American Psychological Association
Emotion

1528-3542/20/$12.00
2020, Vol. 20, No. 2, 217–235
http://dx.doi.org/10.1037/emo0000538

Attending to Emotional Faces in the Flanker Task: Probably Much Less Automatic Than Previously Assumed

Swantje Tannert and Klaus Rothermund
Friedrich Schiller University Jena

Attention to emotional faces was tested in a series of 5 experiments using the flanker paradigm. Distraction and compatibility effects that were stronger for emotional compared to neutral faces were found in only one of the studies. No reliable differences were found between faces displaying different emotions. The data suggest that attentional capture of emotional faces depends on emotion being a task relevant feature, indicating that attention has to be intentionally allocated to emotional information for those effects to materialize. Our findings also indicate that attending to emotions due to task requirements is not a sufficient condition for an attentional bias towards emotional faces. Even within emotion classification tasks, we only found reliable attentional prioritizing of emotional faces when the position of the target stimulus varied across trials and had to be identified on the basis of an additional feature, thus rendering the processing of the flanker stimuli obligatory in the task. In sum, these findings indicate that automatic attentional capture by emotional faces is a highly conditional phenomenon.

Keywords: facial expressions of emotion, attentional capture, flanker task, flanker compatibility effect, flanker distraction effect

Does emotional information in faces automatically capture attention? For mostly theoretical reasons, this was claimed by several researchers. A general processing advantage for emotional faces was postulated based on the assumption that emotion is a highly relevant signal in social communication and interaction (Frischen, Eastwood, & Smilek, 2008; Williams, Moss, Bradshaw, & Mattingley, 2005). In addition, it was also claimed that facial expressions showing emotions relating to threat (anger or fear) are particularly relevant due to their function as danger signals (Eastwood, Smilek, & Merikle, 2001; Fox, Lester, Russo, Bowles, Pichler, & Dutton, 2000; Hansen & Hansen, 1988; Öhman, Lundqvist, & Esteves, 2001; Öhman & Mineka, 2001; Schupp, Öhman, Junghöfer, Weike, Stockburger, & Hamm, 2004; Tipples, Atkinson, & Young, 2002): "Because facial threat provides a warning that aversive consequences are likely, the evolved [fear] module should be biased for orienting attention to salient facial gestures that convey threat." (Öhman et al., 2001, p. 381).

Convincing empirical evidence for these claims, however, is scarce. Some researchers found effects of attentional capture of (negative) emotional faces in socially anxious people or other clinical samples (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van Ijzendoorn, 2007; Liu, Qian, Zhou, & Wang, 2006; Rapee & Heimberg, 1997). Because we are not concerned with psychopathology in this article but rather want to investigate processes of emotional processing in the general population, we will not deal with these findings from clinical samples in further detail. In recent studies with nonclinical/nonanxious samples, only little evidence can be found for an automatic attention allocation to emotional faces. Some neuropsychological findings show a preferred or especially deep processing of threatening stimuli (Morris, Öhman, & Dolan, 1998; Schupp et al., 2004). The vast majority of researchers using behavioral measures, however, finds either no attentional capture for facial expressions of emotions or finds those effects only in specific contexts or samples (Bar-Haim et al., 2007; Yiend, 2010; see our brief review of the literature on assessing capture for emotional faces with different paradigms below).

Recently, the assumption of an unconditionally automatic allocation of attention to emotional faces has also been criticized on theoretical grounds. Based on their feature specific attention allocation model, Spruyt and colleagues (Spruyt, De Houwer, Hermans, & Eelen, 2007) claimed that an automatic processing of affective information for irrelevant stimuli (i.e., primes or distractors) depends on an intentional processing of valence that is instructed in the task. In support of this claim, they found that affective priming effects only emerged when attention was directed to emotion in the main task (Spruyt et al., 2007).

Taken together, there is reason to doubt the unconditionality of attentional capture by emotional faces. However, the literature on this issue to date is far from being conclusive. Opinions diverge, and the field is characterized by scattered findings that were gathered with different paradigms, using different procedures, parameters, and materials that are hardly comparable. In general, there is a lack of approaches to pinpoint the phenomenon systematically and with sufficient power across different variants of a paradigm. In the following we will give a brief overview of the

This article was published Online First December 13, 2018.
Swantje Tannert and Klaus Rothermund, Institute of Psychology, Friedrich Schiller University Jena.
This research was supported by a grant from the German Research Foundation (LA 3275/1-1) to Oliver Langner and Klaus Rothermund.
Correspondence concerning this article should be addressed to Swantje Tannert, Institute of Psychology, Friedrich Schiller University Jena, Am Steiger 3, Haus 1, D-07743 Jena, Germany. E-mail: swantje.tannert@uni-jena.de

major paradigms that were used to study attentional capture by emotional faces, and of the findings that were reported in the literature. The most widely used paradigms for this purpose are (a) the search-task (Treisman & Souther, 1985), (b) the dot probe paradigm (MacLeod, Mathews, & Tata, 1986), and (c) the flanker-task (Eriksen & Eriksen, 1974).

