

# Beware how you compare: comparison direction dictates stimulus-valence-modulated presentation-order effects in preference judgment

Mats P. Englund · Åke Hellström

Published online: 17 April 2013  
© Psychonomic Society, Inc. 2013

**Abstract** Englund and Hellström (Journal of Behavioral Decision Making 25:82–94, 2012a) found a tendency to prefer the left (first-read) of two attractive alternatives but the right (second-read) of two unattractive alternatives—a valence-dependent word-order effect (WOE). They used stimulus pairs spaced horizontally, and preference was indicated by choosing one of several written statements (e.g., “apple I like more than pear”). The results were interpreted as being due to stimulus position, with the magnitude of the left stimulus having a greater impact on the comparison outcome than the magnitude of the right. Here we investigated the effects of the positioning of the stimuli versus the semantics of the response alternatives (i.e., comparison direction) on the relative impacts of the stimuli. Participants rated preferences for stimuli spaced horizontally with the response alternatives not dictating a comparison direction (Exp. 1), and stimuli spaced vertically using Englund and Hellström’s (Journal of Behavioral Decision Making 25:82–94, 2012a) response alternatives, which dictate a comparison direction semantically (Exp. 2). The results showed that the valence-dependent WOE found by Englund and Hellström (Journal of Behavioral Decision Making 25:82–94, 2012a) was not due to the horizontal stimulus positioning per se, but to the induced comparison direction, with the effect probably being mediated by attention directed at the subject of the comparison. We concluded that a set comparison direction is required for the valence-dependent WOE to appear, and that using Hellström’s sensation-weighting model to determine stimulus weights is a way to verify the comparison direction.

**Keywords** Preference judgment · Stimulus valence · Feature matching · Comparison direction · Sensation weighting · Word-order effect

In the decision-making literature, numerous reports have demonstrated that the presentation order of the choice options affects the outcome systematically (e.g., Bruine de Bruin & Keren, 2003; Houston & Sherman, 1995; Wänke, Schwartz, & Noelle-Neumann, 1995). For example, presentation order has been shown to influence the results in contexts as various as opinion polls (Wänke et al., 1995), preference for political candidates (Houston & Roskos-Ewoldsen, 1998), preference for job options (Slaughter & Highhouse, 2003), evaluation of food products (Dean, 1980), and professional judges’ ratings in figure-skating contests (Bruine de Bruin, 2005, 2006). Analogous effects have been found by psychophysicists for one and a half centuries (see, e.g., Guilford, 1954, and Hellström, 1985, for reviews), and the psychophysical tradition to use many stimulus pairs of varying magnitudes has made it possible to discover that the size and direction of presentation-order effects vary with, among several factors, the stimuli’s magnitude level. For example, Hellström (2003) found that, in length comparisons of lines of relatively long durations, subjects overestimated the left out of two short lines but the right out of two long lines. Similarly, in comparisons of musical excerpts, Koh (1967) found a tendency to prefer the first of two unpleasant excerpts but the second of two pleasant excerpts; the size of this effect varied linearly with the rated pleasantness of the stimulus pairs. Analogous results were found by Englund and Hellström (2012b) for pairs of successive jingles as well as for color patterns. These are but a few examples of what Fechner (1860) called *space-order error* (SOE) and *time-order error* (TOE). As is hinted at by the two terms, a SOE appears when stimuli are separated spatially, and a TOE when stimuli are separated temporally.

M. P. Englund  
cut-e Sweden, Stockholm, Sweden

Å. Hellström (✉)  
Department of Psychology, Stockholm University,  
106 91 Stockholm, Sweden  
e-mail: hellst@psychology.su.se

Fechner defined these effects as being positive whenever the left (for SOEs) or the first (for TOEs) stimulus is overestimated relative to the other, and negative for the opposite case.

Inspired by these psychophysical results, Englund and Hellström (2012a) investigated preference choices between everyday objects and phenomena denoted by printed labels (e.g., *apple–pear*, *headache–stomach ache*) and found analogous magnitude-level dependent order effects. Specifically, the stimuli in each pair were presented at the left and right margins, respectively, of a page, and preference statements were printed in-between them (e.g., “I like [*much more, more, somewhat more, somewhat less, less, much less*] than . . .”). Preference was indicated for each stimulus pair by marking the statement that agreed most with the participant’s opinion. Participants also rated their overall opinions of the stimuli—that is, the stimuli’s valence levels. The results showed a tendency to prefer the left/first-read of two attractive stimuli (e.g., *apple–pear*) and the right/last-read of two unattractive stimuli (e.g., *headache–stomach ache*); the size of the effect varied linearly with the rated valence level of the stimuli. Englund and Hellström (2012a) dubbed this effect the *word-order effect* (WOE); in analogy to the SOE and TOE, a positive WOE indicates a tendency to prefer the left/first-read stimulus. One possible explanation for this valence-level-dependent WOE is that it is due to the left–right positioning of the words—that is, that it simply amounts to a SOE. Another possibility is that the effect is due to the semantics of the preference statements dictating a comparison direction (Tversky, 1977; Wänke, 1996). Obviously, pitting these two hypotheses against each other is of great theoretical, but also of practical, importance (e.g., in domains of product marketing and evaluation). Therefore, in the present article we present two experiments that were designed to test the two alternative explanations for the valence-level-dependent WOE reported by Englund and Hellström (2012a). First, however, we present the framework for description of the phenomenon and then the two alternative explanations.

### The sensation-weighting model

Englund and Hellström (2012a) found that the valence-level-dependent WOE that they observed could be described well by the *sensation-weighting* (SW) model, which has proved a powerful tool for describing the effects of temporal and spatial presentation order (i.e., TOEs and SOEs) in comparisons of psychophysical stimuli (e.g., Hellström, 1979, 1985, 2003; Patching, Englund, & Hellström, 2012), and also of aesthetic stimuli (Englund & Hellström, 2012b). This model rests, essentially, on the simple notion, common across many disciplines, that the impacts of two independent variables  $X_1$  and  $X_2$  (here, the valences of the compared stimuli) on a dependent

variable  $Y$  (here, the preference judgment) can often be efficiently analyzed by calculating the weights in a linear regression of  $Y$  on  $X_1$  and  $X_2$ . According to the SW model, the subjective difference,  $d_{12}$ , between two compared stimuli can be described as the difference between two weighted subjective magnitudes (the subscripts 1 and 2 denote the left and right [first and second] stimuli, respectively):

$$d_{12} = k\{[s_1\psi_1 + (1 - s_1)\psi_{\text{ref}}] - [s_2\psi_2 + (1 - s_2)\psi_{\text{ref}}]\} + b, \quad (1)$$

where  $k$  is a scale constant,  $s_1$  and  $s_2$  are weighting coefficients,  $b$  is a constant to account for effects not attributable to the weighting process (e.g., a judgment bias),  $\psi_1$  and  $\psi_2$  are the subjective stimulus magnitudes (here, valences), and  $\psi_{\text{ref}}$  is the magnitude corresponding to the current *reference level* (ReL). The ReL is conceived similarly to the *adaptation level* (Helson, 1964), which results from the pooling of focal, background, and residual stimulation, creating a subjectively neutral point—that is, an internal standard. According to the SW model, the size and sign of the order effect equals  $d_{12}$  when  $\psi_1 = \psi_2 = \psi$ , and with the often reasonable assumption that  $b = 0$ , Eq. 1 reduces to

