Effects in the Affect Misattribution Procedure are Modulated by

Feature-Specific Attention Allocation

Tom Everaert, Adriaan Spruyt, and Jan De Houwer

Ghent University

Keywords: feature-specific attention allocation, Affect Misattribution Procedure, implicit

measures, automatic stimulus evaluation, automaticity, attitudes

Address correspondence to:

Tom Everaert

Department of Psychology

Ghent University

Henri Dunantlaan 2

B-9000 Ghent, Belgium

TEL: ++32-9-264-91-07

FAX: ++32-9-264-64-89

E-mail: Tom.Everaert@UGent.be

Abstract

We examined whether automatic stimulus evaluation as measured by the Affect Misattribution Procedure (AMP) is moderated by the degree to which attention is assigned to the evaluative stimulus dimension (i.e. feature-specific attention allocation, FSAA). In two experiments, one group of participants completed a standard AMP while attending to evaluative stimulus information. A second group of participants completed the AMP while attending to non-evaluative stimulus information. In line with earlier work, larger AMP effects were observed when participants were encouraged to attend to evaluative stimulus information than when they were not. These observations support the idea that the impact of FSAA on measures of automatic stimulus evaluation results from a genuine change in the degree of automatic stimulus evaluation rather than a change in the degree to which automatic stimulus evaluation is picked up by these measures.

Effects in the affect misattribution paradigm are modulated by feature-specific attention allocation

The capacity to select information as a function of task demands and (long-term) goals is an essential requirement for adaptive behavior. One mechanism that allows for such an automatic selection of goal-relevant information is feature-specific attention allocation (hereafter referred to as *FSAA*). It is the extent to which attention is assigned to one specific stimulus dimension (e.g., size, shape, color, etc.) or even one particular stimulus feature (e.g., large, circular, red, etc.) as a function of current goals and task-demands (Medin & Schaffer, 1978; Nosofsky, 1986; also see Kiefer & Martens, 2010; Kiefer, 2012; Spruyt, De Houwer, Everaert, & Hermans, 2012).

FSAA has been found to influence the degree to which participants process the evaluative meaning of task-irrelevant stimuli (e.g., Everaert, Spruyt, & De Houwer, 2013; Spruyt, De Houwer, & Hermans, 2009). As an example, consider the evaluative priming studies by Spruyt, De Houwer, Hermans, and Eelen (2007). In a typical evaluative priming experiment, participants are presented with a series of evaluative 'target' stimuli (e.g., a picture of a kitten), each of which is preceded by a brief presentation of an evaluative 'prime' stimulus (e.g., a picture of a corpse). Participants are usually asked to categorize the target stimuli according to their evaluative connotation (positive or negative) and to disregard the prime stimuli. Performance is typically better when the prime stimulus has the same evaluative connotation than when it has the opposite evaluative connotation of the target. This effect can come about only if participants process the evaluative connotation of the primes and can therefore be used as an index of automatic evaluation of the primes. Despite numerous studies attesting to the unconditional, automatic nature of this so-called 'evaluative priming effect' (Herring et al., 2013; Klauer & Musch, 2003), Spruyt and colleagues demonstrated that the occurrence of this effect is restricted to conditions that maximize

selective attention for the evaluative stimulus dimension. In their experiment, participants were asked to switch between two categorization tasks on a trial-to-trial basis: the (traditional) evaluative categorization task (i.e., positive vs. negative discriminations) and a non-evaluative semantic categorization task (i.e., animals vs. object discriminations). Crucially, one group of participants (i.e., the evaluative group) performed the evaluative categorization task on 75 % of the trials and the non-evaluative semantic categorization task on the remaining 25 % of the trials. Another group of participants (i.e., the non-evaluative group) performed the evaluative categorization task on just 25 % of the trials and the non-evaluative semantic categorization task on the remaining 75 % of the trials. As a result of the different base rates of the two categorization tasks, participants in the evaluative group were encouraged to maximize selective attention for the evaluative stimulus dimension whereas participants in the non-evaluative group were encouraged to maximize selective attention for non-evaluative stimulus information. In line with the hypothesis that automatic stimulus evaluation depends on FSAA, a reliable evaluative priming effect was found in the evaluative group only (for related findings, see Gast, Werner, Heitmann, Spruyt, & Rothermund, 2014; Spruyt, 2014; Spruyt & Tibboel, 2015).

Importantly, while this pattern of results may be sufficient to conclude that the evaluative priming effect can be dependent on FSAA (i.e., an observable phenomenon), it is insufficient to draw strong conclusions concerning the degree to which automatic stimulus evaluation is in general contingent upon FSAA (i.e., one of the processes underlying this phenomenon). For the evaluative priming effect to occur, two sequential steps need to occur. First, participants need to process the evaluative tone of the (task-irrelevant) prime stimulus to some extent. Second, the outcome of the prime-evaluation process needs to impact target performance by means of some mechanism. While such a mechanism can set in before the evaluative tone of a prime is processed completely, partial completion of the prime-evaluation

process is necessary before any second mechanism can be initiated (see Spruyt et al., 2011).

Accordingly, the observation that the evaluative priming effect is dependent upon FSAA may

in principle result from three different scenarios (see also Deutsch & Gawronski, 2009; Spruyt

et al., 2011): (a) FSAA may impact the very process of automatic stimulus evaluation, (b)

FSAA may impact task-specific processes that translate the outcome of the prime-evaluation

process into an observable effect, or (c) FSAA may impact both the automatic evaluation

process and task-specific processes that drive the evaluative priming effect. Accordingly, to

validate the hypothesis that FSAA exerts an influence on automatic stimulus evaluation (i.e.,

scenario a or c), one needs to demonstrate that the impact of FSAA on automatic stimulus

evaluation is not restricted to one specific measure of automatic stimulus evaluation. Instead,

if the impact of FSAA on automatic stimulus evaluation replicates across a wide range of

experimental tasks, each characterized by the operation of unique, task-specific processes, one

can safely conclude that the observed effects of FSAA must be driven by a genuine

modulation of the degree to which automatic stimulus evaluation occurs.[1]

    In line with this logic, Everaert et al. (2013) demonstrated that attentional capture by

emotional stimuli, as measured with the emotional Stroop task and the dot probe task, occurs

only if participants assign attention to the evaluative stimulus dimension. Similarly, Everaert,

Spruyt, Rossi, De Houwer, and Pourtois (2014) showed that EEG correlates of early orienting

responses towards rare, emotional stimuli were larger when the evaluative stimulus dimension

was selectively attended to as compared to when a non-emotional stimulus dimension was

selectively attended to. In sum, FSAA seems to modulate the very process of automatic

---

[1] There is currently an intense debate concerning the question whether effects in the standard evaluative priming paradigm can arise in the absence of overlap between the response set and the prime set (see Becker; Klauer; & Spruyt, in press; Klauer, Becker, & Spruyt, in press; Werner & Rothermund, 2013; Rothermund & Werner, 2014; Spruyt, 2014; Spruyt & Tibboel, 2015). Crucially, this discussion concerns the mechanism(s) that translate the outcome of the prime-evaluation process into an observable effect (e.g., response compatibility, encoding facilitation, etc.), not the prime-evaluation process itself. Given that these mechanisms are not operative in the AMP, we refer the reader to the above-cited papers for an extensive treatment of this topic.

stimulus evaluation rather than the mechanisms that are at play in specific experimental paradigms.