The Visual Search Task

In the search task (Treisman & Souther, 1985), participants have to detect the presence or absence of a target within a crowd of distractors. Typically RTs grow linearly with increasing number of distractors when search is inefficient (Treisman & Gelade, 1980). Different studies with the search task provide evidence for attentional capture of emotional compared to neutral faces in general (Eastwood et al., 2001), as well as for threatening faces in particular (Fox et al., 2000; Hansen & Hansen, 1988; Horstmann & Bauland, 2006; Huang, Chang, & Chen, 2011; Öhman et al., 2001), which has been assumed to reflect an evolutionary important danger detection mechanism. Nevertheless many researchers argue that low-level stimulus characteristics can account for the findings in this task, which undermines an interpretation of the findings in terms of emotional attentional capture (Coelho et al., 2010; Harms & Bundesen, 1983; Horstmann & Bauland, 2006; Miller, 1991; Purcell & Stewart, 2010; Purcell, Stewart, & Skov, 1996).[1] Furthermore, there are only few studies comparing the attentional capture of emotional versus neutral stimuli. Because most studies compared angry versus happy faces, these studies thus do not allow any strong conclusions regarding general processing advantages (i.e., attentional engagement or difficulty of disengagement) for emotional compared to neutral faces.

The Dot Probe Paradigm

Compared to the search task, the dot-probe paradigm (MacLeod et al., 1986) has many advantages for studying automatic attention allocation to emotional faces. Typically, two horizontally aligned cue-faces, one of which is emotional, are presented briefly on the screen. Immediately following the presentation of the cues, a target appears at the former location of one of the cue faces. The attentional capture is measured by the difference of RTs between invalid trials, where the target appears at the position of the neutral face, and valid trials, where the target appears at the position of the emotional face (validity effect). Consequently, emotion is fully task-irrelevant in this paradigm and in contrast to the search task, because a direct competition between two cue stimuli is generated, it is possible to compute a validity effect that directly compares attention allocation to emotional versus neutral stimuli. Nevertheless, attentional capture for emotional faces in the dot probe task has proven very difficult to replicate (Bar-Haim et al., 2007; Puls & Rothermund, 2018) and some researchers question the reliability of this task (Schmukle, 2005).

The Flanker Paradigm

Another often used task to investigate automatic attention allocation to task-irrelevant stimuli is the flanker paradigm (Eriksen & Eriksen, 1974). In its original form, the paradigm contained letters but was later used to study attentional capture for emotional stimuli (e.g., Fenske & Eastwood, 2003). In the classical flanker task, participants are simultaneously presented with a central target and flanking stimuli (flankers) on both sides. Although they are supposed to classify the target, they are told to ignore the flankers.

In contrast to cuing tasks like the dot probe paradigm (see above), the flanker task typically does not compare flanker effects between two different flankers within the same trial (but see Moeller & Frings, 2014). Instead, comparisons regarding the strength of attentional effects between different types of flankers are made by comparing flanker effects that obtained in different trial types. Nevertheless, because target and flanker stimuli are presented simultaneously in each trial, the flanker paradigm also creates a situation in which the flanker and target stimuli compete for attention, thus allowing for an analysis and comparison of attentional biases toward different types of flanker stimuli.

Importantly, there are fundamentally different ways of computing emotion-related effects in the flanker paradigm, which it is important to distinguish. Because the core aim of our study is to investigate attention allocation for emotional facial expressions with different variants of the flanker paradigm, we will devote some space to describe the different ways how indicators of