$$d_{12} = k(s_1 - s_2)(\psi - \psi_{\text{ref}}). \quad (2)$$

This simplification is particularly useful in studying effects on preferences, using pairs of stimuli that are rather close in their positions on the valence continuum. According to Eq. 2, with the weight relation  $s_1 > s_2$ , the left (first) stimulus is overestimated when comparing two stimuli above the ReL, and the right (last) is overestimated when comparing two stimuli below the ReL. This weight relation could then account for the valence-level-dependent order effects obtained by Englund and Hellström (2012a). That is, the left stimulus had a greater impact than did the right on the outcome of the preference comparison, which led to the tendency to prefer the leftmost of two attractive stimuli and the rightmost of two unattractive stimuli.

### The space-order hypothesis

As we noted above, the first possible explanation for the valence-dependent WOE, described by the weight relation  $s_1 > s_2$ , is that this order effect is due to the horizontal spacing of the words—that is, to a SOE. This interpretation is in line with previous psychophysical results as well as with the right-hemisphere lateralization hypothesis; a higher weight for the left than for the right stimulus of two horizontally spaced stimuli has been found previously for comparisons of line length (Hellström, 2003; Masin & Agostini, 1991), and more efficient processing by the right hemisphere (receiving stimulation from the left visual field) than by the left hemisphere has been found for emotional stimuli (e.g., Borod et al., 1998; Lang, Bradley, & Cuthbert, 1990). A modified version of this hypothesis would point to the possible importance of the

successiveness of stimulus presentation that arises from the habitual reading order (for Swedish participants, from left to right), which would make the WOE akin to the TOE. (It should be noted that self-administered presentation of printed stimuli cannot be characterized as strictly simultaneous or successive.)

### The comparison direction hypothesis

The second explanation—that the valence-dependent WOE, described by the weight relation  $s_1 > s_2$ , is due to the comparison direction—comes from cognitive psychology. In the cognitive research on presentation-order effects in preference comparisons, the most prominent paradigm is that of *feature matching* (e.g., Houston & Sherman, 1995; Houston, Sherman, & Baker, 1989, 1991; Wänke et al., 1995), which is based on Tversky's (1977) *contrast model* of similarity comparison. The basic assumption of feature matching is that each stimulus can be represented by a set of features, and that the comparison is directed so that one stimulus, the *subject*, is compared to another stimulus, the *referent*. Furthermore, Houston and Sherman (1995) argued that, for judgments of preference, the shared features of the subject and referent are cancelled out because they do not carry any distinguishing information. Then, the unique features of the subject work as a checklist, and these features will be looked for actively among the features of the referent. Given equally many unique features of both stimuli, this checklist procedure will induce the perception that the subject possesses a larger number of the unique features. Therefore, the subject will be chosen whenever the unique features of the two stimuli are attractive, whereas the referent will be chosen whenever the unique features are unattractive. Consequently, one critical assumption behind the accuracy of feature-matching models based on Tversky's (1977) ideas is that the direction of the order effect depends on the comparison direction—that is, on which stimulus is the subject in the comparison. In line with this interpretation of the effect, Englund and Hellström (2012a) suggested that the higher weight for the left stimulus ( $s_1 > s_2$ ) in their results was due to the semantics of the preference statements pointing out the left stimulus as the subject, to be compared to the referent (the right stimulus). Such an effect could be mediated by more attention being allotted to the subject than to the referent (cf. Hellström, 1985; Tversky, 1977).

### The present study

To the best of our knowledge, in only one study (Wänke, 1996) have WOEs been studied by pitting a word-order hypothesis against a comparison direction hypothesis. Wänke investigated the matter explicitly by using a factorial design with comparison requests with different semantic structures and different stimulus presentation orders. She concluded from her results

that semantic determination of the comparison direction affects the responses, but that mere word order does not. A few important points need to be made regarding Wänke's study, however. First, participants were asked the questions by interviewers, thus making the stimuli separated temporally rather than spatially. Thus, her results did not address the present issue of whether the valence-dependent WOE is analogous to a SOE—that is, whether it is induced by the spatial order of the stimuli. Second, the possibility exists that the different wordings used to vary the comparison direction “may have changed the comprehensibility of the sentence” (p. 404). Third, out of the three choices that Wänke's participants made, only one was of a preferential nature. Consequently, before generalizing the interpretation of semantic dictation of the comparison direction to preference choices—and, in particular, to the valence-level-dependent WOE—it is of utmost interest to investigate this hypothesis further. Therefore, in the present article, we present two experiments that were designed to investigate, using an alternative method to Wänke's, whether the valence-level-dependent WOE is due to the horizontal stimulus positioning (i.e., the space-order hypothesis) or to the semantics dictating a comparison direction by pointing out one stimulus as the subject of the comparison (i.e., the comparison direction hypothesis).

### General method

The participants, who were different for each experiment, were Stockholm University psychology students who participated as volunteers or to fulfill a partial course requirement. In each experiment, the participants received a booklet consisting of three sections—a preference judgment task, a filler task (two personality tests not related to the experimental hypotheses: the Barratt Impulsiveness Scale of Patton, Stanford, & Barratt, 1995, and the Wender Utah Rating Scale of Ward, Wender, & Reimherr, 1993), and a valence-rating task. In the preference task, the participants were to make comparisons of stimulus pairs. The stimuli (e.g., *apple–pear*, *headache–stomach ache*) were spaced either horizontally (Exp. 1), by being printed at the respective margins, or vertically (Exp. 2), by being printed centrally on the page. In both experiments, the response alternatives were printed centrally on separate lines in-between the stimuli of each pair. The order of the response alternatives was counterbalanced, and the order of the stimulus pairs was randomized for each participant.

In the stimulus-valence rating task of Experiments 1 and 2, participants were to rate their general opinion on each stimulus separately. In Experiment 1, the stimulus was printed in the left margin, and there were seven statements (from uppermost to lowermost: “... I generally [*like greatly*, *like*, *like somewhat*, *neither like nor dislike*, *dislike somewhat*, *dislike*, *dislike greatly*] \_\_\_\_\_”) written on separate

lines, with the stimulus on a level with the neutral alternative. Participants indicated their general opinions of the stimuli by marking with an “X” the dashed line to the right of the valence statement that most accurately represented their opinions. In Experiment 2, the response alternatives were the same as for Experiment 1, and in the same vertical order, but the stimulus was printed centrally above the response alternatives instead of in the left margin. The presentation order of the stimuli was randomized for each participant with the restriction that two stimuli belonging to the same pair occurred at least two full pages from one another.