Nevertheless, this general conclusion is challenged by a series of recent studies in which the Affect Misattribution Procedure (AMP) was used as a measure of automatic stimulus evaluation. In a typical AMP study, participants are asked to judge whether a briefly presented, neutral Chinese ideograph is more pleasant or less pleasant than the average Chinese ideograph (Payne, Cheng, Govorun, & Stewart, 2005). Crucially, each ideograph is preceded by the short presentation of an evaluative prime stimulus. Results consistently show that the (neutral) ideographs are more likely to be evaluated as being pleasant rather than unpleasant when they are preceded by a positive prime stimulus as compared to when they are preceded by a negative prime stimulus. Although the AMP is quite similar to the evaluative priming paradigm in terms of its procedural features, several studies suggest that both paradigms rely on different task-specific processes (e.g., Deutsch & Gawronski, 2009; Deutsch, Kordts-Freudinger, Gawronski, & Strack, 2009; Gawronski, Cunningham, LeBel, & Deutsch, 2010; Gawronski & Ye, 2014; Gawronski & Ye, 2015). While the evaluative priming effect is thought to result from processes operating at the response interference and/or target encoding stage (e.g., De Houwer, Hermans, Rothermund, & Wentura, 2002; Klauer, Roßnagel, & Musch, 1997; Klinger, Burton, & Pitts, 2000; Spruyt et al., 2007), it is typically argued that AMP effects result from an (unintentional) misattribution of the feelings evoked by the primes to the targets (e.g., Payne et al., 2005; but see Bar-Anan & Nosek, 2012).

Recently, Gawronski and colleagues compared the evaluative priming task and the AMP with regard to their sensitivity to variations in FSAA (Gawronski et al., 2010; Gawronski & Ye, 2015). In both paradigms, the primes were face pictures that varied along two, orthogonal dimensions: race (black vs. white) and age (young vs. old). One group of participants (i.e., the race group) was asked to count how many prime pictures depicted black

or white faces and a second group of participants (i.e., the age group) was asked to count how many prime pictures depicted young or old faces. Participants in the race group were thus encouraged to assign selective attention to the race dimension while participants in the age group were encouraged to assign selective attention to the age dimension. When participants performed the evaluative priming task, Gawronski et al. (2010) observed a data pattern that was consistent with the results reported by Spruyt et al. (2007, 2009, 2012). Evaluative priming of race was found only when participants assigned attention to the race dimension while evaluative priming of age was found only when participants assigned attention to the age dimension. Moreover, indices of reliability and validity mirrored these observations, with priming effects showing greater reliability and validity for the dimension that was selectively attended to. In contrast, the priming scores of the AMP were found to be unaffected by the manipulation of FSAA. Both the race group and the age group showed similar AMP scores, reliability coefficients, and validity indices along both stimulus dimensions.

To account for these divergent findings, Gawronski et al. (2010) argued "that the impact of unattended category cues [is dependent] on conditions inherent in specific tasks" (p. 1008). That is, the divergent findings obtained with the AMP and the standard evaluative priming paradigm were ascribed to a differential impact of FSAA on task-specific mechanisms whereas automatic stimulus evaluation itself was assumed to be unaffected by FSAA (i.e., scenario b). Clearly, these results and their interpretation are in direct opposition to earlier studies showing an impact of FSAA on a wide variety of measures of automatic stimulus evaluation.

It must be noted, however, that the procedures used to manipulate FSAA in the studies by Spruyt and colleagues were quite different from those used by Gawronski and colleagues. Most notably, in the studies by Spruyt and colleagues, participants were always encouraged to assign selective attention either to the evaluative stimulus dimension (e.g., positive vs.

negative) or to a non-evaluative control dimension (e.g., animal vs. object). In contrast, in the studies by Gawronski and colleagues, participants were encouraged to assign attention to either race or age, two dimensions that are intrinsically related to valence (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; McConnel & Leibold, 2001; Rudman, Greenwald, Mellott, & Schwartz, 1999; Stangor, Lynch, Duan, & Glas, 1992; Taylor, Fiske, Etcoff, & Ruderman, 1978). Crucially, the AMP always induces selective attention for the evaluative stimulus dimension as it requires participants to make explicit evaluative judgments. It could be hypothesized that participants were encouraged to process both the age dimension and the race dimension because their inherent evaluative connotation was task-relevant. It thus remains an open question whether a more traditional manipulation of FSAA (i.e., selective attention or stimulus valence vs. selective attention for a neutral stimulus dimension) might impact automatic stimulus evaluation as measured by the AMP.

Accordingly, we ran two experiments in which the AMP was used to capture automatic stimulus evaluation under conditions that, consistent with our earlier studies, either promoted selective attention for the evaluative stimulus dimension or a non-evaluative stimulus dimension (e.g., Everaert et al., 2013; Spruyt, 2014; Spruyt et al., 2007, 2009, 2012). Similar to Spruyt et al. (2007), participants performed either the standard AMP or were asked to decide whether the target Chinese ideograph referred to an animal or an object (Experiment 1), or a human or animal (Experiment 2). One group (i.e., the evaluative group) performed the standard AMP on 75 % of all trials and the aforementioned non-evaluative semantic categorization task on the remaining 25 %. The other group (i.e., the non-evaluative group), performed the non-evaluative semantic categorization task on 75 % of all trials and the standard AMP on the remaining 25%. The high proportion of standard AMP trials in the evaluative group and the semantic categorization task in the non-evaluative group encouraged participants to assign to attention to the evaluative stimulus dimension or a non-evaluative

stimulus dimension, respectively. We therefore predicted the evaluative group to have stronger AMP effects compared to the non-evaluative group, both in terms of overall effect size and the reliability of the measure.

Experiment 1

Method

Participants. Participants where 69 students at Ghent University who were sampled using an online recruitment system (10 men, 59 women, M_{age} = 22.8 years, SD = 4.7 years). All participants signed an informed consent form prior to participation. The experiment lasted approximately 20 minutes and was followed by a short, unrelated experiment. Participants were paid €5 in exchange for their participation.

Materials. A set of 20 positive pictures ($M_{valence}$ = 7.97, $SD_{valence}$ = 0.43) and 20 negative pictures ($M_{valence}$ = 2.49, $SD_{valence}$ = 0.67) was selected from the IAPS database (Lang, Bradley, & Cuthbert, 1999) to be used as primes in the experiment (see the Appendix for the picture numbers). A set of 200 neutral, Chinese ideographs, taken from Payne et al. (2005), was used as a target set. The ideographs were presented in white on a black background. All pictures were resized to have a height and width of 256 pixels and where presented in the center of the screen in a frame that had a height and width of 300 pixels. The lines of the frame were 20 pixels wide. The color of the frame could either be white or green and served as a cue for the task that participants were required to perform.