[1] Low level perceptual features can also bias effects in the flanker paradigm (e.g., Horstmann et al., 2006; we discuss this topic in more detail below). It should be noted, however, that such a biasing influence of low-level features is much stronger and more fatal in the search task than in the flanker paradigm: In the search paradigm, participants have to search for dissimilarities between stimuli in order to distinguish the target from the surrounding distractors, and to decide whether a target is present or absent. Target detection in the search task typically is achieved by detecting a difference between the stimuli that are shown in the display (similarly, the absence of a target is evident if all stimuli are similar or equal). Search efficiency is thus a direct (inverse) function of the similarity between targets and distractors (Duncan & Humphreys, 1989). If emotional and neutral stimuli differ with regard to low-level perceptual features (e.g., spots of darkness), this feature is directly relevant for the task because it can be used to decide whether a target is present or not and will thus have a huge influence on target-distractor similarity. This in turn will produce strong asymmetries in search efficiency between emotional and neutral stimuli that are completely unrelated to their emotional meaning. This problem is not present in the flanker paradigm, because there the task is to categorize the target, rather than to detect its presence. The similarity/dissimilarity between target and flanker stimuli thus does not allow for an identification of the required response. Thus, although low level features may affect the perceptual salience of emotional and neutral flanker stimuli, and may produce differences in attentional capture, these effects are much weaker than in the search task, because they cannot be used to infer the correct response and thus influence responding only indirectly. Another important caveat that has to be kept in mind is that most of the currently used variants of the search task nearly always require processing of the emotional content of the faces, because this is what defines and separates targets from distractors. These versions of the search task thus do not allow for an investigation of fully unconditional attentional capture for emotion, because emotion is task relevant. ITo overcome this difficulty, the search task has to introduce another dimension on which faces differ and that defines the target detection task (e.g., person identity, gender, age), with emotion varying orthogonally to this dimension. Variants of the additional singleton paradigm (Theeuwes, 1992) might help to investigate automatic attentional capture for emotional facial expressions in the search task without making emotion task relevant. Several studies used this variant of the paradigm to investigate automatic attentional capture for various types of valent or relevant compared to neutral stimuli (e.g., Hickey, Chelazzi, & Theeuwes, 2010, 2011; Müller, Rothermund, & Wentura, 2016; Wentura, Müller, & Rothermund, 2014), but to our knowledge the additional singleton paradigm has not yet been applied to investigate these effects for facial expressions of emotion.

emotion-specific attention allocation can be computed within this paradigm, and we will discuss the rationale and interpretation as well as the pros and cons of each of these methods in some detail. We will also review previous findings of flanker studies that used these indicators to measure attention allocation to emotional faces in the flanker paradigm in this section.

Compatibility Effects

A standard effect that is typically computed in a flanker design is the response compatibility effect that measures the extent to which the presence of a flanker activates a corresponding response that either facilitates or conflicts with the target response. Subtracting reaction times (RTs; or error rates) in the compatible condition from those in the incompatible condition results in the standard overall flanker compatibility effect. As a measure of emotional attentional biases, response compatibility effects for emotional versus neutral (or positive vs. negative) stimuli have to be compared with regard to their strength. Although such a comparison of compatibility effects for different types of stimuli seems to be a simple and straightforward idea, it is in fact not easy to decide how exactly this should be done. Most studies using the flanker task with emotional faces as stimuli in fact compared compatibility effects that were computed for different types of targets (Barratt & Bundesen, 2012, Exp. 1; Fenske & Eastwood, 2003; Grose-Fifer, Rodrigues, Hoover, & Zottoli, 2013; Horstmann, Borgstedt, & Heumann, 2006). The result that emerged consistently across these studies is that compatibility effects were larger for positive than for negative targets. This result is typically taken as evidence that negative target faces attract more attention than positive faces and thus are less influenced by compatible and incompatible flankers. Although this may be a legitimate way to analyze and interpret the findings, it has to be noted that this is not what one originally wanted to investigate with the flanker paradigm, for several reasons: First and foremost, finding unequal compatibility effects for different targets is mute with regard to attention allocation to different types of flanker stimuli. The core idea of the flanker paradigm, which is to study attention allocation to stimuli that are task-irrelevant (i.e., to the flankers), is lost when compatibility effects are compared for different targets. Second, differences in the processing of positive and negative targets may have a strategic basis (e.g., participants may decide to put more weight on the detection and correct classification of negative compared to positive targets), and thus may not reflect processes of truly automatic (in the sense of unintentional and purely stimulus-driven; Bargh & Gollwitzer, 1994; Moors & De Houwer, 2006) attention allocation. Finally, target-related compatibility effects provide an (inverse) index of the amount of cognitive processing resources that are bound by a certain type of target reflecting familiarity or ease of processing rather than attention. Relatedly, target-related compatibility effects are not a sensitive indicator for processes of attentional capture because the target position is typically fixed across trials, which is why attention is directed to the target position by default regardless of the type of target.

