Assessing Automatic Activation of Valence

A Multinomial Model of EAST Performance

Christoph Stahl[1] and Juliane Degner[2]

[1]University of Freiburg, Germany, [2]Saarland University, Saarbrücken, Germany

Abstract. The Extrinsic Affective Simon Task (EAST; De Houwer, 2003) has been introduced as an indirect measure of automatic activation of valence. EAST effects provide nonrelative valence measures of single stimuli compared to relative measures (e.g., Implicit Association Test) that imply a comparison between two stimuli or concepts. However, EAST effects can be biased by response tendencies. A multinomial process dissociation model of EAST performance is proposed and successfully validated in four experiments. Its parameters provide pure and unbiased measures of automatic valence activation, controlled processing of task-relevant features, and response tendency. A first application of latent-class hierarchical multinomial models reveals a significant amount of parameter heterogeneity resulting from interindividual differences in accuracy motivation.

Keywords: automatic processing of valence, EAST, evaluative conditioning, latent-class hierarchical multinomial processing tree models

Introduction

When we encounter an object, its valence can be activated and processed automatically and can have – sometimes unintended – behavioral consequences (e.g., Zajonc, 1980). Recently, a number of research tools has been developed that capture such automatic activation of stimulus valence, for example the Affective Priming Paradigm (Fazio, Sanbonmatsu, Powell, & Kardes, 1986), affective versions of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) or affective variants of the Simon task (De Houwer, Crombez, Baeyens, & Herman, 2001; De Houwer & Eelen, 1998; Voss, Rothermund, & Wentura, 2003). Likewise, the Extrinsic Affective Simon Task (EAST; De Houwer, 2003) was put forward as such a measure of the valence of stimuli.

The Extrinsic Affective Simon Task

Originally, the EAST was introduced as a modified variant of the IAT (Greenwald et al., 1998), which has already been widely used as an indirect measure of the valence of stimuli or concepts but suffers from some limitations. For instance, IAT effects always imply a comparison of two concepts. Such relative IAT effects can vary remarkably depending on which contrast category is used (e.g., Karpinski, 2004). Another potential disadvantage of the IAT effect as a measure of valence is that it is based on a comparison of performance in two different tasks. This makes it vulnerable for a distortion due to different strategies used in the two tasks (e.g., Brendl, Markman, & Messner, 2001; Rothermund & Wentura, 2001, 2004). The EAST was designed to overcome these limitations: It is capable of assessing the nonrelative valence of single concepts, and is less vulnerable to distortion because EAST effects are computed from performance in one single task.

The procedure of the EAST is directly derived from the affective Simon task (De Houwer & Eelen, 1998). In an affective Simon task, participants are to respond to a series of target stimuli by giving valent responses such as pronouncing the words “good” or “bad.” The required response depends on a stimulus feature that is unrelated to valence (e.g., color). However, if stimuli vary in valence, responses are typically faster and more accurate when stimulus and required response are congruent with regard to their valence (e.g., participants are to respond “good” to a positive word), compared to trials in which their valence is incongruent (e.g., participants are to respond “bad” to a positive word). In an affective Simon task, the responses (i.e., the words “good” and “bad”) possess *intrinsic* valence. In contrast, in an EAST, responses are assumed to acquire *extrinsic* valence by means of task instructions. Participants again respond to a task-relevant feature other than valence, such as the ink color of the target stimulus. However, responses are given by pressing one of two response keys that are arbitrarily assigned the labels “positive” and “negative.” To assure that these key-presses acquire valence, a second set of valenced stimuli is introduced (attribute stimuli), that have to be explicitly evaluated. For example, a set of positive and negative at-

DOI 10.1027/1618-3169.54.2.99

tribute words printed in white ink (on a black screen) would be presented intermixed with the colored target stimuli. Participants would be instructed to press the positive key for positive white words and for all green words and the negative key for negative white words and all blue words.

As each target stimulus is presented at least once in each color, positive and negative reactions have to be given to the same stimulus; it thus serves as its own control. Consequently, the valence of a single stimulus can be determined by subtracting error rates (or response latencies) of trials with positive required responses from those of trials with negative required responses, resulting in a score with values larger than zero for positive and smaller than zero for negative evaluations.

The EAST comes with an additional benefit, namely that the underlying processes contributing to an EAST effect are much better understood than those operating in the IAT (which are subject to a considerable amount of debate; e.g., Brendl et al. 2001; De Houwer, 2001; Klauer & Mierke, 2005; Mierke & Klauer, 2003; Rothermund & Wentura, 2001, 2004). As a variant of the affective Simon task, the EAST effect is based on the compatibility between valence of stimuli and required responses (De Houwer & Eelen, 1998; De Houwer et al., 2001). Although stimulus valence is a task-irrelevant feature, it is automatically processed and causes an activation of the corresponding response (i.e., the positive response is activated upon presentation of a positive stimulus). Responding is thus facilitated in trials where stimulus and response valence are compatible, compared to trials in which both are incompatible (Beckers, De Houwer, & Eelen, 2002). Accordingly, EAST effects can indicate automatic processing of the task-irrelevant feature of stimulus valence. However, responses in the EAST are not only driven by automatic processing of stimulus valence, but also by the controlled processing of response-relevant features (e.g., color) as well as by guessing (e.g., preferences for right-hand responses). To obtain a pure measure of the valence activation process, that process needs to be dissociated from these additional processes that also contribute to EAST task performance.

Dissociating Automatic and Controlled Processes

At the core of cognitive psychology is the endeavor to disentangle the underlying processes contributing to behavior. One way to address this problem is to postulate and test a formal model of underlying cognitive processes. Multinomial processing tree models are a successful family of such models (for a review, see Batchelder & Riefer, 1999). For example, the process-dissociation procedure put forward by Jacoby (1991) was the first attempt to dissociate the contribution of conscious and unconscious processes to memory outcomes within a single task. It soon became obvious that the process-dissociation procedure could benefit from a more complex modeling approach that also incorporated guessing as an additional process that could potentially influence outcomes (e.g., Buchner, Erdfelder, & Vaterrodt-Plünnecke, 1995). In addition to their many successful applications to memory phenomena, multinomial processing tree models have also been applied to paradigms that are more central to this article's topic, for example priming paradigms, and the IAT (Payne, 2001; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005).

The ABC Model of EAST Performance

In the present article, we introduce the ABC model as a new multinomial processing tree model for EAST accuracy data. The model postulates that EAST task performance is a combined result of three cognitive processes: Automatic activation of valence (A), controlled processing of the task-relevant feature (C), and guessing (B). The model's parameters provide quantitative estimates of the contribution of each of these processes to EAST performance.

When the task-irrelevant feature of stimulus valence is processed involuntarily, an EAST effect based on stimulus valence can result. In such cases, EAST effects reflect unintended interference from an automatic activation of stimulus valence. The model posits that when stimulus valence is automatically activated, it determines the response. This automatic activation of valence is measured by the model's A parameter.

From a participant's perspective, one is best advised to concentrate on task-relevant stimulus features in order to perform accurately. This controlled processing of the task-relevant feature is captured by the model's C parameter.[1]

Guessing processes come into play when neither the task-relevant feature nor the automatic activation of valence succeed in determining the response. Response biases, for example, tendencies toward pressing the right-hand or the positive key, can influence performance. Response biases are captured by the model's B parameter. By estimating one single B parameter for all stimuli, the model equations incorporate the assumption that those biases are independent from stimulus valence.

