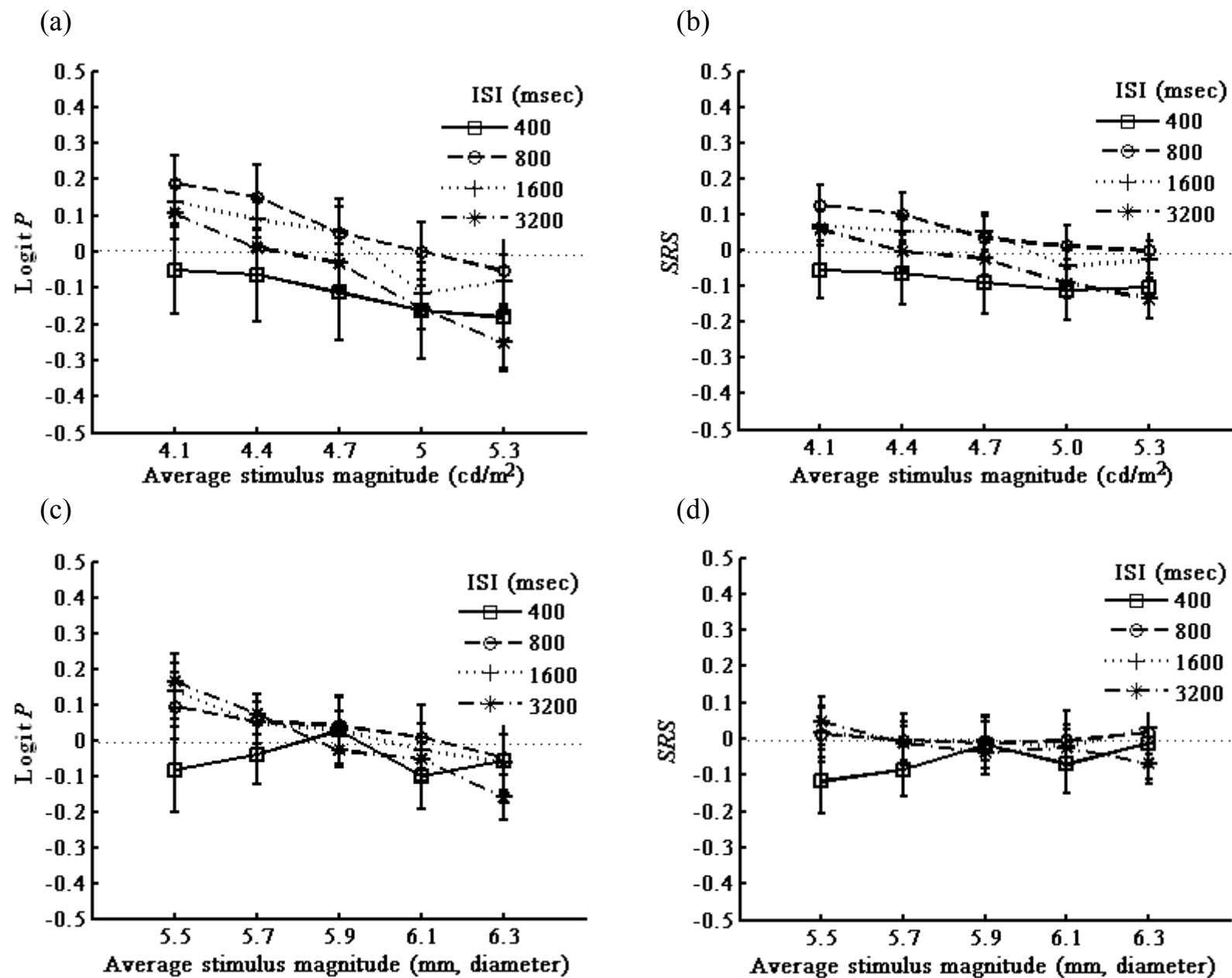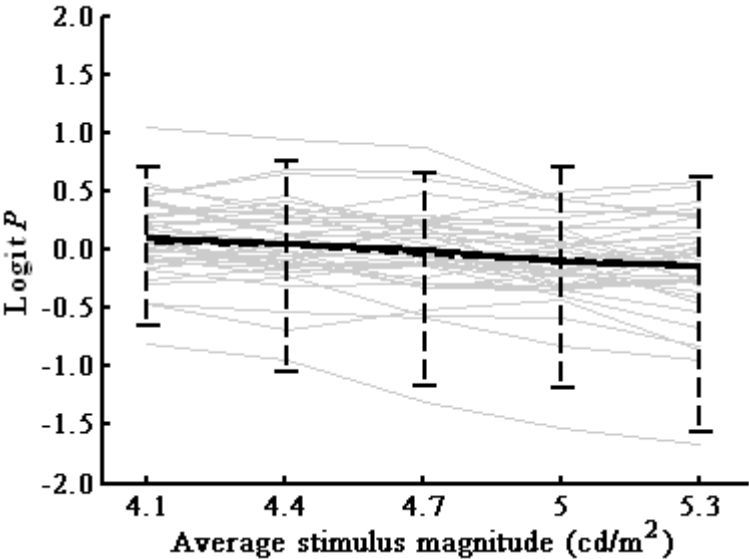
Figure 4   (a)

(b)
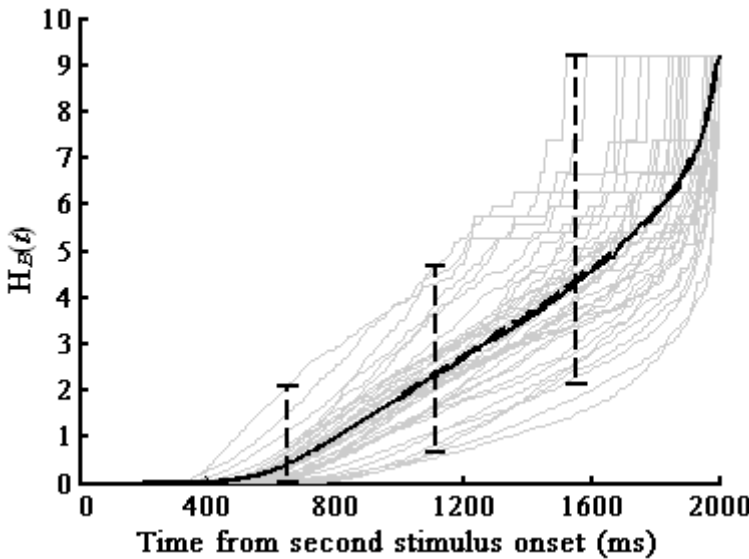
(c)

(d)

Figure 5

Figure 6

(a)