All stimuli were presented against the black background of a 19-inch computer monitor with a 60-Hz refresh rate and a screen resolution of 1024 by 768 pixels. Stimulus presentation was controlled with a Dell Optiplex GX520 desktop computer running Affect 4.0 (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010). Responses were registered using a standard computer keyboard.

Procedure. Participants were assigned at random to the evaluative group ($n = 37$) or the non-evaluative group ($n = 32$) and were tested individually in a dimly-lit room, either individually or in pairs. The sample size of both groups was slightly different because the experimenter assigned two participants to the evaluative group instead of the non-evaluative group by accident. The instructions stated that the aim of the experiment was to investigate the interpretation of foreign symbols in the presence of distracting stimuli. Depending on the color of the frame surrounding the Chinese ideograph (see below), participants were asked to judge whether they thought the symbol was less or more pleasant than the average symbol (i.e., the evaluative decision task) or referred to an animal or an object (i.e., the semantic decision task). In line with Payne et al. (2005), the prime pictures were explicitly described as irrelevant to the task at hand.

Each of the 200 Chinese ideographs was presented once and each prime stimulus was presented 5 times. The total number of trials was thus 200 and the proportion of positive and negative primes was exactly 50 percent. Each trial started with the presentation of a fixation cross in the middle of the screen for 500 ms. Immediately after the offset of the fixation cross, a prime picture was presented for 5 refresh cycles of the computer screen (approximately 83 ms) in the center of a white frame. After a stimulus onset asynchrony of 200 ms, the target ideograph was presented for 100 ms in the middle of a frame that could be either white or green. The presentation of the ideograph was followed by the presentation of a backward mask that depicted a noisy pattern of black and white pixels.

The white frame surrounding the targets turned green on a set of 50 randomly selected trials (i.e., 25 %) containing an equal portion of positive and negative primes. Participants in the evaluative group were asked to perform the evaluative decision task when a white frame surrounded the target ideograph and to perform the semantic decision task when a green frame surrounded the target ideograph. Conversely, participants in the non-evaluative group

were asked to perform the semantic decision task when a white frame surrounded the target ideograph and to perform the evaluative decision task when a green frame surrounded the target ideograph. The participants in the evaluative group thus performed the evaluative decision task on 75 % of all trials while the participants in the non-evaluative group performed the evaluative decision task on just 25 % of all trials.

Participants used the "a" and "z" key, and the "o" and "p" key on the standard AZERTY keyboard with their left and right hands, respectively, to perform the two tasks. The mapping of the tasks onto different hands was counterbalanced across participants. Similar to the procedure employed by Gawronski et al. (2010), an error message was presented for 2 seconds after the response when participants responded after 1500 ms had elapsed ("!!!TOO SLOW!!!"), applied the wrong task (!!!WRONG TASK!!!), or both ("!!!WRONG TASK AND TOO SLOW!!!"). When the trial ended, an inter-trial interval was initiated that varied randomly between 500 ms and 1500 ms.

Results

Before proceeding with the analyses, data of 3 participants were excluded because their error rates exceeded the limit of 1.5 times the interquartile range above the third quartile of the distribution of the errors. As a result, the evaluative group consisted of 37 participants whereas the non-evaluative group consisted of 29 participants. Analyses were restricted to the standard (evaluative) AMP trials. Trials in which the wrong task was performed were excluded from the analyses (4.22 %). Trials with reaction times below 150 ms or above the response deadline of 1500 ms were removed from the dataset, as were trials in which the reaction time deviated more than 2.5 standard deviations from a participant's mean reaction time in a particular condition (3.19 %; Ratcliff, 1993).

Main analyses. A mixed effects approach was adopted for two reasons. First, the proportion of valid AMP trials was necessarily larger in the evaluative group (75 %) as compared to the

non-evaluative group (25 %). Second, as a result of the uneven distribution of AMP trials in both conditions, the number of AMP trials that required a task-switch was necessarily larger in the non-evaluative group as compared to the evaluative group. To take these between-group differences into account, we performed a logistic mixed effects analysis wherein the binary response was regressed on the fixed factors prime valence (positive vs. negative), switch type (repetition vs. switch), and group (evaluative vs. non-evaluative). In addition, participants and prime stimuli were treated as crossed random effects (Baayen, Davidson, & Bates, 2008; Jaeger, 2008; Judd, Westfall, & Kenny, 2012). Significance tests were performed using standard Chi-squared tests for mixed-effects models. It may be noted that the maximum-likelihood approach precluded the calculation of effect sizes. As an alternative, confidence intervals are reported where appropriate.

The mixed-effects model yielded several significant effects (see Table 1, for the mean predicted probabilities). A highly significant main effect of prime valence revealed that participants were more likely to emit a positive response after a positive prime than after a negative prime (i.e., the AMP effect), $\chi^2_1 = 78.24, p < .0001$. Crucially, a significant interaction between group and prime valence was observed, $\chi^2_1 = 10.37, p = .001$. This interaction showed that the AMP effect was larger in the evaluative group than in the non-evaluative group. The AMP effect, however, was significant in the evaluative group, $\chi^2_1 = 123.75, p < .0001$, as well as the non-evaluative group, $\chi^2_1 = 28.00, p < .0001$. Finally, a significant three-way interaction between group, switch type, and prime valence was observed, $\chi^2_1 = 4.36, p = .037$, showing that the interaction between group and prime valence was present on the repetition trials, $\chi^2_1 = 10.17, p < .001$, but not on the switch trials, $\chi^2_1 = 0.89, p = .34$.

Error analyses. Analysis of the proportion of errors (wrong task selection or slow responses) revealed main effects of group, $\chi^2_1 = 64.05, p < .0001$, and switch type, $\chi^2_1 = 41.02, p <$

.0001. Higher error probabilities were observed in the non-evaluative group, $\hat{\pi} =$ .074, 95% $CI = (.052, .104)$ compared to the evaluative group, $\hat{\pi} = .015, 95\%, CI = (.011, .020)$ and in switch trials, $\hat{\pi} = .038, 95\% \, CI = (.028, .051)$, compared to repetition trials, $\hat{\pi} = .015, 95\% \, CI = (.011, .020)$. The effect of prime valence did not reach significance nor did it interact significantly with other effects in the model, all $\chi_1^2 s < 2.40, ps > 0.12$.