Figure 1 illustrates how these processes are postulated to interplay in determining EAST performance. Upon encountering a target stimulus, its valence can automatically be activated and then determines the response (with probability A). If a congruent response is required, this leads to a correct response; if an incongruent response is required,

1 We do not, on a theoretical level, assume mutual exclusivity of both processes. However, a model that assumes C overriding A would result in the same pattern of estimates.

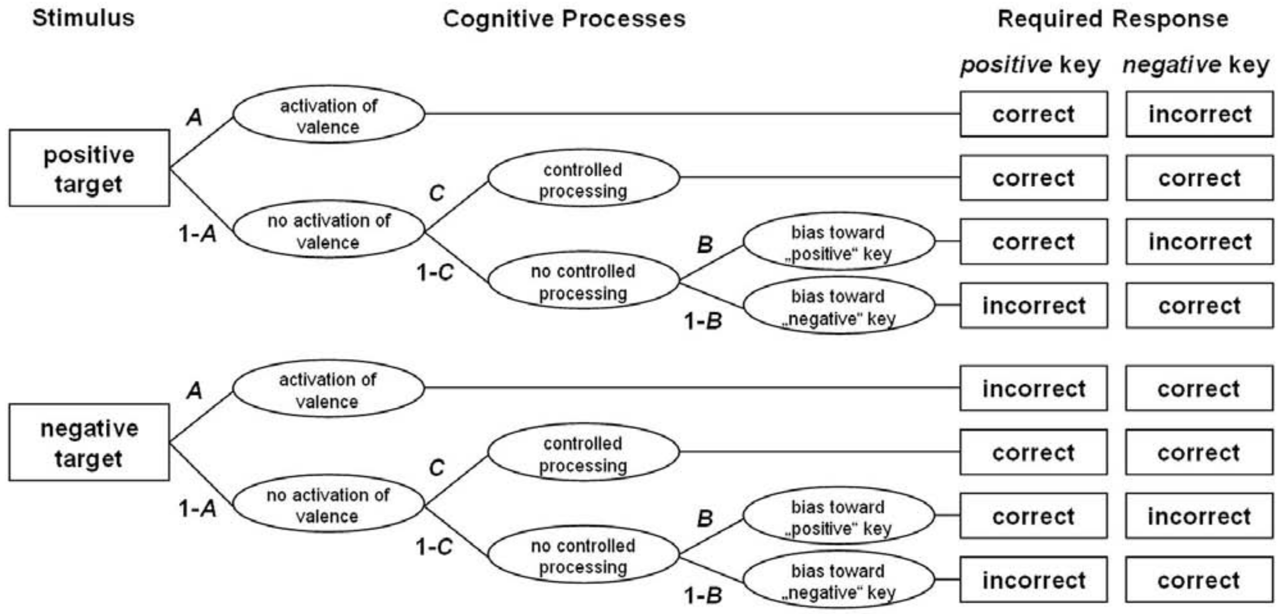

Figure 1. The ABC model of EAST performance.

an incorrect response results. Given that the task-irrelevant feature of stimulus valence is not automatically activated (with probability 1 – *A*), controlled processing of the task-relevant feature can determine the response with probability (1 – *A*)*C*; in this case, a correct response results. With probability (1 – *A*)(1 – *C*), neither automatic activation of valence nor controlled processing of the task-relevant feature determine the response; instead, a response is selected by guessing: With probability *B*, the positive key is chosen, and with probability 1 – *B*, the negative key is chosen.

Before this model can be applied to analyze EAST data, two requirements have to be met: First, it has to be established that the model is identifiable and can fit EAST data. Second, and most importantly, its parameters have to be validated: It is to be shown that the separation of the postulated processes in an EAST is psychologically meaningful and valid. For this purpose, validation experiments were conducted in which each of the model's parameters was targeted by an experimental manipulation. The parameter values of a psychologically valid model are expected to respond to experimental manipulations as predicted from previous knowledge about the effects of those manipulations on the processes in question. For the ABC model, this means that the *A* parameter (and only the *A* parameter) should respond to manipulations of stimulus valence, that the *B* parameter (and only the *B* parameter) should respond to response bias manipulations, and that the *C* parameter (and only the *C* parameter) should respond to manipulations affecting the processing of the task-relevant features. The fit and the validity of the ABC model for analyzing EAST data is demonstrated in four experiments.

Experiment 1

The first study aimed at validating the *A* parameter of automatic valence activation and the *C* parameter of controlled task processing in a standard EAST. We predicted a double dissociation: Manipulations that affect the valence of stimuli should be reflected only in changes in estimates of the *A* parameter and should not affect the *C* parameter. Manipulations that affect the controlled processing of the task-relevant feature (color) should be reflected only in changes in estimates of the *C* parameter and should have no effect on the *A* parameter. The *B* parameter of response bias should be affected by neither of those manipulations.

Three experimental conditions were realized. As a control condition, a standard EAST procedure was realized. A set of positive and negative words served as attribute stimuli and another set of positive and negative words served as target stimuli. As a task-relevant feature, target words were presented in blue or green ink. Highly similar hues of blue and green were used, such that the colors were not easy to discriminate. In this condition, we predicted the *A* parameter to capture the valence of stimuli, resulting in a parameter estimate significantly different from zero.

In the dissimilar condition, more dissimilar hues of blue and green were used. Thus, colors were easy to discriminate. Processing the task-relevant feature of ink color should be easier if the to-be-discriminated colors are less similar to each other than when they are very similar. This manipulation should increase the probability that the task-relevant feature of color determines participants' responses. We predicted that the value of the *C* parameter should increase compared to the control condition.

A nonword condition was realized in which pronounceable nonwords (e.g., *uwimor*) served as target stimuli. No automatic activation of valence should be observed for nonwords. It was thus predicted that the *A* parameter should not differ from zero in this condition. For data analysis purposes, half of the nonwords were arbitrarily categorized as "positive" or "negative."

Method

Participants

Seventy-two Saarland University students participated. Two participants' data were incomplete due to computer failure; data from three participants in the nonwords condition were excluded because of performance at chance level (error rates of .54, .56, and .48). The remaining sample consisted of 23 men and 44 women (ages 18 to 40, *Mdn* = 24). There were 21 participants in the nonwords condition and 23 participants each in the control and dissimilar conditions.

Materials

Two sets (A and B) of five positive and five negative words were used (pleasantness norms were taken from Hager & Hasselhorn, 1994). Positive words, (M = 12.81 on a scale from –20 to +20) were rated as more pleasant than negative words (M = –13.55), $t(18) = 26.85$, $p < .001$. No differences in pleasantness ratings were observed between the positive words of Set A and Set B (M = 11.85 and M = 13.76, respectively), $t(8) = 1.34$, *ns*, or between the negative words of Set A and Set B, (M = –14.68 and M = –12,41, respectively), $t(8) = 1.96$, *ns*. Assignment of sets to target versus attribute stimuli was balanced across conditions. Similarly, two sets (A and B) of nonwords were used. Nonwords were chosen for their neutrality according to pretests in which they obtained valence ratings around the midpoint of a scale from 1 *(negative)* to 7 *(positive)*. No difference was obtained in the valence ratings of both sets (M = 3.88 and M = 3.88, respectively), $t(8) = .07$, *ns*.

Attribute stimuli were presented in white ink on a black background; target stimuli were presented in blue and green colors on a black background. In the control and nonword conditions the RGB values of the blue and green colors were, respectively, 0, 125, 150, and 0, 150, 125. In the dissimilar condition, the RGB values of the blue and green colors were, respectively, 31, 82, 235 and 56, 235, 31. All word stimuli were presented in a sans serif font; letters were 8 mm of height.