(b)
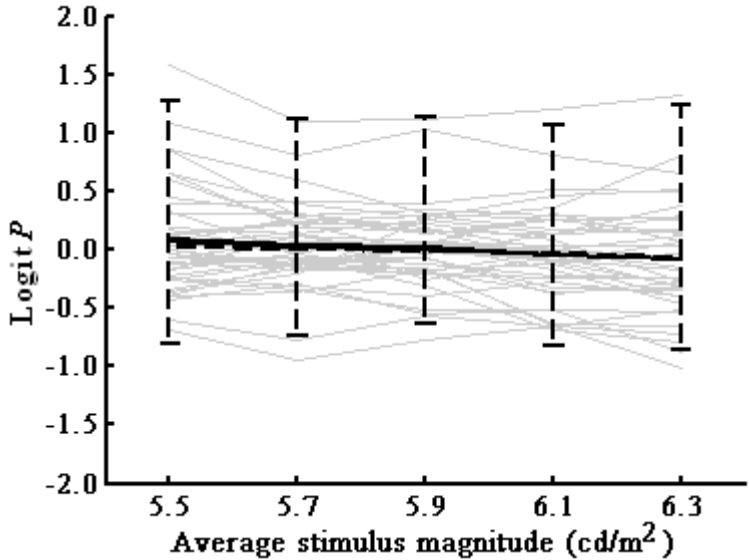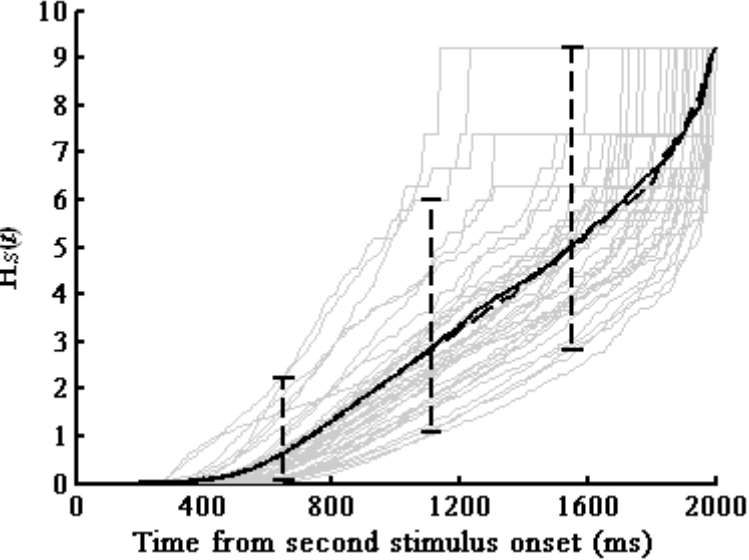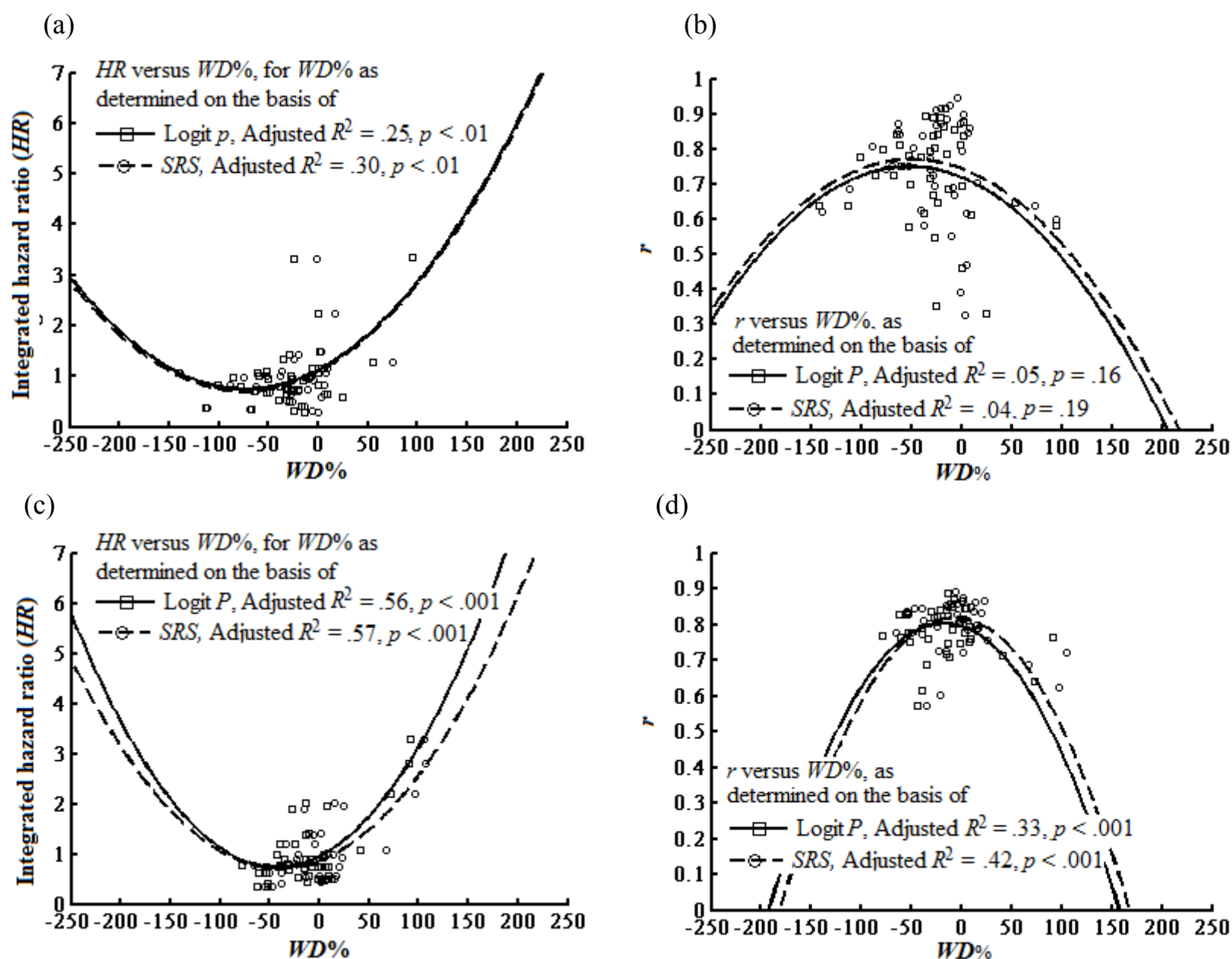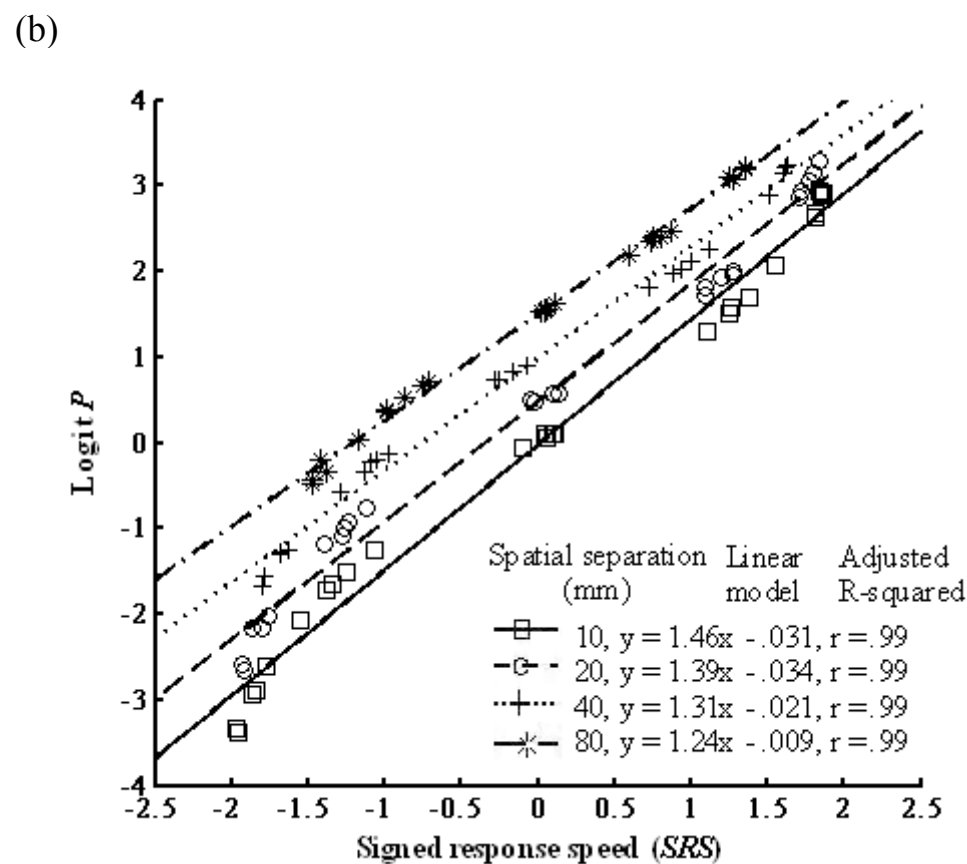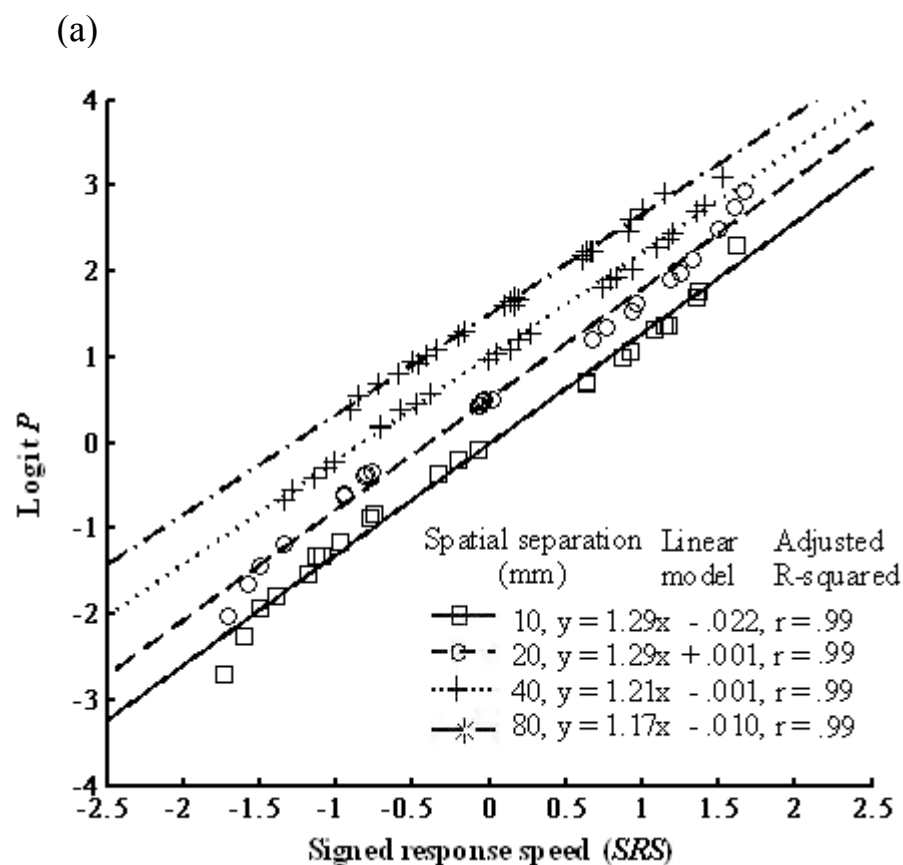
Figure 7   (a)