Reliability analyses. A bootstrapping approach was adopted to examine the reliability of the AMP effect, both across and within groups. For each of 10,000 runs, the data of an individual participant were split in two random sets. To ensure that these sets were (roughly) of equal size in both groups, all AMP data were used for participants in the non-evaluative group (i.e., 50 trials) whereas a random subset of about 50 data points (i.e., 1/3 of all trials) was used for participants in the evaluative group. For each participant, the two sets were then used to calculate two AMP scores per participant, which were subsequently correlated and Spearman-Brown corrected. Finally, the reliability coefficient was estimated by taking the mean of all correlations obtained. The sampling distribution of the correlations was used to perform further non-parametric statistical tests. Overall, the reliability was high, $r = .76, p < .0002$, both in the evaluative group, $r = .74, p < .0002$, and the non-evaluative group, $r = .76, p < .0002$. The difference between the reliability coefficients observed both groups was statistically unreliable, $p = .89$.

Discussion

According to the FSAA framework developed by Spruyt et al. (2009; also see, Spruyt et al., 2007; Spruyt et al., 2012), automatic stimulus evaluation occurs only if and to the extent that attention is allocated to evaluative stimulus information. In line with this framework, we observed that automatic stimulus evaluation as measured by the AMP was more pronounced when participants were encouraged to selectively attend to evaluative stimulus information as

compared to when participants were encouraged to selectively attend to non-evaluative stimulus information.

Our analyses also revealed that the modulation of the AMP effect by FSAA was much more pronounced on repetition trials as compared to switch trials. This finding is important as it seems to rule out rivaling interpretations in terms of task-switching effects. Assume, for instance, that participants were always inclined to perform the task that was presented on the majority of the trials (i.e., the evaluative task in the evaluative condition and the non-evaluative task in the non-evaluative condition). If so, switching to the evaluative task must have been easier in the evaluative group as compared to the non-evaluative group, as task preparation has been found to make switching between tasks easier (e.g. Meiran, 2000, Monsell, 2003). One might thus expect the switch trials to produce more noisy data in the non-evaluative group as compared to the evaluative group, thereby potentially producing a reduced AMP effect in non-evaluative group as compared to the evaluative group. The present findings show, however, that the impact of FSAA on the AMP was more pronounced on the repetition trials as compared to the switch trials, not the other way around. We can thus safely rule-out the hypothesis that the observed impact of FSAA of the AMP effect was simply a by-product of task-switching effects.

Experiment 2

Whilst Experiment 1 revealed a clear-cut modulation of the AMP effect by FSAA, an additional experiment was performed, for a number of reasons.

First, in Experiment 1, the nature of the required classification task was signaled by means of a green frame that was presented on the exception trials only. If it is assumed that the green frame was perceived as a go-signal, it may have interfered with a quick and efficient switching from the dominant to the non-dominant task. Alternatively, one could simply argue that participants may have been confused by the occasional presentation of an additional

stimulus on a small subset of trials. In both scenario's, the likelihood of obtaining a clear-cut AMP effect in the non-evaluative condition would be reduced, thereby producing an artifactual moderation of the AMP effect by FSAA. Accordingly, in Experiment 2, a frame was presented on all trials. The neutral color of this frame (i.e., purple vs. yellow) signaled which task was required on a given trial and the assignment of the frame color to the tasks was counterbalanced across participants.

A second limitation of Experiment 1 concerns the nature of the prime set. As the primes varied on the evaluative stimulus dimension only, the results of Experiment 1 are insufficient to rule out the possibility that selectively attending towards the evaluative stimulus dimension increases misattribution effects in general rather than affect misattribution effects in particular. In this respect, it is important to point out that, according to the FSAA account, the impact of FSAA on the semantic analysis of task-irrelevant stimuli is by no means confined to the evaluative stimulus dimension. In principle, any stimulus dimension or stimulus feature can be prioritized as a function of concurrent goals and task-demands (see also Kiefer, 2010; Kiefer & Martens, 2012; Medin & Schaffer, 1978; Nosofsky, 1986). In line with this viewpoint, Spruyt et al. (2009) found significant (non-evaluative) semantic priming effects only when FSAA was directed towards a (non-evaluative) semantic stimulus dimension (e.g., animals vs. objects). Likewise, Everaert et al., (2013, Experiment 1) observed effects akin to the emotional Stroop effect for neutral semantic stimuli when participants were encouraged to attend selectively to a semantic stimulus dimension. However, both Spruyt et al. (2012) and Everaert et al., (2013, Experiment 2) failed to replicate this data pattern when, respectively, using a subliminal (sequential) priming procedure or a dot probe paradigm. It thus remains to be seen whether FSAA can impact non-evaluative misattribution effects.

Accordingly, in the second experiment, the primes varied along two, orthogonal stimulus dimensions: the evaluative stimulus dimension (i.e., positive vs. negative) and a non-

evaluative, semantic stimulus dimension (i.e., human vs. animal). Just like in Experiment 1, participants were asked to switch between an evaluative judgment task (i.e., the AMP) and a non-evaluative judgment task (hereafter referred to as the Semantic Misattribution Procedure, SMP). In this experiment, however, the nature of the non-evaluative judgment task (i.e., humans vs. animals) was related to the non-evaluative stimulus dimension of the primes that was manipulated. In line with prior work (e.g., Deutsch & Gawronski, 2009; Imhoff, Schmidt, Bernhardt, Dierksmeier, & Banse, 2011), we thus expected the response rates in the SMP to vary as a function of the primes. More specifically, we expected participants to select the response 'human' more frequently than the response 'animal' after the presentation of a person as a prime. Likewise, we expected participants to select the response 'animal' more frequently than the response 'human' after the presentation of an animal as a prime. In line with previous research, AMP effects were predicted to be larger in the evaluative group than in the non-evaluative group whereas SMP effects were predicted to be larger in the non-evaluative group than in the evaluative group. Although no differences with regard to the reliability of the AMP were found in Experiment 1, we also examined whether the reliability of the AMP and the SMP was dependent on FSAA.

Method

Participants. Participants were 90 students at Ghent University who were sampled using an online recruitment system (58 women, 32 men, $M_{age} = 21.9$ years, $SD = 3.0$ years). All participants gave informed consent prior to participation and were paid €5.

Materials. Both the evaluative stimulus dimension (positive vs. negative) and the semantic non-evaluative stimulus dimension (humans vs. animals) of the prime stimuli were manipulated by selecting 10 IAPS pictures for each of the following four stimulus categories (see the Appendix for the picture numbers): positive humans ($M_{valence} = 7.95$, $SD_{valence} = 0.34$), positive animals ($M_{valence} = 8.10$, $SD_{valence} = 0.40$), negative humans ($M_{valence} = 2.31$, $SD_{valence} = $

0.64), and negative animals ($M_{valence} = 2.53$, $SD_{valence} = 0.72$). The target set now included 260 Chinese ideographs, of which 20 were used in the practice phase and 240 were used during the actual experiment. The 240 ideographs were paired randomly with the 40 prime pictures, which were presented six times each. The experiment was run on laptop computers connected to an external keyboard, mouse, and a 19-inch LCD monitor (60 Hz refresh rate, screen resolution 1024×768 pixels). In all other aspects, the materials of Experiment 2 were identical to those used in Experiment 1.