An alternative way to compare compatibility effects in this paradigm is to compare compatibility effects for different types of flanker stimuli. At first sight, this way of analyzing the data comes much closer to the basic rationale of the flanker paradigm, which is to investigate automatic and unintentional attentional capture and processing of the irrelevant flanker stimuli. Such a strategy of analyzing the data has been followed in a study by Chen, Yao, Qian, and Lin (2016), who reported stronger interference for negative compared to neutral and positive face flanker stimuli in a sample of socially anxious individuals, whereas no differences between the three types of flanker-related compatibility effects were obtained for a control group of nonanxious individuals (for the sake of completeness it should be mentioned that the difference in findings between the two groups was not statistically reliable either). Although the authors of the other previously mentioned flanker studies did not report statistical tests for flanker-related compatibility effects in their studies, one can at least visually inspect the data in this regard. The pattern of findings is maximally heterogeneous, with two studies suggesting (descriptively) stronger compatibility effects for positive than for negative flankers (Fenske & Eastwood, 2003; Grose-Fifer et al., 2013), another study suggesting stronger compatibility effects for negative than for positive flankers (Horstmann et al., 2006), whereas for the third study, compatibility effects for positive and negative flankers were of equal magnitude (Barratt & Bundesen, 2012, Exp. 1). Rather than trying to search for meaning in this enigmatic pattern of results, we hasten to add that all of these direct comparisons between flanker-related compatibility effects suffer from a fatal flaw because flanker compatibility effects are confounded with main effects of the target stimuli. That is, the compatibility effect for a positive flanker consists in subtracting the RT for a positive target (flanked by a positive face) from the RT of a negative target (flanked by a positive face), whereas the compatibility effect for a negative flanker is computed by subtracting the RT for a negative target (flanked by a negative face) from the RT of a positive target (flanked by a negative face). Such a confounding of flanker compatibility effects with (reverse-scored) differences in responding to positive and negative targets prevents any unambiguous interpretation of these comparisons. This is also the reason why most previous studies refrained from explicitly reporting or testing these comparisons. There is a way, however, to circumvent the problem of confounding the comparison of flanker-related compatibility effects with differences in target-related RTs, which is to control for overall main effects of different targets before computing and comparing the flanker-related compatibility effects. We will apply this technique in the present article to compute flanker-related compatibility effects for different types of emotional face stimuli, which allows us to compute and compare the strength of automatically (i.e., unintentionally) elicited response tendencies for different types of flankers without confounding these effects with differences in target-related RTs.

Finally, it should be noted that even flanker-related compatibility effects are not a pure indicator of attention allocation to the flanker. What is assessed by a flanker-specific response compatibility effect is the tendency to translate this flanker stimulus into a corresponding response, which may depend on attention allocation to this flanker but is not a pure measure of attention allocation to the flanker per se. To assess attention allocation pure and proper in the flanker paradigm, we recommend flanker distraction effects, which are described in the next section.

Distraction Effects

The rationale of an emotional flanker distraction effect is similar to an Emotional Stroop effect (Gotlib & McCann, 1984; Williams & Nulty, 1986): To compute emotional distraction effects in a flanker paradigm, the average RT for trials with neutral flankers is subtracted from the average RT for trials with emotional flankers (averaging across compatible and incompatible conditions for each type of flanker). Delayed responding in the emotional flanker condition (compared to the condition with neutral flankers) indicates that attention is attracted toward these task-irrelevant emotional flanker stimuli and away from the target stimuli. The same rationale can be used to compare distraction effects for different types of emotional flankers (e.g., to compare flanker interference for angry vs. fearful faces).

Flanker distraction effects offer a straightforward way to assess automatic attention allocation to emotional flankers. Distraction in this case does not imply and is in fact completely independent of response competition, because compatible as well as incompatible trials enter equally into each condition of the comparison that defines the flanker distraction effect. In some cases, when the main task is unrelated to emotion classification, response compatibility effects based on emotional expressions are completely eliminated and thus cannot be responsible for flanker distraction effects (e.g., see Experiments 4a and 4b below). Therefore, as in the emotional Stroop paradigm, emotional flanker distraction is nonspecific and is due to an allocation of attention to a spatial location, stimulus, or stimulus feature that is irrelevant for the task at hand, thus withdrawing resources from the processing of task-relevant stimuli and information (Algom, Chajut, & Lev, 2004).

Because of a focus on compatibility effects that is the core feature of standard flanker paradigms in cognitive psychology (Eriksen & Eriksen, 1974), flanker distraction effects have not been the main dependent variable in previous studies using the flanker task to assess attention allocation to emotional faces. There is only one study that directly assessed and compared flanker distraction effects for emotional and neutral face stimuli that were presented as irrelevant flanker stimuli in a letter classification task (Barratt & Bundesen, 2012, Exp. 2). Perhaps not surprisingly, this study did not reveal any evidence for emotional biases (if anything, letter categorization RTs in trials with negative flankers were even slightly faster than in the conditions with neutral and positive face flankers). This lack of an effect may be attributed to the fact that emotional processing was not required in this task at all (Spruyt et al., 2007), but it could also be due to the use of schematic instead of natural faces or to any of the specific parametric or procedural details of this study.