## Experiment 1

In Experiment 1, the aim was to test whether the valence-level-dependent WOE (Englund & Hellström, 2012a) is due to the left–right spatial positioning of the stimuli (i.e., to a SOE). This was tested by attempting to eliminate the semantically dictated comparison direction but still using a left–right stimulus positioning. If the valence-dependent WOE is due to the stimulus positioning (in accordance with the space-order hypothesis), it should occur even under these conditions, but if it is due to a comparison direction that is dictated semantically (in accordance with the comparison direction hypothesis), it should disappear.

## Method

**Participants** A group of 168 participants took part (48 men and 119 women, plus one participant who did not state gender and age), ranging in age from 19 to 54 years ( $M_{\text{age}} = 26.3$ ). (Two participants, uncounted here, failed to finish the booklets, and their data were discarded from all calculations.) The participants were assigned randomly to one of two groups receiving opposite orders of the preference response alternatives.

**Stimuli and procedure** In the preference comparison task, participants were to make paired comparisons of 24 stimulus pairs (see Table 1); three stimulus pairs were printed on each page. The stimuli in each pair were spaced horizontally by being printed at the opposite margins of the page. The six response alternatives were printed on separate lines in-between the stimuli and consisted of a short preference statement, together with an open arrow pointing at the left or the right stimulus (see Fig. 1). Participants were instructed that the arrow pointed to the stimulus that was preferred according to the statement. The arrow designations were counterbalanced between participants; for half of the participants, the three topmost alternatives’ arrows

pointed at the left stimulus and the three bottommost pointed at the right stimulus, and for the other half of the participants, the arrow allocations were the reversed. The participants were to indicate their preferences by marking the response alternative that agreed most with their own opinions.

The within-pair presentation order was randomized so that, for each participant, a random set of 12 pairs were presented in the within-pair order specified in Table 1, and the remaining 12 pairs were presented in the opposite within-pair order. Thus, each participant received a unique combination of which stimulus pairs were presented in the within-pair orders  $A_i B_i$  and  $B_i A_i$ . The order of the stimulus pairs was randomized for each participant.

## Results and discussion

Preference ratings were scaled from 2.5 (*maximum preference for the left stimulus*) to  $-2.5$ , in steps of 1. The valence ratings were scaled from 3 (*highest positive valence*) to  $-3$ , also in steps of 1. The means and standard deviations of the preference and valence ratings for each word order (AB or BA) are displayed in Table 1.

The relative effect of the presentation order on the preference judgments, the WOE, for a stimulus pair  $\{A, B\}$  was defined theory-independently for this design by Englund and Hellström (2012a) as one-half of the difference between the mean preference judgments in the two stimulus-presentation orders AB and BA. The WOE was defined as being positive (in accordance with the SOE definition of Fechner, 1860) whenever the left tended to be preferred over the right stimulus, and as being negative when the right stimulus tended to be preferred. Englund and Hellström (2012a) also showed that the WOE can be expressed in terms of the SW model (Hellström, 1979, 1985, 2003); a first step would be to rewrite Eq. 1 to yield the simpler formulation (with the subscripts L and R denoting the *left* and *right* stimulus, respectively)

$$d_{LR} = W_L \cdot \psi_L - W_R \cdot \psi_R + C, \quad (3)$$

where  $W_L = k \cdot s_L$ ,  $W_R = k \cdot s_R$ , and  $C = (W_R - W_L)\psi_{\text{ref}} + b$ . Then, using Eq. 3, the preference  $d_{LR}$  in the pair  $\{A, B\}$  can be expressed for the two possible presentation orders, respectively, as

$$d_{AB} = W_L \cdot \psi_A - W_R \cdot \psi_B + C \quad (4a)$$

and

$$d_{BA} = W_L \cdot \psi_B - W_R \cdot \psi_A + C, \quad (4b)$$

where the subscripts AB and BA denote the presentation orders. Finally, using Eqs. 4a and 4b, the WOE for the pair  $\{A, B\}$  can be expressed as the predicted  $d_{LR}$

**Table 1** Experiment 1: Means (and standard deviations) of preference and valence ratings for horizontally spaced stimulus pairs with unique within-pair stimulus randomization (in translation from the original Swedish, where each stimulus was denoted by a single noun)

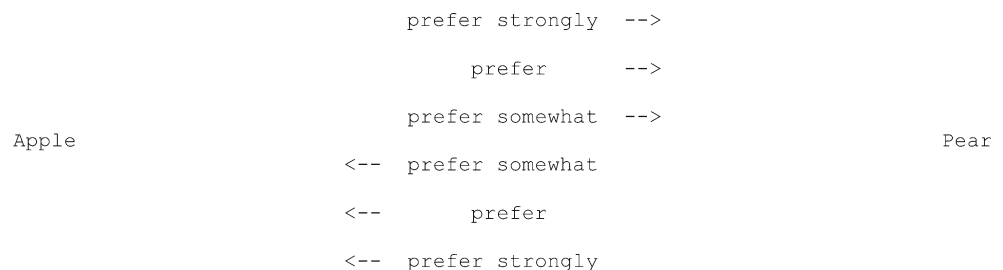
Stimulus Pair $i$ ( $A_i$ – $B_i$ )	Relative Preference ( $A_i$ – $B_i$ ) <sup>a</sup>		Mean Valence of Stimuli in Pair ( $A_i$ , $B_i$ ) <sup>b</sup>
	Order AB $M$ ( $SD$ )	Order BA $M$ ( $SD$ )	
1. Aftonbladet–Expressen	0.73 (1.41)	0.69 (1.11)	–0.31 (1.36)
2. Apple–Pear	0.64 (1.30)	0.54 (1.36)	1.61 (0.81)
3. Bus–Subway	0.28 (1.52)	0.19 (1.59)	0.83 (0.96)
4. Card games–Board games	0.17 (1.24)	0.34 (1.24)	1.27 (1.10)
5. Cinema–Concert	–0.27 (1.45)	–0.15 (1.74)	2.07 (0.64)
6. Circles–Triangles	0.94 (1.16)	0.49 (1.08)	0.62 (0.70)
7. Cough–Runny nose	–0.49 (1.06)	–0.32 (1.10)	–2.06 (0.67)
8. Deciduous forest–Coniferous forest	0.94 (1.17)	0.89 (1.10)	1.65 (0.87)
9. Dogs–Cats	0.40 (1.81)	0.37 (1.68)	1.36 (1.25)
10. Fall–Winter	0.52 (1.57)	0.51 (1.52)	0.85 (1.26)
11. Fever–Vomits	1.43 (1.07)	1.34 (1.12)	–2.41 (0.59)
12. Filth brown–Dust gray	–0.73 (1.02)	–0.54 (1.03)	–0.49 (1.06)
13. Finland–Norway	–0.55 (1.53)	–0.85 (1.22)	0.91 (0.81)
14. Forest green–Sea blue	–0.87 (1.49)	–1.08 (1.38)	1.94 (0.72)
15. Headache–Stomach ache	0.38 (1.02)	0.22 (1.14)	–2.46 (0.59)
16. High jump–Long jump	0.13 (1.33)	–0.01 (1.16)	0.25 (1.04)
17. Ink pens–Lead pencils	0.62 (1.32)	0.55 (1.41)	1.26 (0.83)
18. Pasta–Rice	0.95 (1.10)	0.83 (1.41)	1.86 (0.73)
19. Piano–Guitar	0.26 (1.49)	0.15 (1.44)	2.03 (0.80)
20. Plastic bag–Paper carrier	–0.23 (1.53)	0.09 (1.35)	0.67 (0.75)
21. Toothache–Nausea	0.04 (1.23)	0.12 (1.04)	–2.53 (0.61)
22. Traffic noise–Radio static	0.60 (1.01)	0.39 (1.22)	–1.69 (0.86)
23. TV news–Newspaper news	–0.16 (1.35)	–0.18 (1.47)	1.33 (0.90)
24. Wasps–Mosquitoes	0.12 (1.42)	–0.13 (1.35)	–1.99 (0.80)
$M$ ( $SD$ ) of means	0.24 (0.58)	0.19 (0.55)	0.27 (1.60)