Procedure. Participants were assigned at random to either the evaluative group ($n = 46$) or the non-evaluative group ($n = 44$). The experimental procedures used in Experiment 2 were virtually identical to those used in Experiment 1, with the following exceptions. First, the SMP required participants to indicate whether they thought the target symbols referred to humans or animals. Second, the color of the frame that signaled the task was purple for one task and yellow for the other task, with color assignment counterbalanced across participants. Third, to familiarize participants with the response assignments, we included a brief practice phase (20 trials) in which no primes were presented. Fourth, the actual experiment consisted of 240 trials (instead of 200). Each prime picture was presented exactly 6 times. Each of the 240 ideographs that had not been presented during the practice phase was presented exactly once. In the evaluative group, the AMP was performed on 180 trials (75 %) while the SMP was performed on 60 trials (25 %). Conversely, the non-evaluative group performed the AMP on 60 (25 %) trials and the SMP on 180 trials (75 %). Different trial types were divided randomly across two 120-trial blocks, separated by a short, self-paced break. Finally, at the end of the experiment, participants were asked to use two 5-point rating scales to indicate (a) the extent to which they intentionally judged the primes instead of the targets and (b) the extent to which they felt influenced by the presentation of the primes (see Bar-Anan & Nosek, 2012; Payne et al., 2013).

Results

      Due to an experimenter error, the data of 2 participants were unusable. Inspection of the error rates identified two participants whose error rates exceeded the limit of 1.5 times the interquartile range above the third quartile of the distribution of the errors. The final sample therefore consisted of 86 participants, 45 in the evaluative group and 41 in the non-evaluative group. Similar to Experiment 1, trials in which the wrong task was performed were excluded from the analyses (5.33 %). Trials with reaction times below 150 ms or above the response deadline of 1500 ms were removed from the dataset, as were trials in which the reaction time deviated more than 2.5 standard deviations from a participant's mean reaction time in a particular condition (2.64 %).

Main analyses. Separate mixed effects models were estimated for the AMP and the SMP (see Table 2 and Table 3, for the predicted means). For the AMP , the binary response was regressed on the fixed factors prime valence (positive vs. negative), switch type (repetition vs. switch), and group (evaluative vs. non-evaluative). For the SMP, the binary response was regressed on the fixed factors prime type (human vs. animal), switch type (repetition vs. switch), and group (evaluative vs. non-evaluative). In addition, similar to Experiment 1, both models included participant identity as a crossed random effect. Unlike Experiment 1, however, prime identity was not included as a crossed random effect as its associated variance was so small ($s^2 < 0.05$) that the model failed to converge for at least a subset of analyses.

      The analysis of the AMP data revealed a significant main effect of prime valence (i.e., the AMP effect), $\chi^2_1 = 168.61, p < .0001$, that was qualified by a significant interaction with the group factor, $\chi^2_1 = 23.29, p < .0001$. Replicating the results of Experiment 1, the AMP effect was larger in the evaluative group as compared to the non-evaluative group. Nevertheless, also similar to Experiment 1, the AMP effect reached significance both in the evaluative group, $\chi^2_1 = 343.16, p < .0001$, and the non-evaluative group, $\chi^2_1 = 21.69, p <$

.0001. In contrast with the previous experiment, the three-way interaction between group, switch type, and prime valence did not attain significance, $\chi_1^2 = 1.08, p = .28$. Finally, we also obtained a significant main effect of the group factor, $\chi_1^2 = 14.76, p < .001$. Participants were more likely to emit a 'positive' response than a 'negative' response in the evaluative group compared to the non-evaluative group. A significant interaction between the group factor and the factor switch type, $\chi_1^2 = 11.55, p < .001$, revealed that this group difference was more pronounced on switch trials than on repetition trials.

A similar analysis of the SMP yielded a significant main effect of prime type, $\chi_1^2 = 159.38, p < .0001$. The probability of the response 'human' was larger when the prime picture depicted a human than when the prime picture depicted an animal. As anticipated, this main effect of prime type was qualified by a significant interaction with the group factor, $\chi_1^2 = 6.54, p = .011$. To our surprise, however, the nature of this interaction effect was opposite to our predictions: the SMP effect was larger, not smaller, in the evaluative group, $\chi_1^2 = 103.34, p < .0001$, as compared to the non-evaluative group, $\chi_1^2 = 76.40, p < .0001$. The three-way interaction between group, switch type, and prime type did not attain significance, $\chi_1^2 = 1.93, p = .16$. Finally, participants tended to select the 'human' response more frequently in the non-evaluative group as compared to the evaluative group, $\chi_1^2 = 3.19, p = .074$. A significant interaction between the group factor and the factor switch type indicated that this group difference was larger on switch trials than on repetition trials, $\chi_1^2 = 4.34, p = .037$.

Error analysis. Additional analyses revealed that participants were more likely to execute the wrong task or to respond after the response deadline had already passed in the task that was less frequent, $\chi_1^2 = 79.44, p < .0001$, and $\chi_1^2 = 74.81, p < .0001$, for the AMP and SMP respectively. Likewise, more errors were produced on switch trials than on repetition trials,

both in the AMP and the SMP, $\chi_1^2 = 56.58, p < .0001$, and $\chi_1^2 = 71.25, p < .0001$, respectively.

Reliability analysis. Reliability estimates were computed in the same way as for Experiment 1. However, since more trials were available in the current experiment, two sets of 30 trials instead of 25 trials were randomly drawn for each simulation run. A high reliability coefficient was obtained for the AMP, $r = .73, p < .0002$. The reliability of the AMP was high in the evaluative group, $r = .80, p < .0002$, and moderate in the non-evaluative group, $r = .53, p = .0004$. The difference between both reliability coefficients attained significance, $p = .012$. For the SMP, a high reliability coefficient was obtained as well, $r = .78, p < .0002$. Contrary to our expectations, the reliability coefficient was high in the evaluative group, $r = .83, p < .0002$, and moderate in the non-evaluative group, $r = .69, p < .0002$. The difference between both coefficients was not statistically significant, $p = .08$.

Discussion

Replicating Experiment 1, the present experiment revealed a larger AMP effect when participants were encouraged to assign selective attention to evaluative stimulus information than when they were encouraged to assign selective attention to non-evaluative semantic stimulus information. Unlike Experiment 1, the present experiment also revealed that this modulation was equally strong on task-switch trials and repetition trials. The present experiment therefore adds further weight to the idea that the impact of our experimental manipulation on the AMP effect is not simply a by-product of task-switching effects. In addition, the present experiment confirms that the results obtained in Experiment 1 were not just a by-product of the fact that the (green) task cues were presented on exception trials only.