Design

A 3 (Condition: control, nonword, dissimilar) × 2 (Stimulus set: A vs. B) × 2 (Color assignment: blue-negative vs. green-negative) × 2 (Valence: positive vs. negative) × 2 (Required response: positive vs. negative) design was implemented with repeated measures on the last two factors.

Procedure

The experiment was conducted on laptop computers in different locations on campus. To ensure that participants could discriminate the colors on the laptop screen, they were first presented with sample patches of the respective colors and instructed to adjust the screen to optimize discriminability. Participants were instructed to respond with a positive and a negative key to the valence of the words printed in white ink, and to respond with the same keys to the color of the words printed in blue and green ink. The positive key [M] was to be pressed with the right index finger, the negative key [C] was to be pressed with the left index finger.

Because multinomial model analyses are based on accuracy data, we modified procedure and instructions to encourage speed over accuracy. Similar to a response deadline, stimuli were presented for 200 ms and then masked, and participants received a "faster" message on late responses. These modifications should largely remove any variance of interest from latency data; RT effects are thus not reported.

Practice blocks of 20 trials each were run for the task of discriminating the valence of the attribute words and for the task of discriminating the color of the target words. Four mixed blocks of 44 trials each followed, of which the first was introduced as a practice block. Order of trials was fully randomized, except that a mixed block always started with four attribute trials. Of the remaining 40 trials, 20 were attribute trials (10 positive, 10 negative) and 20 were target trials (10 positive, 10 negative). Within one mixed block, each attribute stimulus appeared twice; each target stimulus appeared once in each color.

A trial started with a 1,000 ms pause. A fixation cross was presented in the center of the screen for 350 ms and was replaced by the stimulus, which was removed from screen after 200 ms and replaced by a black screen until a response was registered. In case of an incorrect response, an error message was displayed for 400 ms. The next trial commenced 200 ms after a response was registered or after the offset of an error message.

Results

Traditional EAST error difference scores and model-based analyses are reported. EAST error scores were computed by subtracting the proportion of errors in target trials that

Table 1. EAST scores [95% confidence intervals] for Experiments 1–4

		Stimulus valence		
Experiment	Condition	Positive	Negative	Neutral
1	Control	0.09 [0.03 0.15]	–0.07 [–0.14 0.00]	
	Nonwords	–0.07 [–0.15 0.01]	–0.05 [–0.13 0.04]	
	Dissimilar	0.06 [0.01 0.12]	–0.07 [–0.10 –0.03]	
2		0.01 [–0.02 0.04]	–0.07 [–0.10 –0.04]	–0.01 [–0.04 0.02]
3	Words	0.07 [0.03 0.10]	–0.07 [–0.11 –0.02]	
	Nonwords	–0.02 [–0.05 0.02]	–0,07 [–0.11 –0.03]	
4	Negative-Majority	–0.02 [–0.06 0.02]	–0.14 [–0.19 –0.09]	–0.08 [–0.11 –0.04]
	Positive-Majority	0.07 [0.04 0.11]	–0.05 [–0.09 –0.01]	0.06 [0.03 0.08]

Table 2. Empirical probabilities, estimated parameters, and degrees of freedom. Number of empirical probabilities and number of estimated parameters are given in brackets; degrees of freedom are computed as the difference between these numbers.

Exp.	Empirical probabilities[1]	Estimated parameters	*df*
1	Condition (control, nonwords, dissimilar) × stimulus valence (positive, negative) × required response (positive, negative) (12)	A_{Control}, A_{Nonwords}, $A_{\text{Dissimilar}}$, B, C_{Control}, C_{Nonwords}, $C_{\text{Dissimilar}}$ (7)	5
2	Stimulus valence (positive, negative, neutral) × required response (positive, negative) (6)	$A_{\text{Pos/Neg}}$, A_{Neutral}, B, $C_{\text{Pos/Neg}}$, C_{Neutral} (5)	1
3	Stimulus type (word, nonword) × stimulus valence (positive, negative) × required response (positive, negative) (8)	$A_{\text{Words,Negative}}$, $A_{\text{Words,Positive}}$, $A_{\text{Nonwords,Negative}}$, $A_{\text{Nonwords,Positive}}$, B, C_{Words}, C_{Nonwords} (7)	1
4	Majority (positive-key, negative-key) × stimulus valence (positive, negative, neutral) × required response (positive, negative) (12)	$A_{\text{NegMaj,Negative}}$, $A_{\text{NegMaj,Positive}}$, $A_{\text{NegMaj,Neutral}}$, $A_{\text{PosMaj,Negative}}$, $A_{\text{PosMaj,Positive}}$, $A_{\text{PosMaj,Neutral}}$, B_{NegMaj}, B_{PosMaj}, C (9)	3

Note. Exp. = Experiment, Neg = Negative, Pos = Positive, Maj = Majority, *df* = degrees of freedom. [1]For each cell of the design, one probability *p*(correct) is obtained.

required a positive response from the proportion of errors in trials requiring a negative response. Positive scores thus reflect positive evaluations; negative scores reflect negative evaluations.

EAST Scores

EAST scores are given in Table 1. A 3 (Condition) × 2 (Color assignment) × 2 (Valence) ANOVA with repeated measures on the last factor revealed a main effect of valence, $F(1, 61) = 20.93$, $p < .001$, $MSE = 0.012$. It was qualified by a Condition × Valence interaction, $F(2, 61) = 8.39$, $p = .001$. Separate analyses revealed that valence effects were obtained in the control and dissimilar conditions, $F(1, 21) = 21.24$, $p < .001$, and $F(1, 21) = 16.47$, $p = .001$, but not in the nonword condition, $F < 1$.

A positive EAST score was obtained for positive words in the control condition, $t(22) = 2.94$, $p < .05$, and in the dissimilar condition, $t(22) = 2.38$, $p < .05$. A negative score was obtained for negative words in the control and the dissimilar conditions, $t(22) = 2.06$, $p = .05$, and $t(22) = 4.29$, $p < .001$. The scores for the arbitrarily assigned positive and negative nonwords did not differ from zero, $t(20) = 1.80$ and $t(20) = 1.06$, *ns.*

Model Analyses

A joint model was computed for the data from all three conditions. Separate A and C parameters were estimated for the control, the dissimilar, and the nonword conditions respectively; a single B parameter was estimated (see Table 2). With $N = 5360$ data points and $\alpha = \beta = .01$, the goodness-of-fit test was able to reliably detect small effects ($w = .08$). Model fit was good, $G^2(5) = 6.10$, critical $G^2(5) = 15.09$. The model thus describes the data well, justifying the equality restriction on the B parameters. Two necessary conditions for identifiability were given: Sufficient degrees of freedom were available and repeated runs of the estimation algorithm consistently returned the same parameter estimates. We thus conclude that our model is identifiable. Parameter estimates are given in Table 3.