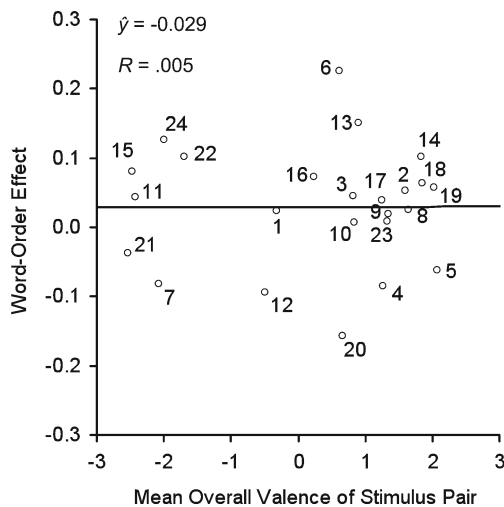
Positive preference values denote preference for A over B. Stimulus Pair 1 denotes two Swedish evening newspapers. <sup>a</sup> Means and standard deviations were calculated for each stimulus pair over those participants who received the specific order (roughly half of the participants for each stimulus pair). <sup>b</sup> Calculated over all participants ( $n = 168$ )

in a hypothetical pair where each stimulus has the valence  $(\psi_A + \psi_B)/2$ :

$$\begin{aligned}
 d_{LR} = \text{WOE} &= \left\{ \frac{d_{AB} - (-d_{BA})}{2} \right\} = \left\{ \frac{d_{AB} + d_{BA}}{2} \right\} \\
 &= \left\{ \frac{(W_L - W_R)(\psi_A + \psi_B) + 2U}{2} \right\} \\
 &= (W_L - W_R) \left[ \left\{ \frac{\psi_A + \psi_B}{2} \right\} - \psi_{\text{ref}} \right] + b.
 \end{aligned}
 \tag{5}$$

**Fig. 1** Experiment 1: Layout of response sheet for horizontal stimulus positioning





**Fig. 2** Experiment 1: Horizontal stimulus positioning. The word-order effect (WOE) is plotted against the mean valence level of each stimulus pair (labels are from Table 1). The fitted regression line is also displayed. A positive WOE means a tendency to prefer the left over the right stimulus

based on the space-order hypothesis, according to which the spatial separation of the stimuli should yield a valence-level-dependent WOE—is falsified by the results in Fig. 2. According to this prediction and Eq. 5, a linear relationship should hold between WOE and valence level. However, no such relationship was found. As can be seen in Fig. 2 (and Table 1), the WOE values for some stimulus pairs are different from zero, but these effects seem to portray chance rather than systematic effects. Indeed, we found no significant linear ( $p = .986$ ) or quadratic ( $R^2 < .001$ ,  $p = .9998$ ) relationship between WOE and valence level. These results are in opposition to the predictions of the space-order hypothesis, but compatible with the comparison direction hypothesis.

For each participant, Eq. 3 was fitted by linear regression of  $d_{LR}$  on  $\psi_L$  and  $\psi_R$ , yielding  $W_L$ ,  $W_R$ , and  $C$  (due to the minus sign in the equation, the sign of  $W_R$  was changed). The mean multiple  $R$  across participants was .68 ( $SD = .13$ , range .05–.91). The mean  $W_L$  was 0.76 ( $SD = 0.25$ , range 0.01–1.42), and the mean  $W_R$  was 0.76 ( $SD = 0.24$ , range 0.03–1.48). The mean  $C$  was 0.019 ( $SD = 0.24$ , range  $-0.68$  to 0.69). The mean of  $W_L - W_R$  was  $-0.005$  ( $SD = 0.13$ ,  $SEM = 0.010$ , range  $-0.52$  to 0.37).

The individual  $W$  values were submitted to a repeated measures analysis of variance (ANOVA) with Stimulus Position as a within-participants factor (left vs. right) and Gender and Scale Order (uppermost arrow pointing left vs. right) as between-participants factors. No main or interaction effects of gender approached significance, so the Gender factor was dropped from further analyses. No main effect of stimulus position was apparent,  $F(1, 166) = 0.23$ ,  $p = .629$ ,  $\eta_p^2 = .001$ ; that is, the left stimulus had the same impact (mean

$W_L = 0.756$ ) on the comparison as did the right stimulus (mean  $W_R = 0.761$ ). We also found no significant overall effect of scale order,  $F(1, 166) = 0.74$ ,  $p = .390$ ,  $\eta_p^2 = .004$ , and stimulus position did not interact significantly with this variable,  $F(1, 166) = 0.00$ ,  $p = .963$ ,  $\eta_p^2 = .000$ . Thus, the lack of a systematic effect of presentation order was mirrored by the absence of systematic differential weighting of the compared stimuli. Next, the individual values of  $C$  (i.e., the predicted WOE for a stimulus pair with valences of zero) were submitted to a two-way ANOVA with Gender and Scale Order as between-participants factors. The main effect of gender and its interaction with scale order did not approach significance. Therefore, the Gender factor was dropped. The effect of scale order did reach significance,  $F(1, 166) = 4.57$ ,  $p = .034$ ,  $\eta_p^2 = .027$ : With the uppermost arrow pointing to the left and the right, respectively, the mean  $C$ s were 0.060 ( $SD = 0.22$ ) and  $-0.017$  ( $SD = 0.25$ ).

It may be noted that in the study by Phaf and Rotteveel (2009), arrows to the right induced more positive affect than did arrows to the left. However, such a tendency, if present, should not have affected the weighting here. The arrows were open, so any association of right-directed arrows with rewarding “Play” buttons, suggested by Phaf and Rotteveel, should be ruled out. If the arrows should still have evoked a tendency to prefer the right alternative over the left, this tendency should have been equally present for all pairs, and thus cancelled out in the data by the thoroughly randomized design. The same applies to a possible tendency to prefer the uppermost statements, which might be predicted by Meier and Robinson’s (2004) finding of a tendency to associate upper and lower positions with positive and negative words, respectively.