Although none of the other studies using the flanker paradigm for an investigation of emotional face processing focused on emotional distraction, these effects can nevertheless be inspected and were even tested (somewhat inadvertently) in some of these studies. Chen et al. (2016) reported stronger flanker distraction for negative compared to positive and neutral faces (which did not differ). However, this difference was present only in the subgroup of socially anxious participants, whereas there were no differences in flanker distraction between negative, positive, and neutral flankers for the nonanxious group. The other studies by Barratt and Bundesen (2012, Experiment 1), Fenske and Eastwood (2003), and Horstmann et al. (2006) found stronger distraction for negative compared to positive schematic faces (because the relevant tests were not conducted, this interpretation rests on visual inspection only). All of these studies, however, used schematic emotional faces. As was shown by Horstmann et al. (2006), asymmetries between positive and negative schematic faces are mostly driven by the low level perceptual features that characterize the geometric elements and shapes from which these faces were constructed, which are devoid of emotional meaning. A desiderate of this research is the necessity to investigate emotional biases in attention for natural face stimuli to be able to make valid conclusions regarding attention allocation to emotional faces in natural social settings.

To summarize, the flanker task is a promising and underresearched paradigm to investigate attention allocation to emotional faces that allows us to differentiate between direct (distraction) and indirect (compatibility) indicators of attention allocation to emotional faces. The flanker paradigm is a highly versatile instrument to vary different degrees of nonobligatory and obligatory processing of emotional information, which allows researchers to specify the exact conditions and the degree of automaticity (conditional vs. unconditional) under which attention allocation to emotional faces can be obtained. Up to now, there is not much literature on this paradigm for automatic emotional face processing, and the existing literature does not clearly differentiate and distinguish between different effect variables to assess and compare attention allocation to emotional and nonemotional flankers. Most of the published studies have relied on designs and/or analytic procedures that focus on target processing, which does not capitalize on the strength of the paradigm and incurs serious interpretational difficulties. Previous research thus has not yet exploited the full range of possibilities that is offered by this paradigm to study automatic attentional allocation to emotional faces.

Our aim in this study was to systematically investigate and compare distraction and compatibility effects for faces showing different emotional (angry, happy, fearful) and neutral expressions in the flanker paradigm. By using different variants of the paradigm varying in the degree to which emotional and flanker processing is obligatory or required in the task, we wanted to delineate the conditions under which attentional distraction or compatibility effects are obtained for emotional faces.

Another goal of our study was to investigate these attentional effects in the flanker paradigm for natural emotional faces. Because most of the existing studies using this paradigm relied on schematic faces (the only exception is the recent study by Chen et al., 2016, which had a strong focus on social anxiety), we wanted to fill this gap to overcome the interpretational and validity problems that are associated with the use of schematic materials (Horstmann et al., 2006).

Experiment 1

The first study consisted of a standard flanker task, that is, in each trial one central target face was flanked by two peripheral flanker faces (see Figure 1). The task was to categorize the emotional expression of the central target face (angry, fearful, neutral) by pressing one of three keys, and to ignore the flanker faces. We

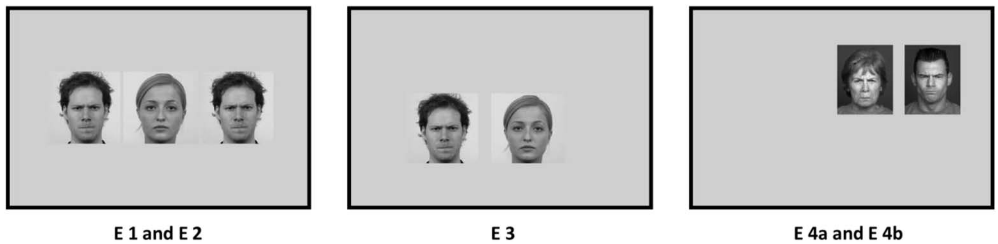

Figure 1. Example stimulus displays for the different experiments. Experiments 1 and 2 (left): Position of target (central) and flanker (left and right) stimuli was fixed across all trials. Participants' task was to identify the target emotion (angry, fearful, or neutral; Experiment 1), or to decide whether the target was emotional (angry, fearful) or neutral (Experiment 2). Experiment 3 (middle): Two horizontally aligned faces were presented in each trial. Position of the target varied across trials. Participants' task was to categorize the emotion (angry, fearful, happy, or neutral) of either the male or the female face (target gender alternated blockwise). Experiments 4a and 4b (right): Two horizontally aligned faces were presented in each trial. Position of the target face varied across trials. Participants' task was to categorize the age of either the male/female face (Experiment 4a) or the sex of the old/young face (Experiment 4b). Photographs are from the Radboud Faces Database (Langner et al., 2010) and from the FACES database (Ebner, Riediger, & Lindenberger, 2010).

decided to focus exclusively on emotional faces with a negative valence (angry, fearful) since previous studies mostly reported emotional biases toward negative emotional expressions. The flanker faces could have either the same or a different emotional expression as the target face. The position of the target and flankers did not change across trials, allowing participants to focus their attention at the central target position. Any effects of the flankers (distraction, compatibility) thus reflect unintended attention allocation.