Nevertheless, other aspects of the present data are difficult to reconcile with the FSAA framework initially developed by Spruyt and colleagues. In addition to the evaluative stimulus

dimension, we also manipulated a non-evaluative, semantic stimulus dimension (i.e., human vs. animal) and participants were asked to judge the targets in terms of this non-evaluative stimulus dimension on either the majority (i.e., the non-evaluative group) or minority of trials (i.e., the evaluative group). Based on the FSAA framework, we had expected that target performance in the non-evaluative task would be affected by the non-evaluative stimulus dimension to a larger extent in the non-evaluative group as compared to the evaluative group. However, the exact opposite was true. If it is assumed that the pattern of results obtained with the SMP are an accurate reflection of the degree to which participants engaged in automatic non-evaluative stimulus processing, the current findings seem to suggest that the requirement to adopt an evaluative processing mindset may enhance automatic semantic processing in general. We will elaborate on this intriguing possibility in more detail in the next section.

General discussion

According to the FSAA framework developed by Spruyt and colleagues (Spruyt et al., 2007, 2009, 2012), the automatic semantic analysis of task-irrelevant stimuli is constrained by top-down attentional control (see also Kiefer & Martens, 2010). Whereas differences in terms of task-relevant stimulus dimensions are assumed to become highly salient, the automatic analysis of task-irrelevant stimulus dimensions is assumed to be reduced. In line with this framework, both in Experiment 1 and 2, we observed that the AMP effect was larger in participants who performed an evaluative categorization task on the majority of trials (i.e., the AMP) as compared to participants who performed a non-evaluative categorization task on the majority of trials (i.e., the SMP). Importantly, this effect is difficult to explain as a by-product of the requirement to switch between two tasks. First, in Experiment 2, the impact of the proportion of evaluative categorization trials on the AMP effect was found independently of whether participants completed a switch trial or a repetition trial. Second, in Experiment 1, it was observed that the proportion of evaluative categorization trials affected the AMP effect

on repetition trials only. While this observation shows that task-switch effects were at play in Experiment 1, the nature of this effect should have reduced (not enhanced) the impact of our experimental manipulation. Taken together then, the results of Experiment 1 and 2 seem to suggest, as hypothesized, that automatic stimulus evaluation as measured by the AMP is dependent upon the degree of attention assigned to the evaluative stimulus dimension. In line with this conclusion, we also observed that the reliability of the AMP effect was consistently high under conditions that promoted selective attention for the evaluative stimulus dimension (i.e., Experiment 1 and 2) and significantly smaller under conditions that promoted selective attention for a non-evaluative stimulus dimension (i.e., Experiment 2).

One might note, however, that the AMP effect reached significance in both groups whereas earlier studies conducted at our lab typically revealed null-findings under conditions that promoted selective attention for non-evaluative stimulus information (Everaert et al., 2011, 2013; Spruyt et al., 2007, 2009, 2012). This discrepancy can be readily explained, however, if one takes into account that the AMP, by definition, requires participants to assign selective attention to the evaluative stimulus dimension in order to comply with task-demands. In addition, we manipulated prime valence in a salient manner, which in itself might be sufficient to encourage participants to attend to evaluative stimulus information (Duscherer, Holender, & Molenaar, 2008; Everaert et al., 2011).

Other discrepancies between our findings and the literature, however, are more difficult to account for. First, while our findings show that the AMP effect depends upon FSAA, Gawronski and colleagues found no modulation of the AMP effect as a function of whether participants selectively assigned attention to the race or the age of persons depicted as primes (Gawronski et al., 2010; Gawronski & Ye, 2015). One way to account for these inconsistent findings concerns the nature of the procedures that were used to manipulate FSAA. In line with our earlier work (e.g., Everaert et al., 2011, 2013, 2014; Spruyt et al.,

2007, 2009, 2012), we pitted selective attention for the evaluative stimulus dimension against selective attention for non-evaluative stimulus information. The use of two separable stimulus dimensions is important because it minimizes the likelihood that selective attention for one stimulus dimension affects the degree to which selective attention is assigned to the other stimulus dimension. In contrast, in the studies by Gawronski and colleagues, participants were required to attend to one of two stimulus dimensions that were both correlated with valence (i.e., race and age). If it is assumed that the requirement to adopt an evaluative processing mindset is sufficient to trigger selective attention for any stimulus dimension that is correlated with stimulus valence, the AMP findings reported by Gawronski and colleagues could, in principle, be accounted for. The same reasoning fails to deal, however, with the observation that Gawronski and colleagues did observe an impact of selective attention assignment in the standard evaluative priming paradigm. This discrepancy might be resolved, however, if one takes into account that participants who complete an AMP are sometimes inclined to judge the primes rather than the targets (e.g. Bar-Anan & Nosek, 2012; Eder & Deutsch, 2015). Such a process would interfere with the FSAA manipulation and might thus explain why FSAA manipulations have been less effective in the AMP as compared to the evaluative priming task. Further research would be required, though, to account for the full pattern of results obtained with the AMP and the evaluative priming paradigm.

A second observation that is relatively difficult to account for concerns the results obtained with the SMP trials in Experiment 2. In theory, the attentional weighting of task-relevant stimulus information should be a general phenomenon (e.g. Kiefer, 2012; Kiefer & Martens, 2010). Accordingly, although some attempts to demonstrate FSAA effects for non-evaluative stimulus dimensions produced null-findings in the past (Everaert et al., 2013; Spruyt et al., 2012), we had predicted that the SMP effect would be more pronounced in participants who were encouraged to assign selective attention to the non-evaluative stimulus

dimension as compared to participants who were encouraged to assign selective attention to the evaluative stimulus dimension. The results showed, however, that the exact opposite was true. Likewise, it was observed that the reliability of the SMP effect was significantly smaller, not larger, in the non-evaluative group as compared to the evaluative group. Clearly, if it is assumed that that this data pattern is an accurate reflection of the degree to which participants engaged in automatic non-evaluative stimulus processing (see above), the FSAA framework has difficulty accounting for these observations as it predicts a facilitation of task-congruent stimulus processing only. It is also inconsistent with earlier work from our research team showing enhanced processing of non-evaluative stimulus dimensions that were task-relevant as compared to non-evaluative stimulus dimensions that were task-irrelevant (e.g., Everaert et al., 2013, Experiment 1; Spruyt et al., 2009).

One way to account for the results obtained with the SMP relates to the nature of the selected stimulus materials. While care was taken to manipulate the evaluative stimulus dimension and a non-evaluative stimulus dimension in an orthogonal manner, one may argue that our choice to pit the category of humans against the category of animals was perhaps not the best decision. For instance, in the literature on animal welfare, it has been argued that the human/animal distinction might constitute an evaluative dimension, with animals becoming increasingly more negative as the degree of similarity with humans decreases (Batt, 2009; Plous, 1993). Moreover, there is ample evidence in the social cognition literature that the human/animal distinction is often used in stereotyping, with members of minority groups being characterized as less human and more animals-like than members of majority groups (e.g., Haslam, Bain, Douge, Lee, & Bastian, 2005; Viki, Winchester, Titshall, & Chisango, 2006). It could thus be argued that the manipulation of semantic category membership (i.e., humans vs. animals) was confounded with variations in evaluative stimulus information. Accordingly, one would predict the SMP effect in the present experiment to increase as the

broader experimental context promotes the use of evaluative stimulus information, as was observed.