## Experiment 2

The results from Experiment 1, where a left–right presentation order was used, weakened seriously the hypothesis that the valence-dependent WOE found by Englund and Hellström (2012a) was simply a SOE with a higher weight for the left stimulus (Hellström, 2003). Instead, the results are compatible with the suggestion made by Englund and Hellström that valence-dependent WOEs occur because the comparison direction is dictated semantically by the preference statements (Wänke, 1996). In Experiment 2, the plausibility of the comparison direction hypothesis was investigated further by using vertical stimulus positioning and the same response alternatives that Englund and Hellström (2012a) had used; if the linear valence dependence of the WOE is due to the semantic dictation of the comparison direction, it should reappear. It was hypothesized that the upper stimulus (the subject) would be given more attention, and therefore receive greater weight, than the nether stimulus (the referent), and that this would lead to

a linearly valence-dependent WOE with a tendency to prefer the upper of two attractive stimuli, or the nether of two unattractive stimuli.

Method

*Participants* A total of 174 participants took part, 41 men and 131 women (plus two participants who did not state gender, one of whom also did not state age), ranging in age from 20 to 53 years ( $M_{age} = 27.2$ ). The participants were assigned randomly to one of two groups (with opposite orders of the response alternatives; see below).

*Stimuli and procedure* The procedure and the stimuli were the same as in Experiment 1, but in the preference comparison task, the stimuli of each pair were printed centrally and spaced vertically, with the response alternatives (those from Englund and Hellström’s, 2012a, Exp. 2—i.e., “. . . I like [*much more, more, somewhat more, somewhat less, less, much less*] than . . .”) printed on separate lines in-between them (see Fig. 3). A check-box was placed just to the left of each response alternative; for each stimulus pair, participants were to respond by choosing the statement that most accurately represented their opinion.

Results and discussion

The scaling of the preference and valence ratings (for the means and standard deviations, see Table 2), as well as the calculation of the WOE values for the respective stimulus pairs, were performed using the methods described for

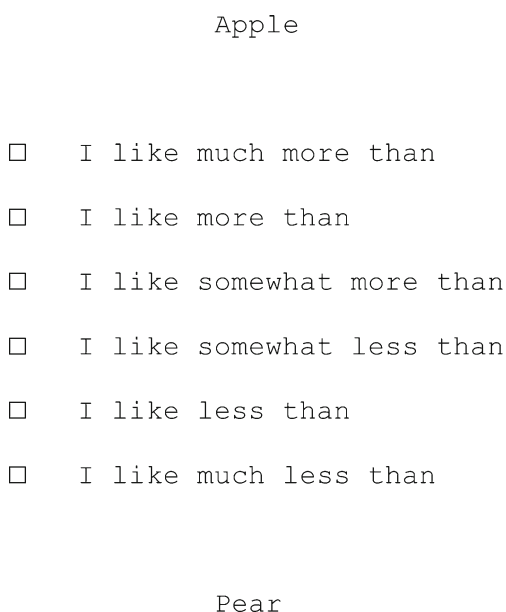


Fig. 3 Experiment 2: Layout of response sheet for vertical stimulus positioning with a dictated comparison direction

Experiment 1. The WOE values are displayed in Fig. 4, plotted against the mean overall valence for the respective pair, and a fitted regression line is also shown. As can be seen from Fig. 4, with the comparison direction dictated semantically, the expected linearly valence-dependent WOE reappeared. The slope of the regression of the WOE on mean valence for the stimulus pairs was statistically significant,  $t(22) = 4.31, p < .001$ , as was the intercept,  $t(22) = -2.38, p = .027$ . The significant regression slope indicates a valence-level-dependent WOE, and the significant intercept suggests a slight bias toward a higher preference for the nether stimulus. The SW model is well able to account for these results, which becomes clear by replacing the subscripts L and R in Eq. 5 with subscripts denoting the upper (U) and nether (N) stimulus, respectively, which yields.

$$WOE = (W_U - W_N) \left[ \left\{ \frac{(\psi_A + \psi_B)}{2} \right\} - \psi_{ref} \right] + b. \quad (6)$$

Thus, in SW terms, the valence-dependent WOE is described as a higher weight for the upper stimulus,  $W_U > W_N$ .

Analyses were also performed on the individual  $W$  and  $C$  values, which were estimated by linear regression as in Experiment 1. The mean multiple  $R$  across participants was .70 ( $SD = .12$ , range .23–.88). The mean  $W_L$  was 0.72 ( $SD = 0.27$ , range -0.91 to 1.50), and the mean  $W_R$  was 0.65 ( $SD = 0.28$ , range -0.88 to 1.38). The mean  $C$  was -0.19 ( $SD = 0.26$ , range -0.61 to 0.88).

The individual  $W$  values were submitted to a repeated measures ANOVA with Stimulus Position (upper vs. nether) as a within-participants variable and Gender and Scale Order (“. . . I like *much more* than . . .” being the uppermost vs. nethermost response alternative) as between-participants factors. No main or interaction effects of gender approached significance, so the Gender factor was dropped from further analyses. The analysis showed that the weight for the upper stimulus was significantly higher than that for the nether stimulus,  $M_{W_U - W_N} = 0.071$  ( $SD = 0.18, SEM = 0.013$ , range -0.44 to 0.84),  $F(1, 172) = 28.82, p < .001, \eta_p^2 = .144$ , indicating that the upper stimulus had a greater impact on the comparison than did the nether stimulus. The effect of stimulus position was nonsignificant,  $F(1, 172) = 0.001, p = .979$ . The interaction of stimulus position and scale order approached significance,  $F(1, 172) = 3.11, p = .080, \eta_p^2 = .018$ , with the mean weight differences ( $W_U - W_N$ ) being 0.048 and 0.094 for participants with *much more* printed uppermost and nethermost, respectively. However, simple analyses confirmed that the mean weight difference was significant for both scale orders,  $t(86) = 2.74, p = .007$ , and  $t(86) = 4.73, p < .001$ , respectively. Next, the individual values of  $C$  (the predicted WOE for a stimulus pair with valences of zero) were submitted to a two-way ANOVA with Gender and Scale Order as between-participant factors.

**Table 2** Experiment 2: Means (and standard deviations) of preference and valence ratings for vertically spaced stimulus pairs with unique within-pair stimulus randomization (in translation from the original Swedish, where each stimulus was denoted by a single noun)