Method

Sample. Forty-one participants (30 female) of different faculties were recruited at the campus of the Friedrich Schiller University Jena and were compensated with a small monetary incentive and a bar of chocolate. In this and in all subsequent experiments, we excluded participants with a high percentage of erroneous responses, using the criterion of "extreme values" (Tukey, 1977), for exclusion, that is, participants with error rates that were more than three interquartile ranges above the third quartile of the distribution of error rates were discarded from the analysis. In this study, all participants were below this criterion, leaving the full sample for analyses.

Materials. All stimuli were presented on a 17" CRT Monitor (XGA resolution 85 Hz) with the Psychopy software (Peirce, 2007, 2009). As targets and distractor stimuli we used greyscale frontal face photographs with different emotional expressions that were taken from the Radboud Faces Database (Langner et al., 2010; see Figure 1 for examples of the images). All face-photographs were presented as rectangles with a size of 250 × 250 pixels on a black background.[2] Flankers as well as targets had an angry, fearful, or neutral facial expression. All emotions appeared equally often as targets as well as flankers, that is, the experiment contained 3 × 3 = 9 different combinations of emotions, each of which was presented in 20 trials, yielding a total of 180 trials.

Procedure. Participants were seated in separate cabins and filled in an informed consent. After that the experimenter started the program, checked seating position, adjusted the chinrest to ensure a fixed distance of 57 cm to the screen, fitted the headset, and left the cabin. Every experiment started with a practice block to familiarize participants with the task. Each trial started with the presentation of a white fixation cross for 1000 ms. The fixation cross was then replaced by three face images, of which the middle one was determined to be the target and was located exactly in the center of the screen. Flankers were horizontally aligned to the target with a space of 10 pixels between the outer edges. The stimuli remained on the screen until a response was given via keyboard. Participants' task was to indicate the emotion of the central image. Keys were the left arrow, down arrow, and right arrow. Mapping of emotions (angry, fearful, neutral) to keys was counterbalanced. Errors were signaled by an error tone that was presented via headphones (100 ms, octave 5). Following a practice block (48 trials), the trials of the main experiment were presented in an individually randomized sequence.[3]

Results

After exclusion of erroneous trials (5.5%), outlier RTs (5.6%) were identified and eliminated (RTs that were more than 1.5 interquartile ranges above the third quartile of the intraindividual RT distribution [outliers according to Tukey, 1977] or that were below 200 ms). We then computed flanker distraction effects, flanker-related compatibility effects, and target-related compatibility effects in the RT and error data, and we tested for differences in the magnitude of these effects between the different types of stimuli. Mean RTs and error frequencies for Experiment 1 are shown in Table 1.

Flanker distraction effects. Flanker distraction effects reflect differences in RTs/errors between flanker conditions, averaging across the target emotion conditions within each flanker condition (see Table 2, Figure 2, and Figure 3 for distraction effects based on the RT and error data; Bayes factors for each type of flanker

[2] Experiments 1 and 2 also included trials in which frequency filtered photographs were presented, since another goal of these studies was to investigate the emotion-specific influence of frequency-filtering on attention allocation to emotional faces. Because this manipulation is not relevant for the present purpose, we excluded these trials from the analyses.

[3] The procedure of this and all other experiments reported in the present article was ethically approved as part of the evaluation procedure of the German Research Foundation (DFG) for the grant LA 3275/1-1 (PIs: Oliver Langner and Klaus Rothermund).

Table 1
Mean RTs (in Ms) and %errors for Combinations of Target and Flanker Emotions