The A parameter was significantly different from zero in the control and dissimilar conditions, $\Delta G^2(1) = 31.69$ and $\Delta G^2(1) = 25.29$, both $p < .05$, but not in the nonword con-

Table 3. Parameter estimates [95% confidence intervals] for Experiments 1–4

			Parameters		
Experiment	Condition	Stimulus valence	A	B	C
1	Control	Pos/Neg	.08 [.05 .11]	*.45 [.40 .49]*	.85 [.82 .88]
	Nonwords	Neutral	.00 [–.04 .04]		.67 [.63 .71]
	Dissimilar	Pos/Neg	.06 [.04 .09]		.88 [.85 .91]
2		Pos/Neg	.04 [.02 .06]	*.38 [.30 .45]*	.88 [.86 .90]
		Neutral	.00 [–.03 .03]		.87 [.84 .90]
3	Words	Negative	.06 [.02 .09]	*.42 [.33 .52]*	*.89 [.86 .92]*
		Positive	.07 [.04 .11]		
	Nonwords	Negative	.04 [.01 .08]		*.87 [.84 .90]*
		Positive	.01 [–.02 .04]		
4	Negative-Majority	Negative	.09 [.05 .13]	*.28 [.20 .36]*	*.88 [.86 .90]*
		Positive	.04 [.02 .07]		
		Neutral	.04 [.00 .05]		
	Positive-Majority	Negative	.09 [.06 .12]	*.72 [.64 .81]*	
		Positive	.02 [–.02 .07]		
		Neutral	.00 [–.03 .04]		

Note. Italicized parameters values also apply to subsequent conditions and/or stimuli where a value for that parameter is omitted.

dition, $\Delta G^2(1) = 0$, *ns*. The color manipulation did not affect the *A* parameter estimates, $\Delta G^2(1) = 0.62$, *ns*. However, the color manipulation affected the *C* parameter as predicted: Dissimilar colors increased the *C* parameter estimate, $\Delta G^2(1) = 4.78$, $p < .05$, reflecting the predicted higher probability for controlled processing of stimulus color in the dissimilar compared to the control condition. Additional tests revealed that the *C* parameter was smaller in the nonword condition than in the control and the dissimilar conditions, $\Delta G^2(1) = 40.82$ and $\Delta G^2(1) = 68.9$, both $p < .05$. The *B* parameter differed from the neutral value of .5, $\Delta G^2(1) = 5.71$, $p < .05$, indicating a response tendency toward the negative key.

Discussion

The results of Experiment 1 demonstrate a dissociation between the automatic activation of valence captured by the *A* parameter and the controlled processing of the task-relevant feature of color captured by the *C* parameter. An automatic activation of valence, as measured with the *A* parameter, occurred only for stimuli that carry valence (i.e., positive and negative words), not for neutral nonword stimuli. The *A* parameter was not affected by manipulations of the task-relevant feature of ink color, whereas the *C* parameter decreased as the similarity of the colors (and thus the difficulty of the discrimination) increased. The good model fit implies that the two manipulations did not have an effect on response bias (i.e., it implies that a single *B* parameter suffices).

However, the dissociation is not a double dissociation because an unexpected effect of the valence manipulation was observed on the *C* parameter: The probability of controlled processing of stimulus color was smaller in the nonword than in the control condition. This effect can be interpreted in two ways: It can be an effect of the valence present in the words but absent in the nonwords, or it could be an effect of the word-nonword difference itself, regardless of their respective valence. In the former case, it would be a threat to the validity of the model because the model does not postulate effects of valence on the controlled processing of the task-relevant feature. In the latter case, when the effect is due to the word-nonword difference, it can be interpreted as an aspect of controlled processing (e.g., differences in strategies or in motivation of processing a standard vs. a nonwords EAST) and thus would not threaten the model's validity.

Experiment 2

To determine which of the two interpretations is correct, we conducted a second experiment in which we replaced the nonwords with neutral words, thus eliminating the valence-word/nonword confound, and implemented a within-subjects design to exclude global strategic or motivational effects. If the unexpected effect on the *C* parameter reflects differences in controlled processing of words versus nonwords, we would expect an effect of the valence manipulation on the *A* parameter and no effect of the manipulation on the *C* parameter.

To further generalize the results, we chose a different task-relevant feature: In Experiment 2, target words were flanked with either the @ or the # symbol as task-relevant stimulus features. All words were presented in white ink on a black screen. To increase the difficulty of the task (and the error rate as a basis for estimating the model parameters) the intertrial interval was reduced to 700 ms.

Subsequent to the EAST, the d2 Test of Attention (Brickenkamp, 2002) was administered. This measure was included to explore individual differences in controlled task processing captured by the *C* parameter. Results are reported in the latent-class analysis section below.

Method

Participants

Forty Saarland University students participated. One participant's data was excluded because of performance at chance level (error rate of .47). The remaining sample consisted of 18 men and 21 women (ages from 20 to 38 years, $Mdn = 24$).

Materials

The positive and negative words from Experiment 1 were used. An additional set of neutral target stimuli was selected from the Hager and Hasselhorn (1994) norms; mean pleasantness rating of $M = -0.19$ did not differ from the scale's neutral point, $t(4) = 0.32$, *ns*. Target stimuli were presented with either one of the symbols @ and # on both sides (e.g., *#happy#*) in white ink on a black screen.

The d2 test consisted of 14 rows of letters with different numbers of dashes above or below each letter. The time participants took to mark all d's with two dashes (i.e., speed) and their omission and commission errors (i.e., accuracy) served as dependent variables.

Design

A 2 (Stimulus set: A vs. B) × 2 (Symbol assignment: @-negative vs. #-negative) × 3 (Valence: positive, negative, neutral) × 2 (Required response: positive vs. negative) design was implemented with repeated measures on the last two factors.

Procedure

Procedure was identical to that of Experiment 1 with the following exceptions: In the symbol discrimination practice blocks, target words were presented once with each of the two symbols, resulting in 30 trials. The mixed blocks consisted of 64 trials each, of which the first four were attribute trials. Of the remaining 60 trials, 30 were attribute trials (15 positive, 15 negative) and 30 were target trials (10 positive, 10 negative, 10 neutral). Within one mixed block, attribute stimuli appeared three times; target stimuli appeared once with each symbol. Trials started with a 700 ms pause. Following the EAST procedure, participants completed the d2 test that was introduced as a concentration test. Participants were instructed to work as fast and as accurately as possible.

Results

EAST Scores

The EAST scores are given in Table 1. A 2 (Stimulus set) × 2 (Symbol assignment) × 3 (Valence) repeated-measures ANOVA revealed only a main effect of valence, $F(2, 70) = 10.00$, $p < .001$. The EAST score for negative words was significantly below zero, $t(38) = 4.84$, $p < .001$, but the EAST score for positive words did not differ from zero, $t(38) = 0.83$, *ns*. The EAST score obtained for neutral words did not differ from zero, $t(38) = 0.88$, *ns*.

Model Analyses

Separate *A* and *C* parameters were estimated for valenced and neutral target words; one single *B* parameter was estimated (see Table 2). With $N = 4680$ and $\alpha = \beta = .01$, the goodness-of-fit test was again able to reliably detect small effects ($w = .07$). Model fit was good, $G^2(1) = 1.21$, critical $G^2(1) = 6.63$; parameter estimates are given in Table 3.

Estimates of the *A* parameter differed between valenced and neutral stimuli, $\Delta G^2(1) = 9.03$, $p < .05$. Additional tests revealed that, as predicted, the estimate of *A* for valenced words was significantly larger than zero, $\Delta G^2(1) = 18.06$, $p < .05$, whereas the estimate for neutral words did not differ from zero, $\Delta G^2(1) = 0.01$, *ns*. Estimates of the *C* parameter did not differ between positive/negative and neutral stimuli, $\Delta G^2(1) = 0.33$, *ns*. The *B* parameter estimate differed from the neutral value of .5, $\Delta G^2(1) = 9.65$, $p < .05$, reflecting again a tendency toward the negative key.

Discussion

As predicted, a dissociation was found between the *A* and *C* parameters. A within-subjects manipulation of stimulus valence (positive/negative versus neutral words) affected the automatic activation of valence, and this was reflected in the *A* parameter. The valence manipulation did not affect the controlled processing of the task-relevant feature; accordingly, and as predicted, no effect was found on the *C* parameter. Response bias *B* was not affected by the valence manipulation, as the model fit implies. Combined with the results of Experiment 1, these findings validate the model parameters: The *A* parameter measures task-irrelevant interference of stimulus valence on valent responses in the EAST and is not affected by manipulations of task difficulty. The *C* parameter measures processing of task-relevant features and is not affected by the valence of stimuli.