Alternatively, one may speculate that the requirement to focus selective attention on the evaluative stimulus dimension somehow led to a general increase of semantic stimulus processing. While this possibility was never envisaged in the initial papers by Spruyt and colleagues, it may be noted that a similar data pattern was reported by Becker et al. (2016). In a recent replication of one of the original studies by Spruyt et al.'s (2009), these authors re-examined the extent to which the semantic priming effects in the pronunciation task are dependent upon FSAA. Although none of the reported effects reached statistical significance, results showed an overall tendency for semantic priming effects (i.e., evaluative and non-evaluative) to emerge under conditions that promoted selective attention for the evaluative stimulus dimension. Taken together then, it may be an inspiring exercise for future research to scrutinize the extent to and the conditions under which selective attention for the valence dimension may impact selective attention assignment to other stimulus dimensions.

Irrespective of the unanticipated pattern of results obtained with the SMP, the present findings already have important implications. First, given that the AMP data revealed clear-cut FSAA effects in the anticipated direction, the present findings add further weight to the hypothesis that automatic evaluation of task-irrelevant stimuli depends on FSAA (Spruyt et al., 2009). Second, given that the AMP effect is much more volatile than previously assumed, researchers are advised to take FSAA in account when using the AMP as a means to capture inter-individual differences in automatic stimulus evaluation. Not only can one expect FSAA to reduce or inflate differences between individuals, it may also influence the predictive validity of the AMP. More specifically, the degree to which a relationship can be found between AMP scores (registered at a *time 1*) and some outcome behavior (registered at a *time*

2) is likely to depend on the degree of overlap between the FSAA conditions realized at both test occasions.

      In conclusion, the present results demonstrate that automatic stimulus evaluation as measured by the AMP is less impervious to variations of FSAA than previously assumed. In addition, our findings suggest that selective attention for the evaluative stimulus dimension can enhance both evaluative and non-evaluative misattribution. Further research would be needed, however, to firmly substantiate and document this latter phenomenon.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Bar-Anan, Y., & Nosek, B. A. (2012). Reporting intentional rating of the primes predicts priming effects in the affective misattribution procedure. *Personality and Social Psychology Bulletin, 38*, 1194-1208.

Batt, S. (2009). Human attitudes towards animals in relation to species similarity to humans: a multivariate approach. *Bioscience Horizons, 2,* 180-190.

Becker, M., Klauer, K. C., & Spruyt, A. (in press). Is Attention Enough? A Re-Examination of the Impact of Feature-Specific Attention Allocation on Semantic Priming Effects in the Pronunciation Task. *Attention, Perception, and Psychophysics*.

Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on "automatic evaluations" as a function of task characteristics of the measure. *Journal of Experimental Social Psychology, 45*, 101-114.

Deutsch, R., Kordts-Freudinger, R., Gawronski, B., & Strack, F. (2009). Fast and fragile: A new look at the automaticity of negation processing. *Experimental Psychology, 56,* 434-446.

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition and Emotion, 16*, 643-666.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510-540.

Duscherer, K., Holender, D., & Molenaar, E. (2008). Revisiting the affective Simon effect. *Cognition and Emotion, 22,* 193-217.

Eder, A. B., & Deutsch, R. (2015). Watch the target! Effects in the affective misattribution procedure become weaker (but not eliminated) when participants are motivated to provide accurate responses to the target. *Frontiers in Psychology, 6.*

Everaert, T., Spruyt, A., & De Houwer, J. (2011). On the (un)conditionality of automatic attitude activation: The valence proportion effect. *Canadian Journal of Experimental Psychology, 65*, 125-132.

Everaert, T., Spruyt, A., & De Houwer, J. (2013). On the malleability of automatic attentional biases: Effects of feature-specific attention allocation. *Cognition and Emotion, 27*, 385-400.

Everaert, T., Spruyt, A., Rossi, V., Pourtois, G., & De Houwer, J. (2014). Feature-specific attention allocation overrules the orienting response to emotional stimuli. *Social Cognitive and Affective Neuroscience, 9*, 1352-1359.

Gawronski, B., Cunningham, W.A., Lebel, E.P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorization influence spontaneous evaluations of multiply categorisable objects? *Cognition and Emotion, 24,* 1008-1025.

Gawronski, B., & Ye, Y. (2014). What drives priming effects in the affect misattribution procedure? *Personality and Social Psychology Bulletin, 40*, 3-15.

Gawronski, B., & Ye, Y. (2015). Prevention of intention invention in the affect misattribution procedure. *Social Psychology and Personality Science, 6,* 101-108.

Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology, 89,* 937-950.

Herring, D.R., White, K.R., Jabeen, L.N., Hinojos, M., Terrazas, G., Reyes, S.M.,…Crites, S.L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin, 139,* 1062-1089.

Imhoff, R., Schmidt, A. F., Bernhardt, J., Dierksmeier, A., & Banse, R. (2011). An inkblot for sexual preference: A semantic variant of the Affect Misattribution Procedure. *Cognition and Emotion, 25*, 676-690.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434-446.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*, 54-69.

Kiefer, M. (2012). Executive control over unconscious cognition: attentional sensitization of unconscious information processing. *Frontiers in Human Neuroscience, 6,* 1-12.

Kiefer, M., & Martens, U. (2010). Attentional sensitization of unconscious cognition: Task sets modulate subsequent masked semantic priming. *Journal of Experimental Psychology: General, 139,* 464-489.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7-49). Mahwah, NJ: Erlbaum.

Klauer, K. C., Roßnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 23*, 246.

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 441-455.

Lang,P. J, Bradley, M. M., & Cuthberth, B. N. (1999). *International Affective Picture System (IAPS): Technical Manual and Affective ratings.* Gainesville, Fl. The Center for Research in Psychophysiology, University of Florida.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435-442.

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207-238.

Meiran, N. (2000). Reconfiguration of stimulus task sets and response task sets during task switching. *Control of cognitive processes: Attention and Performance XVIII,* 377-399.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134-140.

Nosofsky, R.M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39-57.

Payne, B. K., Brown-Iannuzzi, J., Burkley, M., Arbuckle, N. L., Cooley, E., Cameron, C. D., & Lundberg, K. B. (2013). Intention invention and the affect misattribution procedure reply to bar-anan and nosek (2012). *Personality and Social Psychology Bulletin, 39*, 375-386.

Payne, B.K., Cheng, C.M., Govorun, O., & Stewart, B.D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89,* 277-293.

Plous, S. (1993). Psychological mechanisms in the human use of animals. *Journal of Social Issues, 49,* 11-52.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510-532.

Rothermund, K., & Werner, B. (2014). Affective priming in the valent/neutral categorisation task is due to affective matching, not encoding facilitation: Reply to Spruyt. *Cognition and Emotion, 28,* 570-576.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437-465.

Spruyt, A. (2014). Attention please: Evaluative priming effects in a valent/non-valent categorization task (Reply to Werner and Rothermund, 2013). *Cognition and Emotion, 28*, 560-569.