(b)

(c)

(d)

Figure 8

Figure 9

Figure 10   (a)   (b)
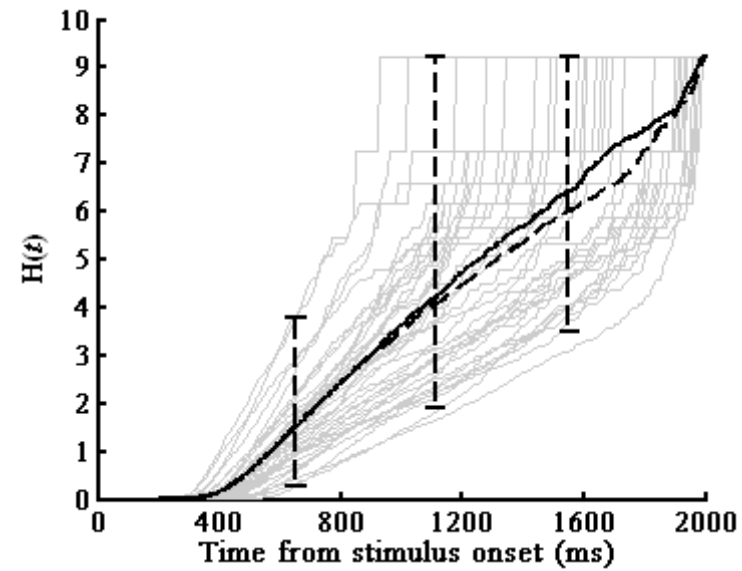
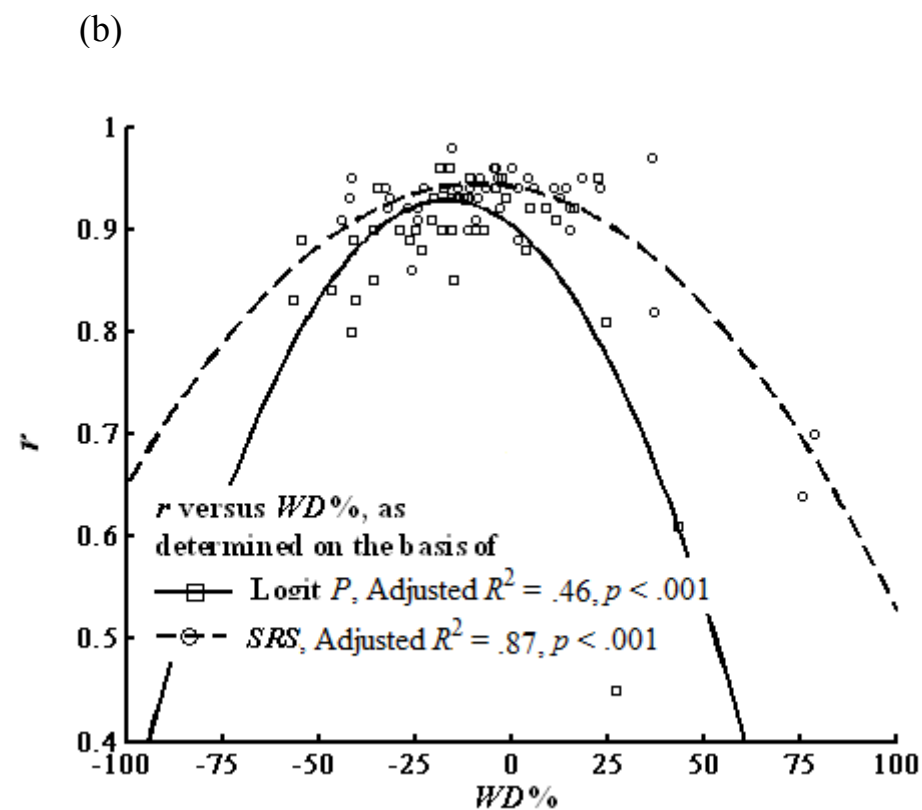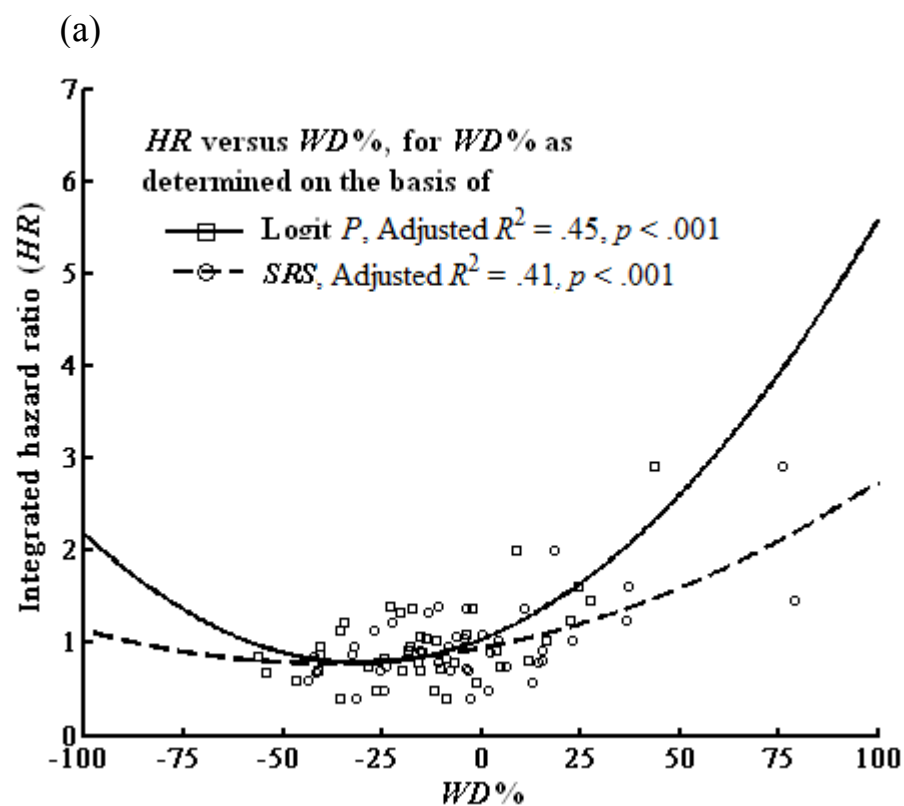**Figure 11** (a)