Experiment 3

A third experiment was conducted to replicate the above dissociation and to demonstrate the model's usefulness as a tool to tap affective processes. In the previous experiments, the finding of a positive evaluation of clearly and inherently positive words and a negative evaluation of clearly negative words is not much of a surprise. To be useful as a measurement tool, the model should be capable of capturing more subtle evaluations. In this experiment we used evaluatively conditioned nonwords as target stimuli, along with positive and negative words. We predicted that the model would accurately reflect effects of the acquired valence of evaluatively conditioned neutral stimuli. To test that prediction, four separate *A* parameters for negative words, positive words, conditioned-negative nonwords, and conditioned-positive nonwords were estimated.

Method

Participants

Twenty-four University of Freiburg students participated. One participant's data was excluded because of performance at chance level (error rate of .51). The remaining sample consisted of 11 males and 12 females (ages from 19 to 32, *Mdn* = 23).

Materials

The positive and negative words and nonwords from Experiment 1 were used. EAST target stimuli were presented in blue or green ink on a black screen. For the evaluative conditioning phase, a set of 25 positive and a set of 25 negative IAPS pictures (Lang, Bradley, & Cuthbert, 2005) were used, mean pleasantness ratings: $M = 7.9$ and $M = 2.5$, respectively, $t(48) = 117.80$, $p < .001$.

Design

A 2 (stimulus set: A vs. B) × 2 (color assignment: blue-negative vs. green-negative) × 2 (stimulus type: words vs. nonwords) × 2 (valence: positive vs. negative) × 2 (required response: positive vs. negative) design was implemented with repeated measures on the last three factors.

Procedure

The evaluative conditioning phase was introduced as a simple learning task. Participants saw pairs of pictures and nonwords on the computer screen and were told to memorize the pairs for a later memory test. Each nonword-picture pair was presented on screen for 2,500 ms; no response was required. Throughout the conditioning phase, each nonword was paired with five pictures of the same valence, resulting in 25 pairs of nonwords and pleasant pictures and 25 pairs of nonwords and unpleasant pictures, presented in a fully randomized sequence. The complete set of 50 pairs was presented twice in immediate succession.

The EAST procedure was identical to that of Experiment 1 with the following exceptions: In the color discrimination practice blocks, each target item was presented once in each of the two colors, resulting in 40 trials. Mixed blocks consisted of 84 trials each. The first four were attribute trials; of the remaining 80 trials, 40 were attribute trials (20 positive, 20 negative) and 40 were target trials (10 positive words, 10 negative words, 10 positive nonwords, 10 negative nonwords). Within one mixed block, attribute stimuli appeared four times, and target stimuli appeared once in each color.

After completing the EAST, participants were asked to rate the pleasantness of each of the target words and nonwords on an 8-point Likert scale (1 = *very unpleasant*, 8 = *very pleasant*).

Results

Pleasantness Ratings

Ratings of the positive and negative words ($M = 6.70$ and $M = 2.63$) significantly differed from the scale midpoint of 4.5, $t(22) = 6.16$ and $t(22) = 5.45$, respectively, both $p < .001$. Ratings of negative nonwords ($M = 3.73$) but not of positive nonwords ($M = 4.58$) differed from the scale midpoint, $t(22) = 3.17$, $p < .05$, and $t(22) = 0.37$, *ns*, respectively. Ratings of positive words and positive nonwords as well as of negative words and negative nonwords differed from each other, $t(22) = 5.54$, $p < .001$, and $t(22) = 2.92$, $p < .05$, for words and nonwords, respectively. Thus, although the mean rating of positive nonwords did not differ from the scale midpoint, and the valence effects were weaker than those obtained for words, the conditioning procedure was successful in manipulating the valence of the presented nonwords.

EAST Scores

EAST scores are given in Table 1. A 2 (Stimulus type: words vs. nonwords) × 2 (Valence: positive vs. negative) repeated-measures ANOVA revealed a main effect of Stimulus type, $F(1, 22) = 7.56$, $p < .05$, $MSE = .006$, indicating a more positive EAST score for words ($M = 0$) than for nonwords ($M = –0.04$). A main effect of Valence, $F(1, 22) = 22.50$, $p = .001$, $MSE = 0.009$, indicated a more positive EAST score for positive items ($M = 0.03$) than for negative items ($M = –0.07$). A marginally significant Stimulus type × Valence interaction was found, $F(1, 22) = 4.29$, $p = .05$, *MSE* = 0.007. Separate *t* tests revealed that it was due to the lack of a significant positive score for positive nonwords: Mirroring the explicit pleasantness ratings, a negative score was obtained for negative words, $t(22) = 2.84$, and negative nonwords, $t(22) = 3.76$, both $p < .05$, and a positive EAST score was obtained for positive words, $t(22) = 4.11$, $p < .05$, but not for positive nonwords, $t(22) = 0.98$, *ns*. However, EAST scores were significantly more positive for positive (compared to negative) words, $t(22) = 4.17$, $p < .001$, and for positive (compared to negative) nonwords, $t(22) = 2.54$, $p = .02$.

Model Analyses

Separate *A* parameters were estimated for positive and negative words and nonwords. Note that an *A* parameter estimate greater than zero is interpreted differently depending on the valence of the stimuli: It reflects the automatic activation of positive valence for positive stimuli, whereas for negative stimuli, it reflects the activation of negative valence.[2] A single *B* parameter was estimated. Separate *C* parameters were estimated for words and nonwords (see Table 2). With $N = 3680$ and $\alpha = \beta = .01$, the goodness-of-fit test was again able to reliably detect small effects ($w = .08$). Model fit was good, $G^2(1) = 2.72$, critical $G^2(1) = 6.63$; parameter estimates are given in Table 3.

As predicted, estimates of the *A* parameter were significantly different from zero for positive and negative words, $\Delta G^2(1) = 18.19$, and $\Delta G^2(1) = 8.66$, respectively, both $p < .05$. For negative nonwords, the estimate of *A* was significantly larger than zero, $\Delta G^2(1) = 5.2$, $p < .05$. For positive nonwords, this was not the case, $\Delta G^2(1) = 0.68$, *ns*, accurately reflecting the absence of an evaluative conditioning effect in the pleasantness ratings. Estimates of the *C* parameter did not differ between words and nonwords, $\Delta G^2(1) = 0.51$, *ns*. The *B* parameter did not differ from the neutral value of .5, $\Delta G^2(1) = 2.36$, *ns*.

Discussion

The evaluative conditioning procedure was partially successful in manipulating participants' evaluation of previously neutral stimuli. Nonwords paired with negative pictures were evaluated more negatively than nonwords paired with positive pictures, as reflected in the pleasantness ratings and in EAST error scores. However, only the evaluation of nonwords paired with negative pictures differed from the neutral point of the scale. The evaluative conditioning effects were accurately captured by the *A* parameter. No difference was observed in controlled processing between words and evaluatively-conditioned nonwords. These effects replicate the dissociations obtained in Experiments 1 and 2, demonstrating that valence manipulations are adequately reflected in the *A* parameter and do not affect the *C* parameter.

The results from Experiment 3 extend the above findings by demonstrating (a) that the *C* parameter is not affected by a within-subjects manipulation of the nature (word or nonword) of the stimulus, and (b) that separate *A* parameter estimates can be obtained within a single EAST experiment and that these separate estimates adequately capture the valence of different sets of stimuli. The present results also demonstrate that the model parameters can not only capture well established valence associations but can also accurately reflect newly acquired valence originating from a short (4 min) acquisition phase. In addition, and equally important, the results demonstrate that parameter estimates accurately reflect the nonrelative valence of single sets of stimuli: Mirroring the pleasantness ratings, an automatic activation of valence was found for negative and positive words, but only for negative, not for positive nonwords.