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0 A Free Software Package for Implementing Psychological and Psychophysiological Experiments. *Experimental Psychology, 57*, 36-45.

Spruyt, A., Gast, A., & Moors, A. (2011). The sequential priming paradigm: A primer. In Klauer, K. C., Stahl, C., & Voss, A. (Eds.) *Cognitive methods in social psychology.* New York: Guilford.

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition, 122*, 91-95.

Spruyt, A., De Houwer, J., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language, 61,* 37-54.

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology, 54,* 44-53.

Spruyt, A., & Tibboel, H. (2015). On the Automaticity of the Evaluative Priming Effect in the Valent/Non-Valent Categorization Task. *PloS one, 10,* e0121564.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207-218.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and

    contextual bases of person memory and stereotyping. *Journal of Personality and*

    Social Psychology, 36, 778-793.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance.

    Psychological Review, 117, 440-463.

Viki, G.T., Winchester, L., Titshall, L., & Chisango, T. (2006). Beyond secondary emotions:

    The infrahumanization of outgroups using human-related and animal-related words.

    Social Cognition, 24, 753-775.

Werner, B., & Rothermund, K. (2013). Attention please: No affective priming effects in a

    valent/neutral-categorisation task. *Cognition and Emotion, 27,* 119-132.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: the role of

    stimulus-attribute and response-set selection. *Journal of Experimental Psychology:*

    Human Perception and Performance, *29,* 455-469.

Author note

Tom Everaert, Adriaan Spruyt, and Jan De Houwer, Department of Psychology, Ghent University, Ghent, Belgium.

Correspondence concerning this article, and requests regarding the experiment, stimulus materials, and the data should be addressed to Tom Everaert, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. Electronic mail can be sent to Tom.Everaert@UGent.be.

Table 1

Mean Predicted Probabilities ($\hat{\pi}$) of Positive Responses and 95 % confidence intervals as a function of Prime Valence (Positive vs. Negative), Switch type (Repetition vs. Switch), and Group (Evaluative vs. Non-evaluative) in Experiment 1.

| | | Prime Valence | | | | AMP effect | |
| | | Negative | | Positive | | | |
Switch Type	Group	$\hat{\pi}$	95 % CI	$\hat{\pi}$	95 % CI	M	95 % CI
Repetition	Evaluative	0.32	(0.27, 0.37)	0.72	(0.66, 0.76)	0.40	(0.37, 0.43)
	Non-Evaluative	0.38	(0.29, 0.48)	0.59	(0.49, 0.69)	0.22	(0.10, 0.34)
Switch	Evaluative	0.36	(0.30, 0.43)	0.66	(0.60, 0.72)	0.30	(0.25, 0.36)
	Non-Evaluative	0.36	(0.29, 0.44)	0.61	(0.54, 0.69)	0.27	(0.20, 0.33)
Total	Evaluative	0.33	(0.28, 0.39)	0.70	(0.65, 0.75)	0.37	(0.34, 0.40)
	Non-Evaluative	0.38	(0.30, 0.46)	0.60	(0.52, 0.68)	0.23	(0.15, 0.31)

Note. AMPs and corresponding 95% CIs were calculated using bootstrapping with 5000 resampling runs.

Table 2

Mean Predicted Probabilities (π̂) of Positive Responses and 95 % confidence intervals in the AMP as a function of Prime Valence (Negative vs. Positive), Switch type (Repetition vs. Switch), and Group (Evaluative vs. Non-evaluative) in Experiment 2.

| | | Prime Valence | | | | | |
| | | Negative | | Positive | | AMP | |
Switch Type	Group	$\hat{\pi}$	95 % CI	$\hat{\pi}$	95 % CI	M	95 % CI
Repetition	Evaluative	0.40	(0.36, 0.44)	0.67	(0.63, 0.71)	0.27	(0.24, 0.30)
	Non-Evaluative	0.42	(0.35, 0.49)	0.51	(0.44, 0.59)	0.10	(0.01, 0.19)
Switch	Evaluative	0.48	(0.42, 0.53)	0.72	(0.67, 0.76)	0.24	(0.20, 0.29)
	Non-Evaluative	0.36	(0.30, 0.41)	0.50	(0.45, 0.56)	0.15	(0.10, 0.20)
Total	Evaluative	0.43	(0.39, 0.47)	0.69	(0.65, 0.72)	0.26	(0.24, 0.28)
	Non-Evaluative	0.40	(0.34, 0.46)	0.51	(0.45, 0.57)	0.12	(0.05, 0.18)

Note. AMPs and corresponding 95% CIs were calculated using bootstrapping with 5000 resampling runs.

Table 3

Mean Predicted Probabilities (π̂) of Human Responses and 95 % confidence intervals in the SMP as a function of Prime Type (Animal vs. Human), Switch type (Repetition vs. Switch), and Group (Evaluative vs. Non-evaluative) in Experiment 2.

| | | Prime Type | | | | | |
| | | Animal | | Human | | SMP | |
Switch Type	Group	$\hat{\pi}$	95 % CI	$\hat{\pi}$	95 % CI	M	95 % CI
Repetition	Evaluative	0.40	(0.34, 0.47)	0.60	(0.54, 0.67)	0.20	(0.11, 0.29)
	Non-Evaluative	0.44	(0.40, 0.48)	0.61	(0.57, 0.65)	0.17	(0.14, 0.20)
Switch	Evaluative	0.37	(0.32, 0.41)	0.60	(0.55, 0.65)	0.24	(0.19, 0.29)
	Non-Evaluative	0.50	(0.45, 0.55)	0.63	(0.58, 0.67)	0.12	(0.07, 0.17)
Total	Evaluative	0.39	(0.34, 0.44)	0.60	(0.55, 0.65)	0.21	(0.16, 0.27)
	Non-Evaluative	0.46	(0.42, 0.50)	0.61	(0.58, 0.65)	0.15	(0.13, 0.18)

Note. SMPs and corresponding 95% CIs were calculated using bootstrapping with 5000 resampling runs.

Appendix

IAPS picture numbers in Experiment 1

Positive: 1440, 1460, 1463, 1500, 1604, 1610, 1620, 1710, 1750, 1920, 2057, 2165, 2311, 2345, 2550, 2655, 4641, 5831, 7325, 7330.

Negative: 1200, 1220, 1300, 1930, 2120, 2692, 2750, 2800, 2900, 3230, 6570, 8230, 9000, 9040, 9041, 9280, 9340, 9440, 9530, 9570.

IAPS picture numbers in Experiment 2

Positive humans: 2070, 2304, 2311, 2340, 2341, 2530, 2550, 7325, 8461, 8496.

Positive animals: 1440, 1460, 1463, 1500, 1600, 1604, 1620, 1710, 1750, 1920.

Negative humans: 2120, 2750, 2800, 2900, 3022, 3230, 6250, 8230, 9040, 9041.

Negative animals: 1052, 1111, 1220, 1274, 1280, 1300, 9181, 9561, 9570, 9571.