Figure 12

(a)

(b)
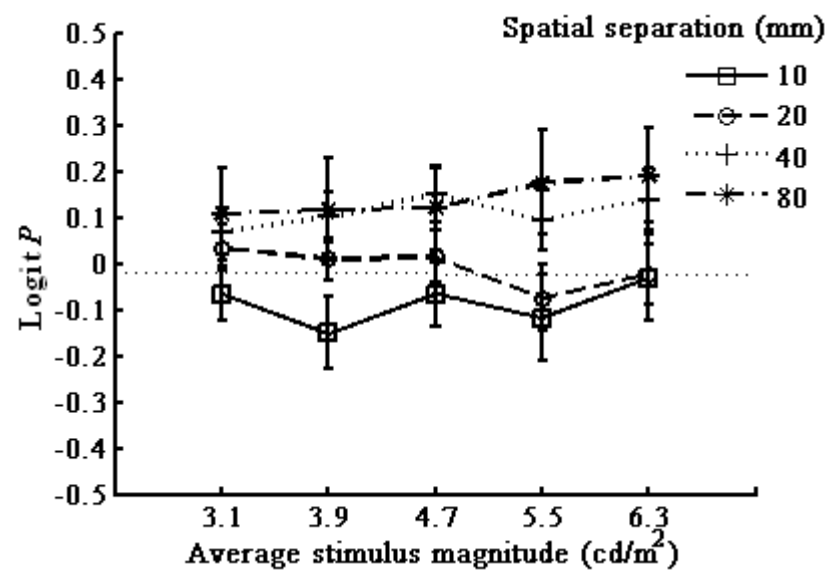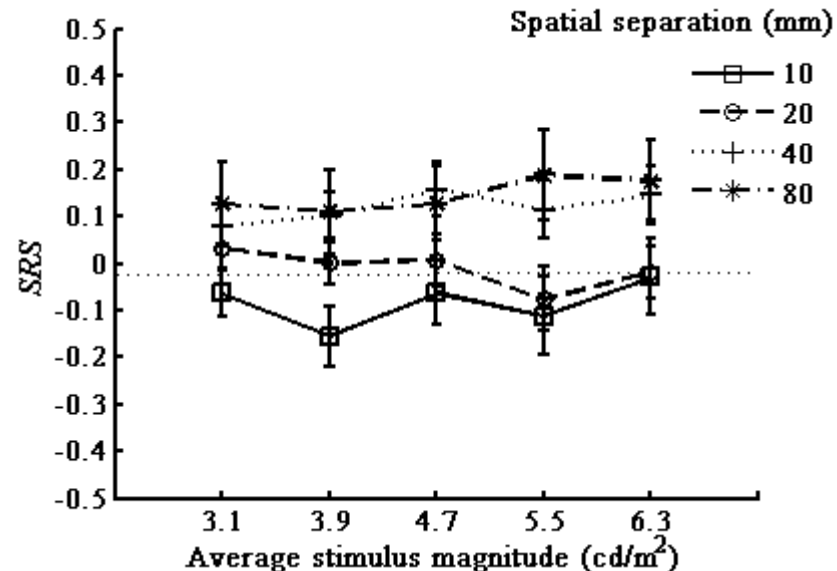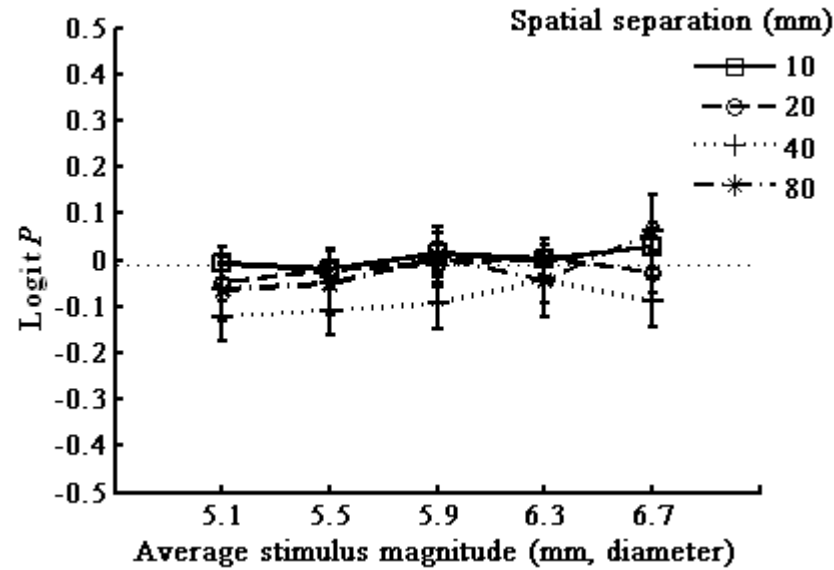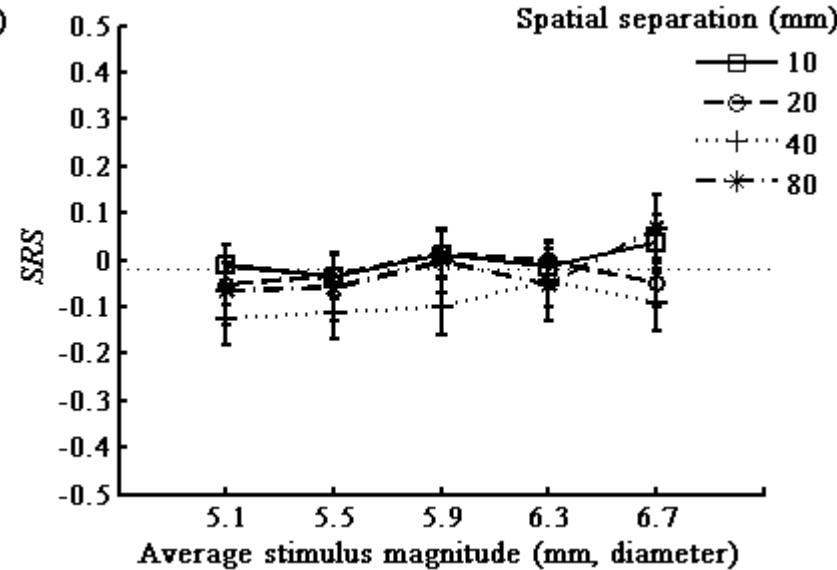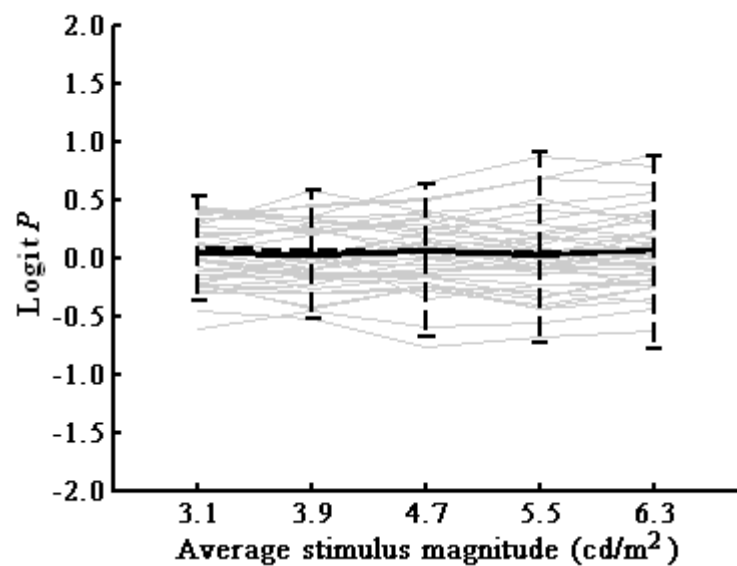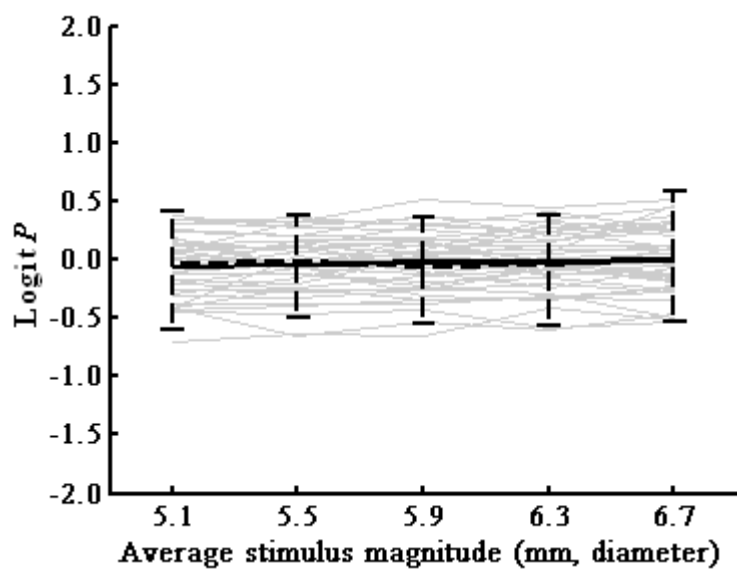