Experiment 4

Response biases pose a potential threat to the validity of the EAST score as a nonrelative measure of valence. To illustrate this, imagine a participant who responds with the positive key to every trial in an EAST. The error rate for stimuli requiring a positive-key response would be zero, whereas the error rate for stimuli requiring a negative-key response would be at a maximum. Thus, a maximally positive EAST score would result, regardless of the valence of the stimuli. The EAST score does not allow one to discriminate whether this result reflects an evaluation of the target stimuli (as one would have to assume) or mere response bias (as would in fact be the case).

Response bias however is unproblematic if the ABC model is used to analyze EAST effects because it will be captured by the *B* parameter and leave the *A* parameter unaffected. This ability to separately assess the contributions of different cognitive processes is a core advantage of the multinomial modeling approach over simple difference scores.

A fourth experiment was conducted to establish the unbiased nature of the model parameters as measures of the underlying cognitive processes and to validate the response bias parameter *B*. To this end, the proportion of positive vs. negative key presses was manipulated between participants to induce a bias toward one of the response keys. Two conditions were realized in which either 41.7% or 58.3% of responses were to be given with the

2 This is due to an a-priori assumption built into the model; it is necessary because parameter estimates range between 0 and 1 and thus do not allow negative values. If, for example, all *A* parameters were coded such that a value greater than zero would reflect a positive valence, effects of negative valence would result in a bad model fit.

negative key. Positive, negative, and neutral words were used as target stimuli. We expected the EAST error difference score to be biased toward the evaluation of the majority key. When the majority of responses is to be given with the positive key, participants should tend to use the positive key more often in cases of uncertainty, even when pressing the negative key is required for a correct response. This should increase the probability of an error in trials that require a negative response, and lead to a more positive EAST score, compared to when the majority of responses is to be given with the negative key. Response bias will be captured by the *B* parameter and leave the *A* and C parameters unaffected.

Method

Participants

Sixty-four University of Freiburg students (27 male, 37 female; ages from 18 to 42, *Mdn* = 22) participated in the study.

Materials

The positive, negative, and neutral words from Experiment 2 were used. In addition, neutral filler words were used for the proportion manipulation. Target stimuli were presented with either one of the symbols @ or # in white ink on a black screen.

Design

A 2 (Proportion of negative-key responses: 58.3% vs. 41.7%) × 2 (Stimulus set: A vs. B) × 2 (Symbol assignment: @-negative vs. #-negative) × 2 (Response key assignment: negative-right vs. negative-left) × 3 (Valence: positive, negative, neutral) × 2 (Required response: positive vs. negative) design was implemented with repeated measures on the last two factors.

Procedure

Procedure was identical to that of Experiment 2 with the following exceptions: In the symbol discrimination practice blocks, each target word was presented once with each of the two symbols. Five filler trials were presented that were mapped onto the majority key, resulting in 35 trials. The mixed blocks consisted of 124 trials each. The first four were attribute trials; of the remaining 120 trials, 60 were attribute trials (30 positive, 30 negative), 30 were target trials (10 positive, 10 negative, 10 neutral), and 30 were neutral filler trials (5 mapped onto the minority key and 25 mapped onto the majority key; five of the latter type were presented immediately after the initial four attribute trials). Within one mixed block, each attribute stimulus appeared six times; each target stimulus appeared once with each symbol. Thus, in a mixed block, 70 of 120 trials (58.7%) required a response with the majority key, and the remaining 50 trials (41.7%) required the alternative response.

Results

EAST Scores

EAST scores are given in Table 1. A 2 (Key proportion: negative-majority vs. positive-majority) × 2 (Key assignment: left = negative vs. right = negative) × 2 (Symbol assignment: @ = negative vs. # = negative) × 2 (Sets: A vs. B) × 2 (Stimulus valence: positive, neutral, negative) ANOVA of the EAST scores with repeated measures on the last factor revealed the predicted main effect of key proportion, $F(1, 48) = 41.81$, $p < .001$, $MSE = 0.013$. In the negative-majority condition, EAST scores were more negative than in the positive-majority condition ($M = -0.08$ and $M = 0.03$, respectively). A main effect of Valence was also found, $F(2, 96) = 28.95$, $p < .001$, $MSE = 0.009$; separate *t* tests revealed that positive stimuli ($M = 0.03$) scored more positively than neutral stimuli ($M = -0.01$), $t(63) = 2.59$, $p < .05$, and that neutral stimuli scored more positively than negative stimuli ($M = -0.09$), $t(63) = 4.94$, $p < .001$. An interaction of stimulus valence with the symbol assignment factor was observed, $F(2, 96) = 4.45$, $p < .05$, indicating that negative stimuli scored more negatively when the @ symbol was assigned to the negative key, $F(1, 48) = 10.43$, $p < .05$; no effects of symbol assignment were observed on positive and neutral stimuli, both $F < 1$. This interaction did not qualify the above main effect: Negative stimuli scored more negatively than neutral and positive stimuli in both symbol assignment conditions, smallest $t(63) = 2.43$, largest $p = .02$.

The EAST difference scores are reported here to illustrate the distorting effect of the proportion manipulation. Independent-sample *t* tests confirmed that the proportion manipulation affected the EAST scores for all stimuli, smallest $t(62) = 3.1$, largest $p = .003$. For neutral words, a negative EAST score was obtained under the negative-majority condition, $M = -0.08$, $t(31) = 4.63$, $p < .001$, whereas a positive score was obtained under the positive-majority condition, $M = 0.06$, $t(31) = 5.04$, $p < .001$. For positive words, the EAST score did not differ from zero under the negative-majority condition, $M = -0.02$, $t(31) = 0.84$ *ns*, whereas it was positive in the positive-majority condition, $M = 0.07$, $t(31) = 4.36$, $p < .001$. For negative words, a strongly negative score was obtained under the negative-majority condition, $M = -0.14$, $t(31) = 6.1$, $p < .001$. It was greatly reduced, but still significantly smaller than zero, under the positive-majority condition, $M = -0.05$, $t(31) = 2.7$, $p < .05$.

Model Analyses

Separate *A* parameters were obtained for positive, negative, and neutral target words in the negative-majority and the positive-majority conditions.[3] Separate *B* parameters were estimated for the negative-majority and positive-majority conditions; one single *C* parameter was estimated across conditions (see Table 2). With $N = 7680$ data points and $\alpha = \beta = .01$, the goodness of fit test was able to reliably detect small effects ($w = .06$). Model fit was again good, $G^2(3) = 4.86$, critical $G^2(3) = 11.34$. Parameter estimates are given in Table 3. The predicted effect of the proportion manipulation on the *B* parameter was obtained, $\Delta G^2(1) = 38.1$, $p < .001$; a *B* estimate smaller than .5 was obtained under the negative-majority condition, $\Delta G^2(1) = 21.92$, $p < .001$, indicating a bias toward the negative-key, and a *B* estimate larger than .5 was obtained under the positive-majority condition, $\Delta G^2(1) = 18.78$, $p < .001$, indicating a bias toward the positive-key. Also as predicted, the *A* parameters were unaffected by the response bias manipulation, $\Delta G^2(3) = 4.49$, *ns*. Instead, the *A* parameters accurately captured the evaluation of target stimuli: For negative words, an estimate significantly larger than zero resulted, $\Delta G^2(2) = 69.13$, $p < .001$, indicating a negative evaluation. For positive words, an estimate significantly larger than zero, $\Delta G^2(2) = 16.21$, $p < .001$, indicated a positive evaluation; whereas for neutral words, the estimate did not differ from zero, $\Delta G^2(2) = 4.91$, *ns*, indicating a neutral evaluation.