Stimulus Pair $i$ ( $A_i$ – $B_i$ )	Relative Preference ( $A_i$ – $B_i$ ) <sup>a</sup>		Mean Valence of Stimuli in Pair ( $A_i$ , $B_i$ ) <sup>b</sup>
	Order AB $M$ ( $SD$ )	Order BA $M$ ( $SD$ )	
1. Aftonbladet–Expressen	0.72 (1.10)	0.59 (1.02)	–0.40 (1.45)
2. Apple–Pear	0.58 (1.30)	0.38 (1.23)	1.58 (0.90)
3. Bus–Subway	0.16 (1.69)	0.28 (1.46)	0.86 (0.93)
4. Card games–Board games	0.25 (1.36)	0.03 (1.33)	1.19 (1.06)
5. Cinema–Concert	0.08 (1.58)	–0.17 (1.22)	2.16 (0.64)
6. Circles–Triangles	0.82 (1.10)	0.88 (1.01)	0.68 (0.83)
7. Cough–Runny nose	–0.55 (1.36)	–0.31 (1.37)	–2.11 (0.74)
8. Deciduous forest–Coniferous forest	0.93 (1.22)	0.88 (1.40)	1.73 (0.92)
9. Dogs–Cats	0.30 (1.68)	0.45 (1.64)	1.51 (1.11)
10. Fall–Winter	0.65 (1.47)	0.78 (1.47)	0.98 (1.29)
11. Fever–Vomits	1.20 (1.43)	1.82 (1.26)	–2.41 (0.73)
12. Filth brown–Dust gray	–0.81 (1.23)	–0.92 (1.41)	–0.51 (1.11)
13. Finland–Norway	–0.92 (1.21)	–0.66 (1.25)	0.90 (0.94)
14. Forest green–Sea blue	–0.29 (1.53)	–0.72 (1.47)	1.82 (0.81)
15. Headache–Stomach ache	0.44 (1.28)	0.48 (1.32)	–2.44 (0.69)
16. High jump–Long jump	0.18 (1.38)	0.15 (1.37)	0.37 (1.09)
17. Ink pens–Lead pencils	0.79 (1.33)	0.65 (1.44)	1.37 (0.76)
18. Pasta–Rice	0.56 (1.38)	0.60 (1.26)	1.81 (0.81)
19. Piano–Guitar	0.10 (1.26)	–0.01 (1.44)	1.90 (0.91)
20. Plastic bag–Paper carrier	–0.04 (1.50)	0.30 (1.46)	0.73 (0.81)
21. Toothache–Nausea	–0.12 (1.61)	0.37 (1.45)	–2.64 (0.49)
22. Traffic noise–Radio static	0.38 (1.41)	0.65 (1.68)	–1.82 (0.86)
23. TV news–Newspaper news	–0.14 (1.40)	–0.27 (1.32)	1.42 (0.97)
24. Wasps–Mosquitoes	–0.27 (1.54)	0.47 (1.92)	–1.89 (0.84)
$M$ ( $SD$ ) of means	0.21 (0.54)	0.28 (0.60)	0.28 (1.62)

Positive preference values denote preference for A over B. Stimulus Pair 1 denotes two Swedish evening newspapers. <sup>a</sup> Means and standard deviations were calculated for each stimulus pair over those participants who received the specific order (roughly half of the participants for each stimulus pair). <sup>b</sup> Calculated over all participants ( $n = 174$ )

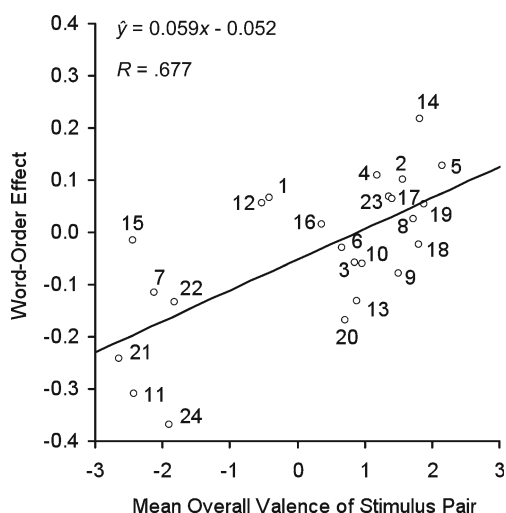
Here, the effect of gender reached significance,  $F(1, 168) = 4.21$ ,  $p = .042$ ,  $\eta_p^2 = .024$ . For the women,  $M_C$  was  $-0.043$  ( $SD = 0.26$ ), indicating a slight tendency to prefer the nethermost stimulus, and for the men,  $M_C = 0.060$  ( $SD = 0.27$ ). The effect of scale order was nonsignificant,  $F(1, 168) = 0.83$ ,  $p = .364$ ,  $\eta_p^2 = .005$ . The Gender  $\times$  Scale Order interaction was also nonsignificant,  $F(1, 168) = 1.24$ ,  $p = .267$ ,  $\eta_p^2 = .007$ .

The upper versus nether position per se might be suspected to have affected the results. As was noted earlier, Meier and Robinson (2004) found a tendency to associate an upper and a nether position with positive and negative words, respectively. Such an association could be expected to lead to a general tendency to prefer the upper stimulus, which is in opposition to the slight tendency that we obtained in the opposite direction. Furthermore, due to the randomized design, such a preference should have applied equally to all

stimulus pairs, and could hardly have changed the weighting. Catrambone, Beike, and Niedenthal (1996) found, consistent with Tversky's (1977) feature-matching theory, higher similarity judgments of pairs of countries (i.e., "How similar is A to B?") when B (the referent) was highly familiar and A (the subject) less familiar than were found in the reverse case. In contrast, no effect was apparent of the vertical position (upper, nether) of the familiar and less familiar country names on nondirectional similarity judgments (i.e., "How similar are A and B?"). By analogy, the higher weight for the upper stimulus in Experiment 2 is not likely to be due to the vertical placement of the stimuli to be compared.

The present results strengthen further the comparison direction hypothesis—that is, that the linear valence-level dependence of the WOE arises as a result of the systematic use of one stimulus, in this case the upper, as the subject in the comparison (e.g., Houston & Sherman, 1995; Wänke,





**Fig. 4** Experiment 2: Vertical stimulus positioning with a dictated comparison direction. The word-order effect (WOE) is plotted against the mean valence level of each stimulus pair (labels are from Table 2). The fitted regression line is also displayed. A positive WOE means a tendency to prefer the upper over the nether stimulus

1996). This would then increase the attention that this stimulus receives, and thereby its weight—that is, its impact on the comparison outcome. At the same time, these results further weaken the space-order hypothesis, the notion that Englund and Hellström’s (2012a) results—a valence-level-dependent WOE expressing itself as a higher weight for the left than for the right stimulus—were due to the space order per se—that is, that they constituted a SOE proper (Hellström, 2003; Masin & Agostini, 1991).

## General discussion

The two experiments presented here were designed to investigate further the potential impact of the semantics of the response alternatives on the valence-level-dependent WOE (Englund & Hellström, 2012a). Specifically, the purpose was to investigate the two rivaling hypotheses that the valence-level-dependent WOE was due to (a) the horizontal stimulus positioning (i.e., a classic SOE), or (b) the semantics of the response alternatives dictating a comparison direction by pointing out one stimulus, the first-read, as the subject to be compared to the referent, the last read. Taken together, the results of these two experiments show that the valence-level-dependent WOE (Englund & Hellström, 2012a) is not due to the spatial positioning of the stimuli—that is, to a SOE proper. Instead, the effect seems to result from the comparison being directed, with a subject being compared to a referent. In the present case, the appointing of the first-read stimulus as the subject is the result of the specific preference statements emphasizing the first stimulus as possessing the property of being liked more or less *than* the second stimulus (“A I like . . . than

B”). In such a directed stimulus comparison, the stimulus in focus, the subject, has a greater impact on the comparison (e.g., Houston & Sherman, 1995; Tversky, 1977), which may be mediated by a higher degree of attention and, in SW terms, is represented by a higher weight for the subject than for the referent. As is illustrated by Eq. 2, letting the subscripts 1 and 2 represent the *subject* and *referent*, respectively, the result is that the WOE varies in size and direction with the valence level ( $\psi$ ) of the compared stimuli relative to a reference level ( $\psi_{\text{ref}}$ ). Although the present results make it unlikely that a valence-dependent WOE would be evoked by lateral stimulus positioning alone, this factor might possibly modulate the effect of the subject–referent relationship of the alternatives. It should therefore be of interest to investigate such possible interactions.