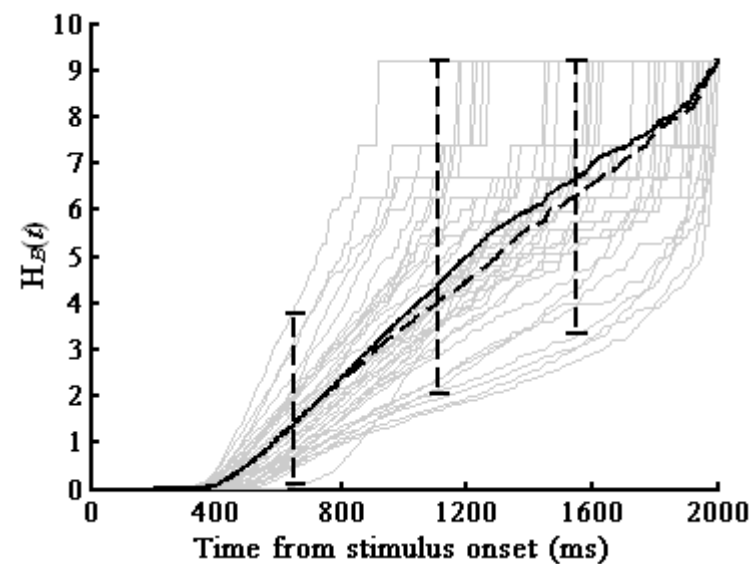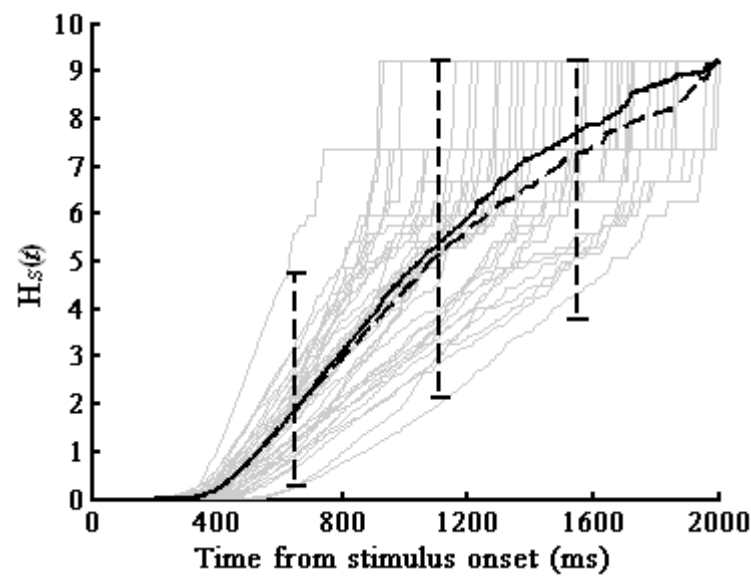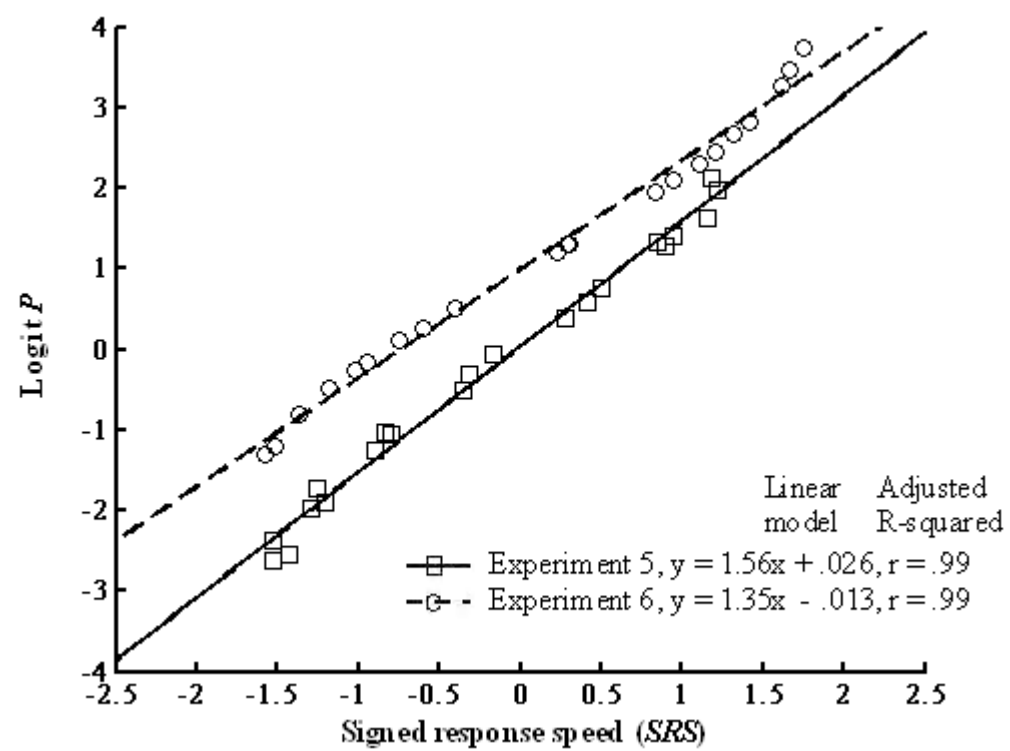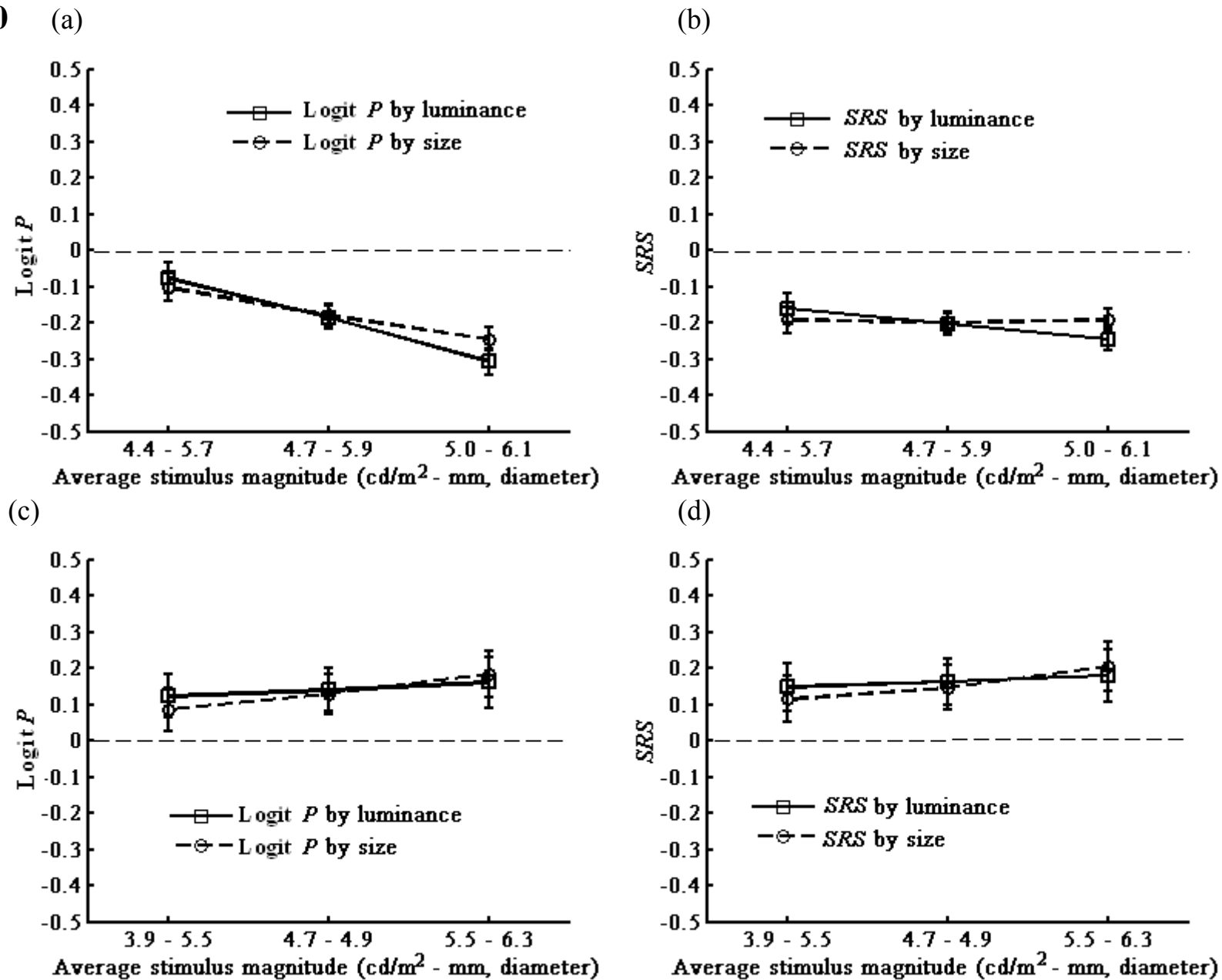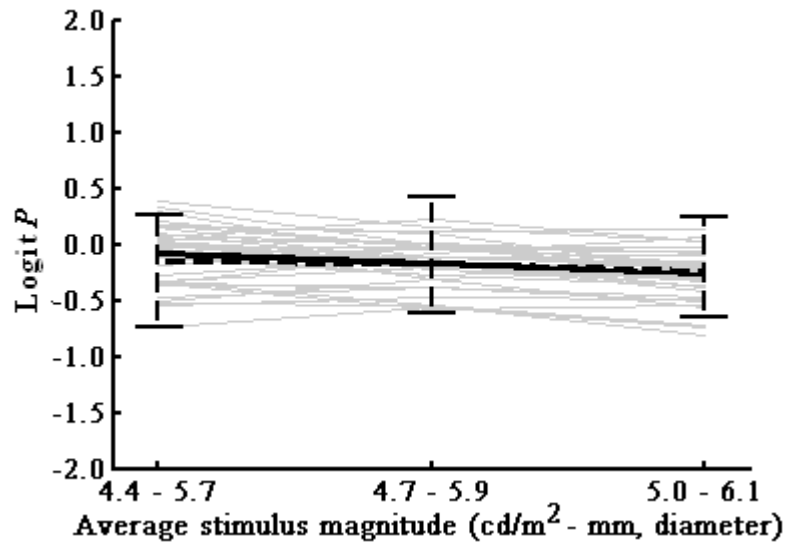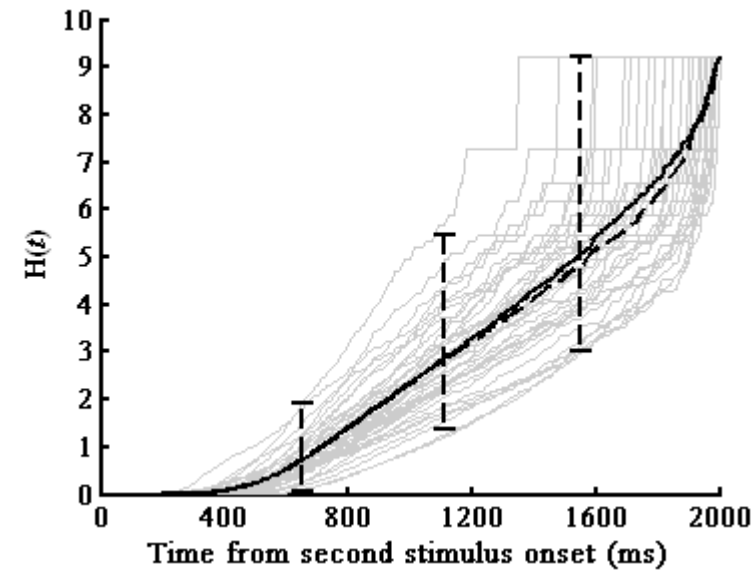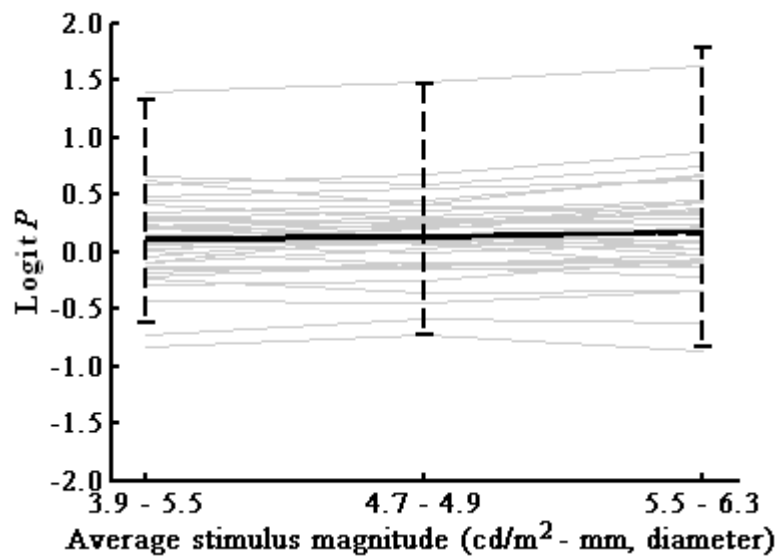
Table 1. Optimal values of s_1 and s_2, and gain percentage ($G\%$), that is, increase of S/N from using optimal s values instead of $s_1 = s_2 = 1$, according to the discrimination optimization model, under various realistic conditions of paired stimulus comparisons.