Experiment	Target emotion	Flanker emotion: Angry		Fearful		Happy		Neutral	
RT									
E1	Angry	806		805		—		792	
	Fearful	787		794		—		782	
	Neutral	742		734		—		739	
E2	Angry	616		608		—		619	
	Fearful	583		582		—		588	
	Neutral	621		618		—		622	
E3	Angry	1,144		1,165		1,157		1141	
	Fearful	1,165		1,145		1,168		1155	
	Happy	966		966		956		955	
	Neutral	1,108		1,104		1,076		1055	
		Comp	Incomp	Comp	Incomp	Comp	Incomp	Comp	Incomp
E4a	Angry	682	705	688	711	686	705	682	704
	Fearful	680	710	682	706	682	711	688	702
	Happy	684	701	691	701	686	702	687	697
	Neutral	684	697	678	701	682	701	686	706
E4b	Angry	710	750	718	742	—	—	715	737
	Fearful	713	745	714	749	—	—	709	740
	Neutral	710	752	711	740	—	—	712	743
ERR									
E1	Angry	7.66		5.73		—		8.05	
	Fearful	6.35		6.71		—		6.85	
	Neutral	3.17		3.17		—		3.54	
E2	Angry	7.00		6.75		—		7.13	
	Fearful	4.00		5.25		—		2.50	
	Neutral	3.88		4.88		—		4.31	
E3	Angry	8.93		9.12		10.39		9.90	
	Fearful	7.22		6.24		6.98		7.61	
	Happy	3.27		2.93		1.90		2.98	
	Neutral	4.68		5.07		4.49		3.61	
		Comp	Incomp	Comp	Incomp	Comp	Incomp	Comp	Incomp
E4a	Angry	3.88	5.19	3.56	5.44	4.00	5.25	4.06	5.63
	Fearful	3.50	4.94	3.12	6.19	3.56	5.25	4.06	4.69
	Happy	4.31	5.25	3.81	5.38	3.56	4.87	3.69	6.00
	Neutral	4.06	5.75	3.25	5.06	3.75	5.63	3.13	4.06
E4b	Angry	2.95	6.22	2.82	6.67	—	—	3.65	6.09
	Fearful	3.59	5.26	3.85	5.71	—	—	2.69	5.83
	Neutral	3.53	5.83	2.82	6.09	—	—	3.14	4.94

Note. Response compatibility (Comp = compatible, Incomp = Incompatible) is another relevant factor that is independent of the flanker and target emotions for Experiments 4a and 4b, because in these experiments responses were determined by gender or age, respectively.

distraction effect are shown in Table A1 in the Appendix). An analysis of variance (ANOVA) with flanker type as a factor did not reveal significant differences between the three flanker conditions, neither for RTs, $F(2, 39) = 2.19$, $p = .13$, nor for errors, $F(2, 39) = 2.15$, $p = .13$, indicating that emotional and neutral flankers did not differ in their general distractive potential. Specific flanker distraction effects for each of the two emotions (angry, fearful) were computed by contrasting the respective emotional flanker condition to the neutral flanker condition. These analyses revealed significant distraction effects for angry compared to neutral faces for both the RT and error data. Importantly, the direction of the effect was opposite for the RT and error data, indicating stronger distraction for angry compared to neutral flanker faces in the RT data, but a reverse effect indicating a stronger distraction of the neutral compared to angry flanker faces for the error data. RTs and errors for the fearful flankers did not differ significantly from the neutral or the angry faces. Finally, there was no significant difference in flanker distraction between the emotional (angry and fearful) versus neutral flanker faces.

Orthogonal flanker distraction effects. Another way to compute flanker distraction effects that is fully independent of and orthogonal to compatibility effects is to compare average RTs/error frequencies for flanker conditions in which compatible and incompatible conditions are equally weighted within each flanker condition by computing the mean of the compatible condition and the average of the incompatible conditions. Like the analyses of flanker-related compatibility effects, computing these compatibility-balanced flanker distraction effects requires controlling for target main effects in a first

Table 2
Flanker Distraction Effects for the Reaction Time (RT) (in Ms) and Error Data (in %errors), Computed as the Difference in Average RTs and %errors Between Emotional and Neutral Flanker Conditions (Positive Values Indicate Stronger Distraction for Emotional Than Neutral Flankers)

Experiment (E)	Emotional vs. neutral	Angry vs. neutral	Fearful vs. neutral	Happy vs. neutral
RT (SE_{diff})				
E1	7 (4)	8^{*} (4)	7 (6)	—
E2	−5 (3)	−3 (4)	-7^{*} (3)	—
E3	17^{**} (5)	19^{**} (6)	19^{**} (5)	13^{*} (6)
E4a	−0 (2)	−1 (3)	1 (2)	0 (2)
E4b	3 (2)	4 (3)	3 (3)	—
ERR% (SE_{diff})				
E1	−.98 (.5)	-1.02^{*} (.5)	−.94 (.6)	—
E2	.65 (.4)	.31 (.5)	.98 (.5)	—
E3	−.089 (.3)	.00 (.4)	−1.83 (.4)	−.85 (.4)
E4a	.11 (.2)	.20 (.3)	.06 (.2)	.07 (.2)
E4b	.22 (.3)	.17 (.3)	.27 (.3)	—

$^{*}\ p < .05.\quad ^{**}\ p < .01.$

step to avoid a confounding with differences in target-related RTs/errors. This way of computing flanker distraction effects yielded a similarly inconsistent picture, with no reliable difference between the three flanker conditions, $F(2, 39) = 1.58$, $p = .22$, for RTs, and $F(2, 39) = 1.86$, $p = .17$, for errors. The pattern of means was again opposite for RTs and errors, with RT data suggesting a tendency for stronger distraction for angry compared to neutral flanker faces, $t(40) = 1.72$, $p = .09$, and an opposite tendency for stronger distraction for neutral compared to angry flanker faces, $t(40) = -1.92$, $p = .06$, in the error data.