As was discussed by Englund and Hellström (2012a), changes of the size and direction of the WOE with the stimulus valence cannot be explained by most common preference models, including expected-utility models (e.g., Bradley & Terry, 1952; Kahneman & Tversky, 1979; Luce, 1959; Marley, Flynn, & Louviere, 2008), because these models do not even address order effects. Even though some suggested model modifications are designed specifically to accommodate order effects (e.g., Beaver & Gokhale, 1975; Davidson & Beaver, 1977), these extended models fail to explain the present data. The reason for that failure is the attempt to account for order effects by merely adding a constant to the model. This kind of extension will only help to explain a positive *or* a negative order effect, but not both (see Englund & Hellström, 2012a, for a fuller discussion). Feature-matching models, though (e.g., Houston & Sherman, 1995; Tversky, 1977), are exceptions to this argument, as these models can be used to explain changes in the direction of order effects. However, as Englund and Hellström (2012a) noted, it is unclear whether and how feature-matching models can be used to account for the valence-level dependence of the WOE. Englund and Hellström (2012a) suggested one possibility, attaching valence values to the features, but argued that feature-matching models generally lead to predictions similar to those of the simpler SW model, which makes the latter the preferred model (see Englund & Hellström, 2012a, for a discussion on the limitations of the feature-matching paradigm).

The present results and those of previous research (Englund & Hellström, 2012a; Wänke, 1996) strongly suggest that comparison direction is a factor that is decisive for the direction of presentation-order effects; thus, it is imperative for the researcher to know in which direction the comparisons are made. In the present case, the direction was dictated explicitly by the preference statements, but it may not be quite as easy to predict the comparison direction when using other methods or instructions. This issue does not seem to have been investigated specifically, but researchers have made suggestions. For example, it has been suggested that the comparison direction can be determined by explicit instructions or questions with an

implied direction (e.g., Houston & Sherman, 1995; Houston et al., 1991; Wänke, 1996; Wänke et al., 1995), or by sequential stimulus presentation (e.g., Brunner & Wänke, 2006; Houston & Sherman, 1995; Houston, Sherrill-Mittleman, & Weeks, 2001; Kardes & Sanbonmatsu, 1993). However, the results in the literature are inconsistent (e.g., Agostinelli, Sherman, Fazio, & Hearst, 1986; Englund & Hellström, 2012a; Houston & Sherman, 1995; Houston et al., 1989; McGill, 1990; Wänke et al., 1995). More importantly, mere instruction or implication does not ensure knowing which stimulus has actually been used as the subject, and this uncertainty is a danger to the researcher; it is all too easy to interpret the results by assuming that a particular stimulus was indeed the subject, because the results support the model given this assumption—a circular argument. The present results, in concert with those of Englund and Hellström (2012a), suggest that estimating the weights ( $W$  values) is one way to determine post hoc which stimulus was used as the subject. This is also in line with Tversky's (1977) original formulation of his contrast model. Therefore, because the issue of determining which stimulus is used as the subject is crucial for the interpretation of results and for the systematic testing of models, the further investigation of the validity of the  $W$  values as reflecting the comparison direction is a natural quest for future research.

The present results suggest important implications for research practice. One commonly used method of handling potential order effects is the use of within-participants counterbalancing of the order of the choice options. This may be an effective method, but it is limited to contexts where it is reasonable. As was pointed out by Englund and Hellström (2012a), this is not the case when stimuli are remembered easily and participants have some motivation to appear consistent, and therefore tend to respond in accordance with previous responses. Decision models that do not account for presentation-order effects are inherently at fault with respect to decision-making reality, because these effects occur even in decisions that are made only once. Indeed, the presentation order of choice options has been shown to systematically influence the results in contexts as diverse and practically relevant as, in addition to those mentioned in the introduction, evaluation of consumer brands (Brunner & Wänke, 2006), preference comparisons of paintings (MacLaughlin & Kermisch, 1997) and of musical excerpts (Koh, 1967), TV audience voting of musical performances (Li & Epley, 2009), and selection of causal explanations (McGill, 1990).

Researchers seeking a maximally unbiased preference measure should use both presentation orders and calculate the arithmetic mean of the two measures (cf. Eqs. 4a and 4b, which yield  $d_{AB} = \{[s_1 + s_1]/2\}[\psi_A - \psi_B]$ ). When this is not feasible for individual participants—for example, in polls or market research—one should give different groups different presentation orders and calculate the arithmetic means (Beike & Sherman, 1998). As is implied by Eq. 2, WOE

become largest at the ends of the valence continuum. Thus, the importance of counterbalancing the presentation order increases with stimulus valence (positive or negative).

In conclusion, the results presented here solidify the valence-dependent WOE as a real effect, and together with those of Englund and Hellström (2012a), indicate that it is caused by differential weighting of the choice options when preference comparisons are directed. This is accounted for well by Hellström's (1979, 1985, 2003) SW model with a higher weight for the subject in the comparison: That is, the subject has a greater impact on the comparison than the referent. Thus, a tendency arises to prefer the subject out of two attractive stimuli, and the referent out of two unattractive stimuli, and this tendency becomes largest at the ends of the valence continuum.

The present results and those of Englund and Hellström (2012a) demonstrate that any researcher investigating order effects in preference comparisons needs to take into account the valence level of the stimuli and to establish which stimulus, if either, is actually being used as the subject in the comparisons. When studying preference comparison of paired stimuli, performing a multiple linear regression of the preference measure on the valences of the stimuli, and determining their respective weights, should be a natural first step in exploring the process behind the comparison.

**Author note** M.P.E. was formerly at Stockholm University, from which support was obtained for this research. Partial results were presented at Fechner Day 2010, the 26th Annual Meeting of the International Society for Psychophysics, Padua, Italy, October 2010.