$\sigma(\psi_2)$	Optimal s_1	Optimal s_2	s_1 / s_2	$G\%$
Series 1: $\sigma(\psi_r) = 0.75$				
0.5	0.67	1.02	0.66	18.09
1	0.79	0.79	1.00	2.06
1.5	0.89	0.63	1.41	10.12
Series 2: $\sigma(\psi_r) = 1.5$				
0.5	0.82	1.01	0.81	10.77
1	0.88	0.88	1.00	2.06
1.5	0.95	0.77	1.23	7.41
Series 3: $\sigma(\psi_r) = 3.0$				
0.5	0.91	1.00	0.91	6.56
1	0.94	0.94	1.00	2.06
1.5	0.97	0.88	0.96	5.02

Common assumptions

$\sigma(\psi_1) = 1$; $\psi_r = \psi_{r1} = \psi_{r2}$; $\rho(\psi_1, \psi_2) = 0.5$; $\rho(\psi_1, \psi_{r1}) = -0.3$; $\rho(\psi_2, \psi_{r2}) = -0.3$;

$\rho(\psi_1, \psi_{r2}) = -0.2$; $\rho(\psi_2, \psi_{r2}) = -0.2$; $\rho(\psi_{r1}, \psi_{r2}) = 0.5$

Table 2. Mean weightings W_1 / W_2, along with their weightings differential percentage, $WD\%$, as determined on the basis of the response probability and RT data obtained in Experiments 1 and 2.

	Logit P				SRS			
ISI (ms)	400	800	1600	3200	400	800	1600	3200
Experiment 1. Brightness discrimination								
W_1 / W_2	0.51 / 0.61	0.47 / 0.63	0.41 / 0.57	0.32 / 0.53	0.30 / 0.33	0.29 / 0.36	0.27 / 0.34	0.22 / 0.32
$WD\%$	-17.86	-29.09	-32.65	-51.16	-9.52	-21.54	-22.95	-37.04
p value for $WD\% \neq 0$	n.s. (.059)	.006	.003	< .001	n.s.	n.s.	.012	< .001
Experiment 2. Size discrimination								
W_1 / W_2	0.75 / 0.76	0.72 / 0.81	0.66 / 0.80	0.57 / 0.78	0.49 / 0.44	0.47 / 0.47	0.47 / 0.47	0.42 / 0.47
$WD\%$	-1.32	-11.76	-19.18	-31.11	10.75	0.00	0.00	-11.24
p value for $WD\% \neq 0$	n.s.	n.s.	n.s. (.079)	< .001	n.s. (.079)	n.s.	n.s.	.028

Note: n.s., not significant ($p > .05$); p values between .05 and .10 are given in parentheses.

Table 3. Mean weightings W_1 / W_2, and their weightings differential percentage, $WD\%$, as determined on the basis of the Ratcliff diffusion model fit to the alternate responses and RTs to the paired successively presented stimuli in Experiments 1 and 2.

	Drift rate μ				Drift rate μ			
	Experiment 1. Brightness discrimination				Experiment 2. Size discrimination			
ISI (ms)	400	800	1600	3200	400	800	1600	3200
W_1 / W_2	.066 / .079	.062 / .080	.055 / .074	.046 / .070	.118 / .114	.111 / .120	.109 / 0.120	.097 / .116
$WD\%$	-17.93	-25.35	-29.46	-41.38	3.45	-7.79	-9.61	-17.84
p value for $WD\% \neq 0$	.010	.006	.004	<.001	n.s.	n.s.	n.s. (.103)	<.001

Note: n.s., not significant ($p > .05$); p values between .05 and .10 are given in parentheses.

Table 4. Mean weightings W_1 / W_2, along with their weightings differential percentage, $WD\%$, as determined on the basis of the response probability and RT data obtained in Experiments 3 and 4.

	Logit P				SRS			
Spatial separation (mm)	10	20	40	80	10	20	40	80
Experiment 3. Brightness discrimination								
W_1 / W_2	1.24 / 1.18	1.19 / 1.24	0.91 / 0.90	0.61 / 0.58	0.75 / 0.74	0.75 / 0.79	0.66 / 0.63	0.45 / 0.42
$WD\%$	4.96	-4.12	1.10	5.04	1.34	-5.19	4.65	6.90
p value for $WD\% \neq 0$	n.s.	n.s. (.097)	n.s.	n.s.	n.s.	n.s. (.057)	n.s.	n.s.
Experiment 4. Size discrimination								
W_1 / W_2	1.85 / 1.81	1.60 / 1.57	1.35 / 1.30	1.05 / 0.99	1.04 / 1.01	0.99 / 0.98	0.90 / 0.87	0.74 / 0.69
$WD\%$	2.19	1.89	3.37	5.88	2.93	1.02	3.39	6.99
p value for $WD\% \neq 0$	n.s.	n.s.	n.s. (.078)	.015	n.s.	n.s.	n.s.	.015

Note: n.s., not significant ($p > .05$); p values between .05 and .10 are given in parentheses.

Table 5. Mean weightings W_1 / W_2, and their weightings differential percentage, *WD*%, as determined on the basis of the Ratcliff diffusion model fit to the alternate responses and RTs to the paired simultaneously presented stimuli in Experiments 3 and 4.

	Drift rate μ				Drift rate μ			
	Experiment 3. Brightness discrimination				Experiment 4. Size discrimination			
Spatial separation (mm)	10	20	40	80	10	20	40	80
W_1 / W_2	.204 / .197	.210 / 0.214	.168 / .164	.109 / .103	.254 / .249	.240 / .238	.207 / .201	.156 / .143
WD%	3.49	-1.89	2.41	5.66	1.99	0.84	2.94	8.70
p value for *WD*% $\neq$ 0	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.004

Note: n.s., not significant ($p > .05$); *p* values between .05 and .10 are given in parentheses.

Table 6. Mean weightings, V_B / V_S and, W_1 / W_2 along with their weightings differential percentage, WD%, as determined on the basis of logit P and SRS for Experiments 5 and 6.

Mean order effect	Experiment 5: TOE		Experiment 6: SOE	
	Logit P	SRS	Logit P	SRS
V_B / V_S	.323 / .677	.294 / .706	.368 / .632	.344 / .656
W_1 / W_2	3.43 / 4.01	2.06 / 2.12	2.58 / 2.50	1.72 / 1.66
WD%	-15.59	-2.87	3.15	3.55
p value for WD% ≠ 0	<.001	n.s.	n.s.	n.s.(.055)

Note: n.s., not significant ($p > .05$); p between .05 and .10 is given in parentheses.