Discussion

The proportion manipulation successfully induced a response bias toward the majority key, resulting in the predicted distortion of EAST scores but leaving the *A* parameters unaffected. This clearly demonstrates the unbiased nature of the *A* parameter as a measure of automatic activation of valence, and thus, the superiority of an analysis of EAST effects within the framework of the ABC model.

A Latent-Class Hierarchical Analysis

A potential threat for the validity of multinomial models is the problem of parameter heterogeneity. In most applications, data are aggregated across participants for analysis. In doing so, one assumes that parameter values are equal across participants. If this assumption is violated, α levels of significance tests are inflated above the nominal level, leading to erroneous rejection of models due to lack of fit (Klauer, 2006). Thus, whenever a multinomial model is rejected by traditional goodness-of-fit tests, this may be due to its not describing the mean category frequencies adequately, or, alternatively, because of parameter heterogeneity, despite its being an adequate description of the data. Whenever a multinomial model achieves a good fit in traditional tests, there might however still be a sizeable amount of parameter heterogeneity, potentially increasing the actual α level of significance tests in hypotheses tests and thereby leading to erroneous substantial conclusions. In applications of multinomial modeling, the homogeneity assumption has rarely ever been evaluated, even in domains where heterogeneity can be expected on theoretical grounds, because tests were not readily available until recently.

Klauer (2006) introduced a new framework, called latent-class hierarchical multinomial models, that provides statistical tests to assess the homogeneity assumption and that constitutes a new family of models that can be applied if the parameter homogeneity assumption is violated. Hierarchical multinomial models differ from traditional multinomial processing tree models in that they provide separate parameter estimates for a specified number of latent classes, and estimates of the proportion of participants that belong to each latent class. They are equivalent to traditional multinomial models in case of a single latent class. If parameter homogeneity is violated in the single-class model, a model with two or more latent classes can be computed, allowing for different parameter values in each class, and thus capturing the parameter heterogeneity. In a nutshell, hierarchical multinomial models provide different sets of parameter estimates for each of two or more latent classes.

To evaluate whether parameter heterogeneity was present and whether it compromised the above validation results, we reanalyzed the data from the four experiments. Although there was substantial parameter heterogeneity in all four data sets, the validation results obtained with the traditional analyses were confirmed. The majority of the extant parameter heterogeneity was located in the *C* parameter assessing controlled task processing. To illustrate the application of hierarchical multinomial models, the reanalysis is reported in detail for Experiment 2.[4]

As a first step, parameters and goodness-of-fit tests for a single class were computed. This single-class solution achieved good fit to the mean category frequencies, as indicated by the traditional log-likelihood ratio statistic $G^2(1) = 1.21$, $p = .11$, and the mean structure test statistic M_1 suggested by Klauer (2006), $M_1(1) = 0.36$, $p = .55$. However, parameter homogeneity was violated, as indicated by the variance-covariance structure tests $S_1(16) = 465.09$, and $S_2(21) = 565.27$, both $p < .05$. In a second step, two latent classes were computed that were free to differ in their parameter values. If the diagnosed parameter heterogeneity is adequately captured by the two-class model, the variance-covariance structure tests S_1 and S_2 should be ren-

3 Model equations for neutral stimuli were coded bias-congruent (i.e., neutral stimuli were treated as negative stimuli in the negative-majority and as positive stimuli in the positive-majority condition) to enable the model to capture valence effects as suggested by the EAST scores.

4 For the sake of brevity, we omit detailed reanalyses of the other data sets.

Table 4. Parameter estimates [95% confidence intervals] of the two-class solution of the latent-class hierarchical model applied to the data from Experiment 2

			Parameters		
Latent class	Class weight	Stimulus valence	A	B	C
1	.61	Pos/Neg	.02 [.01 .04]	*.31 [.17 .45]*	.94 [.91 .97]
		Neutral	.00 [–.05 .05]		.92 [.87 .98]
2	.39	Pos/Neg	.06 [.02 .10]	*.46 [.37 .55]*	.77 [.72 .82]
		Neutral	.00 [–.09 .09]		.70 [.62 .78]

Note. Italicized parameters values also apply to subsequent conditions and/or stimuli where a value for that parameter is omitted.

dered insignificant. This was indeed the case, $S_1(11) = 6.5$, $p = .84$, and $S_2(21) = 28.85$, $p = .12$. Parameter estimates for the two-class solution are reported in Table 4.

To check whether the results obtained with the traditional analyses were distorted by parameter heterogeneity, we repeated the hypotheses tests reported above. Results were confirmed for the *A* parameter: It differed between valenced and neutral words, $\Delta l(2) = 11.59$, $p < .05$, and was larger than zero for valenced words, $\Delta l(2) = 17.17$, $p < .05$, but not for neutral words, $\Delta l(2) = 0$, *ns*.[5] Results were also confirmed for the *C* parameter: No difference between valenced and neutral words was found, $\Delta l(2) = 4.64$, $p = .10$.

Next, the source of parameter heterogeneity was located by testing equality restrictions across classes. Setting the *A* parameters equal across classes did not lead to a significant loss of fit, $\Delta l(2) = 2.33$, *ns*. We conclude that participants did not differ in their automatic activation of valence. The test reached significance for the *B* parameter restriction, $\Delta l(1) = 4.18$, $p < .05$, indicating that response biases differed between classes. The test also reached significance for the *C* parameters, $\Delta l(2) = 76.91$, $p < .05$, indicating that the probability of controlled processing of the task-relevant feature differed between classes.

The latent-class analysis allowed us to test the parameter homogeneity assumption. We found that it was violated: Whereas the larger class of participants showed a very high probability of controlled responses, leaving only a small proportion of responses to be determined by guessing (with a tendency toward the negative key), a minority of participants had a much smaller probability of controlled processing, leading to a much larger proportion of responses determined by guessing; for those participants, no response bias was observed.

The main difference between latent classes was found in controlled processing. This might be due to differences in focus on speed versus accuracy or, to put it differently, differences in motivation to process the task accurately and diligently. To strengthen this interpretation, and to validate the latent-class analysis, we investigated whether differences in the *C* parameter were reflected in differences in speed versus accuracy of performance on the d2 test. Participants were assigned to one of two groups based on their posterior probability of latent-class membership (with a cutoff criterion of .5).[6] The resulting groups were compared with regard to mean processing time, and mean omission and commission errors in the d2 test in a 2 (stimulus set) × 2 (symbol assignment) × 2 (class membership) multivariate ANOVA. The groups differed in the mean time they took for completion of the d2 test, $F(1, 29) = 4.21$, as well as in the mean rates of omission and commission errors, $F(1, 29) = 5.66$, and $F(1, 29) = 5.20$, all $p < .05$. Group 1 participants (those with a high posterior probability of belonging to latent class 1) took about a minute longer to complete the test ($M = 438$ s and $M = 382$ s, for Groups 1 and 2, respectively), and made less errors of omission ($M = 10.54$) and commission ($M = 0.47$) than Group 2 participants ($M = 34.74$ and $M = 1.45$). Group 1 participants processed both the d2 task and the EAST very accurately and diligently, whereas Group 2 participants focused on speed in both tasks and did not reach high levels of accuracy. This result provides support for the validity of the latent-class analysis and corroborates the interpretation that differences in *C* parameter estimates between latent classes in the EAST are due to individual differences in focus on speed versus accuracy of performance.