## References

- Agostinelli, G., Sherman, S. J., Fazio, R. H., & Hearst, E. S. (1986). Detecting and identifying change: Additions and deletions. *Journal of Experimental Psychology. Human Perception and Performance*, *12*, 445–454. doi:10.1037/0096-1523.12.4.445
- Beaver, R. J., & Gokhale, D. V. (1975). A model to incorporate within-pair order effects in paired comparisons. *Communications in Statistics*, *4*, 923–939.
- Beike, D. R., & Sherman, S. J. (1998). Framing of comparisons in research and practice. *Applied and Preventive Psychology*, *7*, 161–180. doi:10.1016/S0962-1849(05)80019-8
- Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C., et al. (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology*, *12*, 446–458. doi:10.1037/0894-4105.12.3.446
- Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*, 324–335. doi:10.1093/biomet/39.3-4.324
- Bruine de Bruin, W. (2005). Save the last dance for me: Unwanted serial position effects in jury evaluations. *Acta Psychologica*, *118*, 245–260. doi:10.1016/j.actpsy.2004.08.005
- Bruine de Bruin, W. (2006). Save the last dance II: Unwanted serial position effects in figure skating judgments. *Acta Psychologica*, *123*, 299–311. doi:10.1016/j.actpsy.2006.01.009

- Bruine de Bruin, W., & Keren, G. (2003). Order effects in sequentially judged options due to the direction of comparison. *Organizational Behavior and Human Decision Processes*, *92*, 91–101. doi:10.1016/S0749-5978(03)00080-3
- Brunner, T. A., & Wänke, M. (2006). The reduced and enhanced impact of shared features on individual brand evaluations. *Journal of Consumer Psychology*, *16*, 101–111. doi:10.1207/s15327663jcp1602\_1
- Catrambone, R., Beike, D., & Niedenthal, P. (1996). Is the self-concept a habitual referent in judgments of similarity? *Psychological Science*, *7*, 158–163. doi:10.1111/j.1467-9280.1996.tb00349.x
- Davidson, R. R., & Beaver, R. J. (1977). On extending the Bradley–Terry model to incorporate within-pair order effects. *Biometrics*, *33*, 693–702.
- Dean, M. L. (1980). Presentation order effects in product taste tests. *Journal of Psychology*, *105*, 107–110.
- Englund, M. P., & Hellström, Å. (2012a). If you have a choice, you have trouble: Stimulus valence modulates presentation-order effects in preference judgment. *Journal of Behavioral Decision Making*, *25*, 82–94. doi:10.1002/bdm.714
- Englund, M. P., & Hellström, Å. (2012b). Presentation-order effects for aesthetic stimulus preference. *Attention, Perception, & Psychophysics*, *74*, 1499–1511. doi:10.3758/s13414-012-0333-9
- Fechner, G. T. (1860). *Elemente der Psychophysik [Elements of psychophysics]*. Leipzig, Germany: Breitkopf & Härtel.
- Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York, NY: McGraw-Hill.
- Hellström, Å. (1979). Time errors and differential sensation weighting. *Journal of Experimental Psychology. Human Perception and Performance*, *5*, 460–477. doi:10.1037/0096-1523.5.3.460
- Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, *97*, 35–61. doi:10.1037/0033-2909.97.1.35
- Hellström, Å. (2003). Comparison is not just subtraction: Effects of time- and space-order on subjective stimulus difference. *Perception & Psychophysics*, *65*, 1161–1177. doi:10.3758/BF03194842
- Helson, H. (1964). *Adaptation-level theory*. New York, NY: Harper & Row.
- Houston, D. A., & Roskos-Ewoldsen, D. R. (1998). Cancellation and focus model of choice and preferences for political candidates. *Basic and Applied Social Psychology*, *20*, 305–312. doi:10.1207/s15324834basp2004\_7
- Houston, D. A., & Sherman, S. J. (1995). Cancellation and focus: The role of shared and unique features in the choice process. *Journal of Experimental Social Psychology*, *31*, 357–378. doi:10.1006/jesp.1995.1016
- Houston, D. A., Sherman, S. J., & Baker, S. M. (1989). The influence of unique features and direction of comparison on preferences. *Journal of Experimental Social Psychology*, *25*, 121–141. doi:10.1016/0022-1031(89)90008-5
- Houston, D. A., Sherman, S. J., & Baker, S. M. (1991). Feature matching, unique features, and the dynamics of the choice process: Predecision conflict and postdecision satisfaction. *Journal of Experimental Social Psychology*, *27*, 411–430. doi:10.1016/0022-1031(91)90001-M
- Houston, D. A., Sherrill-Mittleman, D., & Weeks, M. (2001). The enhancement of feature salience in dichotomous choice dilemmas. In E. U. Weber, J. Baron, & G. Loomes (Eds.), *Conflict and tradeoffs in decision making: Essays in honor of Jane Beattie* (pp. 65–85). Cambridge, UK: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
- Kardes, F. R., & Sanbonmatsu, D. M. (1993). Direction of comparison, expected feature correlation, and the set-size effect in preference judgment. *Journal of Consumer Psychology*, *2*, 39–54. doi:10.1207/s15327663jcp0201\_03
- Koh, S. D. (1967). Time-error in comparisons of preferences for musical excerpts. *The American Journal of Psychology*, *80*, 171–185. doi:10.2307/1420975
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*, 377–395. doi:10.1037/0033-295X.97.3.377
- Li, Y., & Epley, N. (2009). When the best appears to be saved for last: Serial position effects on choice. *Journal of Behavioral Decision Making*, *22*, 378–389. doi:10.1002/bdm.638
- Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.
- MacLaughlin, J. P., & Kermisch, J. (1997). Salience of compositional cues and the order of presentation in the picture reversal effect. *Empirical Studies of the Arts*, *15*, 21–27.
- Marley, A. A. J., Flynn, T. N., & Louviere, J. J. (2008). Probabilistic models of set-dependent and attribute-level best-worst choice. *Journal of Mathematical Psychology*, *52*, 281–296. doi:10.1016/j.jmp.2008.02.002
- Masin, S. C., & Agostini, A. (1991). Attentional scanning and space errors. *Perception & Psychophysics*, *50*, 285–289.
- McGill, A. L. (1990). The effect of direction of comparison on the selection of causal explanations. *Journal of Experimental Social Psychology*, *26*, 93–107. doi:10.1016/0022-1031(90)90070-3
- Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and vertical position. *Psychological Science*, *15*, 243–247.
- Patching, G. R., Englund, M. P., & Hellström, Å. (2012). Time- and space-order effects in timed discrimination of brightness and size of paired visual stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, *38*, 915–940. doi:10.1037/a0027593
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, *51*, 768–774. doi:10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1
- Phaf, R. H., & Rotteveel, M. (2009). Looking at the bright side: The affective monitoring of direction. *Emotion*, *9*, 729–733. doi:10.1037/a0016308
- Slaughter, J. E., & Highhouse, S. (2003). Does matching up features mess up job choice? Boundary conditions on attribute-salience effects. *Journal of Behavioral Decision Making*, *16*, 1–15. doi:10.1002/bdm.428
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352. doi:10.1037/0033-295X.84.4.327
- Wänke, M. (1996). Comparative judgments as a function of the direction of comparison versus word order. *Public Opinion Quarterly*, *60*, 400–409. doi:10.1086/297761
- Wänke, M., Schwartz, N., & Noelle-Neumann, E. (1995). Asking comparative questions: The impact of the direction of comparison. *Public Opinion Quarterly*, *59*, 347–372. doi:10.1086/269481
- Ward, M. F., Wender, P. H., & Reimherr, F. W. (1993). The Wender Utah Rating Scale: An aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, *150*, 885–890.