Table 7. Mean weightings, V_B / V_S and, W_1 / W_2 along with their weightings differential percentage, $WD\%$, as determined by fitting the Ratcliff diffusion model to the alternate responses and RTs obtained in Experiments 5 and 6.

	Drift rate μ	Drift rate μ
	Experiment 5: TOE	Experiment 6: SOE
V_B / V_S	.231 / .769	.206 / .794
W_1 / W_2	.445 / .485	.472 / .457
$WD\%$	-8.60	3.23
p value for $WD\% \neq 0$	.017	.015

Appendix

According to the notion of *adaptive perception* (Hellström, 1986, 1989) it is assumed

that the main aim of the perceptual apparatus is to report task relevant information-carrying

stimulus changes in time and/or space. So in comparing two stimuli the objective is not to

compute Bayesian estimates of the two stimulus magnitudes and compare them, as assumed

by, for instance, Ashourian and Loewenstein (2011), but instead to maximize sensitivity to

information-carrying changes in their relative magnitude. This is supposed to be achieved by

maximizing the signal to noise ratio (S/N) of a change in the perceived difference d_{12}, that is,

$\Delta d_{12} / \sigma_{d12}$, caused by a change in either or both of the stimuli (e.g., from trial to trial). This

leads to using the generalized difference, as defined in Equation 3, rather than the simple

difference between sensory magnitudes, $\psi_1 - \psi_2$, to represent the perceived difference d_{12}

between the paired stimuli. On the basis of work conducted by Hellström (1986, 1989), Δd_{12}

is here taken as the expected absolute amount of change in the subjective difference that

results from adding the quantities δ to ψ_1 and $-\delta$ to ψ_2. On these grounds, the following

quantity is assumed to be maximized:

$$S/N = \Delta d_{12} / \sigma_{d12} = \frac{u\, \delta\, (\partial d_{12} / \partial \psi_1) + (1-u)\, \delta\, (\partial d_{12} / \partial \psi_2)}{\sigma_{d12}} \qquad \text{A1}$$

In this case, S/N is the signal to noise ratio of d_{12} in terms of $\Delta d_{12} / \sigma_{d12}$, where Δd_{12} is

a nonrandom change or signal magnitude of d_{12} and σ_d its standard deviation (Gulliksen,

1958, termed σ_d the 'comparatal dispersion'), δ and $-\delta$ are the changes in ψ_1 and ψ_2, $\partial d_{12} /$

$\partial \psi_1$ and $\partial d_{12} / \partial \psi_2$ the partial derivatives of d_{12} in respect to ψ_1 and ψ_2, and u a coefficient

(between 0 and 1) of relative exogenously driven attention to the stimuli.

As detailed by Hellström (1989) simplifying by setting u to its neutral value of 0.5 and

δ to 2, and using Equation 3, yields

$$S/N = (s_1 + s_2) / (\sigma_{d12}), \qquad \text{A2}$$

and then expressing σ_{d12} in terms of Equation 3 permits calculation of optimal values of s_1 and s_2 that maximize S / N, and which can be compared as a percentage gain of S / N that results when $s_1 = s_2 = 1$; termed, gain percentage ($G\%$).

Calculation of the optimal weights s_1, s_2 and associated signal to noise ratios.
To find the optimal weights s_1, s_2, as referred to in Equation 3, the following equations can be formulated by letting

$$\sigma(\psi_1 - \psi_{r1}) = A; \ \sigma(\psi_2 - \psi_{r2}) = K{\cdot}A; \ \sigma(\psi_{r1} - \psi_{r2}) = C{\cdot}A; \ \rho(\psi_1 - \psi_{r1}, \psi_2 - \psi_{r2}) = R_1;$$

$$\rho(\psi_1 - \psi_{r1}, \psi_{r2} - \psi_{r2}) = R_2; \ \rho(\psi_2 - \psi_{r2}, \psi_{r1} - \psi_{r2}) = R_3;$$

and simplifying by introducing $y = (S / N) / A$, which yields

$$y = (s_1 + s_2) / [(s_1{}^2 + K^2 s_2{}^2 - 2KR_1 s_1 s_2 + C^2 + 2C(R_2 s_1 - KR_3 s_2)]^{\frac{1}{2}}. \qquad \text{A3}$$

Setting the first partial derivatives of y, with respect to s_1 and s_2, equal to zero, and then setting the resultant of the resulting equations, with respect to s_2 and s_1, respectively, equal to zero, yields two third-degree equations in s_1 and s_2, respectively. Each of these equations has one real and two complex roots, whereby the real roots yield a maximum of y and therefore also of S / N. Using S_1 and S_2 to denote the optimal values of s_2 and s_1,

$$S_1 = C(-KR_3{}^2 + K + R_1 - R_2 R_3)/(KR_1 R_3 - KR_2 - R_1 R_2 + R_3)$$

$$S_2 = C(KR_1 - KR_2 R_3 - R_2{}^2 + 1)/[K(KR_1 R_3 - KR_2 - R_1 R_2 + R_3)]$$

S/N for optimal s values (S / N_{opt})

$$= (S_1 + S_2) / [S_1{}^2 + K^2{*} S_2{}^2 - 2KR_1 S_1{}^2 + C^2 + 2C(R_2 S_1 - KR_3 S_2) (1 + V3^2 - 2R_5 V_3)]^{1/2}.$$

Efficiency $= (100{*}S/N_{\text{opt}}) / (S/N_{\text{comp}})$,

where S / N_{comp} is obtained by setting $s_1 = s_2 = 1$ in Equation A3.

The terms of this analysis can be translated by letting

$R_4 = \rho(\psi_1, \psi_2); \ R_5 = \rho(\psi_1, \psi_{r1}); \ R_6 = \rho(\psi_2, \psi_{r2});$

$R_7 = \rho(\psi_1, \psi_{r2}); \ R_8 = \rho(\psi_2, \psi_{r1}); \ R_9 = \rho(\psi_{r1}, \psi_{r2});$

$V_1 = \sigma(\psi_1); \quad V_2 = \sigma(\psi_2); \quad V_3 = \sigma(\psi_{r1}); \quad V_4 = \sigma(\psi_{r2}).$

Setting V_1 to 1 and using this unit for expressing V_2, V_3, and V_4,

$A = (1 + V_3{}^2 - 2R_5V_3)^{\frac{1}{2}};$

$K = (V_2{}^2 + V_4{}^2 - 2R_6V_2V_4)^{\frac{1}{2}} / A;$

$C = (V_3{}^2 + V_4{}^2 - 2R_9V_3V_4)^{\frac{1}{2}} / A;$

$R_1 = (R_4 V_2 - R_7V_4 - R_8V_2V_3 + R_9V_3V_4) / (KA^2);$

$R_2 = (R_5V_3 - R_7V_4 - V_3{}^2 + R_9V_3V_4) / (CA^2);$

$R_3 = (R_8V_2V_3 - R_6V_2V_4 - R_9V_3V_4 + V_4{}^2) / (CKA^2).$

The possible values of R_1, R_2, and R_3 are constrained by the inequality $0 < R_1{}^2 + R_2{}^2 + R_3{}^2 - 2R_1R_2R_3 < 1$, and the corresponding relation must hold for (R_4, R_5, R_8), (R_4, R_6, R_7), (R_5, R_7, R_9), and (R_6, R_8, R_9).

    MATLAB / SPSS code used to test the discrimination optimization model described in this Appendix may be obtained by contacting either Geoffrey R. Patching, E-mail: Geoffrey.Patching@psychology.lu.se, or Åke Hellström, E-mail: hellst@psychology.su.se.