Across all four experiments, analyses within the latent-class hierarchical framework have revealed that (a) whereas the mean structure was described well by the traditional models, parameter homogeneity assumptions were violated in all data sets, (b) nevertheless, the results pertaining to the mean parameter estimates obtained with the traditional analyses were confirmed, (c) participants differed in the degree of diligence or accuracy motivation, as reflected in differences in controlled processing of the task-relevant feature.

General Discussion

A multinomial process dissociation model of EAST performance was introduced and successfully validated in four experiments. In all four experiments, differences in automatic activation of stimulus valence were reflected in dif-

5 $\Delta l = (-2 \log(\text{likelihood})_{\text{restricted}}) - (-2 \log(\text{likelihood})_{\text{unrestricted}})$

6 Probability of an individual's membership in class *c*, given the individual's category frequencies and the parameter estimates for class *c*.

ferences in parameter *A*. In Experiment 1, differences in controlled task processing were captured by parameter *C*. In Experiment 4, response bias was shown to be accurately reflected in estimates of parameter *B*.

Double dissociations were obtained for the *A* and *C* parameters. In Experiment 1, a controlled task processing manipulation affected *C* but left *A* unaffected. In Experiments 2 and 3, valence manipulations affected *A* but left *C* unaffected. The *A* and *B* parameters were also successfully dissociated: In Experiment 4, a response bias manipulation affected *B* but left *A* unaffected, and the good model fit across all experiments implies that response biases were unaffected by the valence manipulations in Experiments 1 through 3.

In Experiments 3 and 4, it was shown that *A* constitutes a *nonrelative* measure of the automatic activation of valence of single sets of stimuli (as opposed to a relative measure comparing two sets of stimuli) that accurately reflects the valence of the stimulus sets (Experiment 3). The *A* parameter was shown to be an *unbiased* measure as it was unaffected by a response bias manipulation (Experiment 4) that substantially distorted the EAST error score.

In sum, the ABC model provides a framework for separately measuring the contribution of the three postulated processes to EAST performance. The model's *A* parameter can be considered superior to the EAST score as a measure of the automatic activation of valence because it is unaffected by differences in controlled processing and/or response biases. By analyzing EAST data with the ABC model, one can benefit from the EAST's ability to assess the nonrelative valence of single (sets of) stimuli, which puts it at an advantage over other, relative measures of automatic activation of valence. This might be especially interesting if the EAST is applied as an indirect measure of automatic components of attitudes, as in clinical or social psychology.[7] With the ABC model, we thus provide a measurement tool that can help researchers and practitioners make use of the advantages of the EAST.

Comparison with Other Models

Related multinomial models have been proposed for similar tasks, for example Jacoby's (1991) model (cf. Payne, 2001), and the QUAD model (Conrey et al., 2005). We consider these models less appropriate for EAST data. As pointed out above, to obtain unbiased parameter estimates, a model must be able to account for guessing tendencies, which the Jacoby (1991) model cannot accomplish. The QUAD model includes such a guessing parameter, and in addition, a fourth parameter *(OB)* that aims at capturing the process of overcoming automatic activation. Capturing this process would also have been of theoretical interest in the present research. However, it appears that *OB* estimates are often not useful because of large confidence intervals. For example, in their Experiment 5, Conrey et al. (2005) reported *OB* estimates of 1.0 and 0.0 that could be set equal without significant loss of fit. Furthermore, for technical reasons (i.e., lack of degrees of freedom), it was not possible to apply the QUAD model to EAST data. As a consequence, we developed and validated the three-parameter ABC model. The results show that EAST data can be accounted for with the three postulated parameters, rendering it unlikely that the process of overcoming automatic activation substantially contributes to EAST performance.

Diffusion models (e.g., Ratcliff & Rouder, 1998) could provide additional insight into the underlying processes of the EAST because they consider accuracy and latency data simultaneously. However, estimating diffusion model parameters requires complex algorithms and extensive computations, and user-friendly software is not currently available. These obstacles render diffusion models practically inapplicable for most EAST researchers and practitioners. For the future, however, we believe diffusion-model approaches to EAST data are a promising route.

Parameter Heterogeneity

The present model can be used to obtain parameter estimates for individual participants; those estimates can subsequently be entered into correlational analyses. From the theoretical standpoint, the *A* parameter clearly is a more valid measure than the traditional EAST score. The question whether the *A* parameter empirically proves to provide a more valid measure of individuals' attitudes needs to be addressed in future research.

Where individual parameter estimates are not required, latent-class hierarchical multinomial models can be used to diagnose and capture interindividual differences in parameter estimates. With an easy-to-use computer program (*HMMTree*; Stahl & Klauer, in press), latent-class hierarchical analyses can be applied to traditional multinomial processing tree models, given that individual category frequencies are available. We presented a first and successful application of this new family of models, and the results clearly show that they are a useful addition to the researcher's toolbox.

Acknowledgments

The authors thank Jan de Houwer, Christoph Klauer, and an anonymous reviewer for valuable comments.

7 The ABC model can in principle be applied to every set of EAST data with at least four empirical categories (e.g., a single set of target stimuli and an experimental factor with two levels, or two sets of target stimuli). A substantial proportion of errors is however required. The present data had mean error rates of 11.7%, 8.2%, 9.9%, and 8.0%, for Experiments 1–4, respectively.

References

Batchelder, W.H., & Riefer, D.M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin and Review, 6*, 57–86.

Beckers, T., De Houwer, J., & Eelen, P. (2002). Automatic integration of nonperceptual action effect features: The case of the associative affective Simon effect. *Psychological Research, 66*, 166–173.

Brendl, C.M., Markman, A.B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 760–773.

Brickenkamp, R. (2002). *Aufmerksamkeits-Belastungs-Test* [The D2 Test of Attention] (9th ed.). Göttingen: Hogrefe.

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General, 124*, 137–160.

Conrey, F., Sherman, J.W., Gawronski, B., Hugenberg, K., & Groom, C.J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology, 89,* 469–487.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443–451.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50*, 77–85.

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition and Emotion*, *15*, 189–206.

De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion, 12,* 45–62.

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., & Kardes, F.R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Hager, W., & Hasselhorn, M. (Eds.). (1994). *Handbuch deutschsprachiger Wortnormen* [Handbook of German word norms]. Göttingen: Hogrefe.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin, 30*, 22–34.

Klauer, K.C. (2006). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika, 71,* 1–31.

Klauer, K.C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin, 31*, 208–217.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). *International affective picture system (IAPS): Digitized photographs, instruction manual and affective ratings.* Technical Report A-6, University of Florida, Gainesville, FL.

Mierke, J., & Klauer, K.C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85*, 1180–1192.

Payne, B.K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*, 181–192.

Ratcliff, R., & Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*, 347–356.

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift für Experimentelle Psychologie, 48*, 94–106.

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General, 133*, 139–165.

Stahl, C., & Klauer, K.C. (in press). *HMMTree: A computer program for hierarchical multinomial processing tree models*. Behavior Research Methods.

Voss, A., Rothermund, K., & Wentura, D. (2003). Estimating the valence of single stimuli: A new variant of the affective Simon task. *Experimental Psychology, 50*, 86–96.

Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.

Received December 22, 2005
Revision received April 11, 2006
Accepted April 11, 2006

Christoph Stahl

Institut für Psychologie
Universität Freiburg
D-79085 Freiburg
Germany
Tel. +49 761 203 2418
Fax +49 761 203 2417
E-mail stahl@psychologie.uni-freiburg.de