Flanker-related compatibility effects. Flanker-related compatibility effects were computed after controlling for target main effects. That is, we subtracted the deviation between the mean RT/error rate for each target emotion (averaging across flanker conditions) and the grand mean from the RTs/errors within the respective target condition. Following this correction, flanker-related compatibility effects were computed by contrasting the compatible and the incompatible target conditions for each flanker (e.g., for angry flankers, the flanker compatibility effect reflects the difference between the average of the incompatible conditions [fearful target/angry flanker, neutral target/angry flanker] and the compatible condition [angry target/angry flanker]).

None of the flanker-related compatibility effects reached significance, neither for the RT nor for the error data (see Table 3, Figure 4, and Figure 5; Bayes factors for each type of flanker related compatibility effect are shown in Table A2 in the Appendix). Nor did we find a significant overall compatibility effect (averaging across all types of flankers), $F(1, 40) = 1.67$, $p = .20$, for RTs, $F < 1$, for errors, or significant differences between the emotional and neutral flanker compatibility effects, all $F < 1$ for both RTs and errors.

Target-related compatibility effects. Target-related compatibility effects were computed by comparing compatible and incompatible conditions within each target condition. Although we argued in the introduction that these target-related effects do not reflect a straightforward index of attention allocation to emotional (or neutral) faces in a standard flanker task, we nevertheless report these effects (a) for the sake of completeness and (b) to allow comparisons with previous studies that also reported these effects.

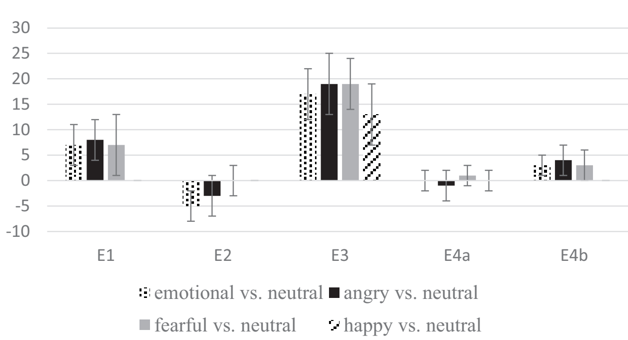

Figure 2. Flanker distraction effects. Reaction time (RT)-effects computed as the difference in average RTs between emotional and neutral flanker conditions (positive values indicate stronger distraction for emotional than neutral flankers).

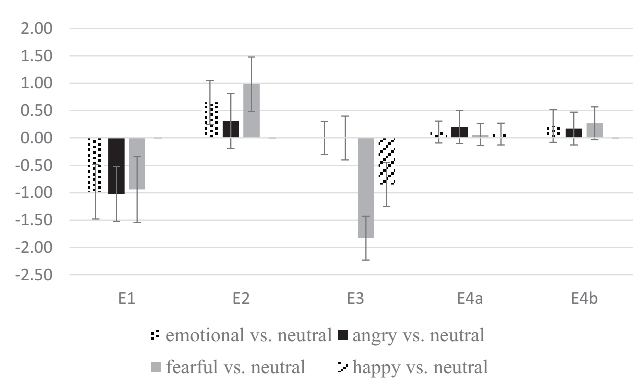

Figure 3. Flanker distraction effects. %error-effects computed as the difference in average %errors between emotional and neutral flanker conditions (positive values indicate stronger distraction for emotional than neutral flankers).

Table 3
Flanker-Related Compatibility Effects for the Reaction Time (RT) and Error Data, Computed as the Difference in Average RTs (in Ms) and %errors Between Compatible and Incompatible Trials Within the Respective Flanker Condition (Positive Values Indicate Facilitation in Compatible Compared to Incompatible Trials)

Experiment (E)	Overall	Angry	Fearful	Happy	Neutral	Compatibility relation
RT (SE_{diff})						
E2	−6 (5)	−3 (6)	−7 (7)	—	−8 (6)	Emotion
E3	2 (3)	1 (4)	2 (5)	—	3 (4)	Emotion
E4a	18^{***} (4)	19 (10)	25^{**} (7)	6 (6)	24^{**} (7)	Emotion
E4b	20^{***} (2)	21^{***} (3)	20^{***} (4)	20^{***} (3)	17^{***} (3)	Age
E2	32^{***} (3)	38^{***} (5)	29^{***} (5)	—	28^{***} (4)	Gender
ERR% (SE_{diff})						
E1	.18 (.5)	.48 (.6)	−.54 (.8)	—	.61 (.9)	Emotion
E2	−.39 (.4)	−.54 (.6)	−.04 (.5)	—	−.58 (.6)	Emotion
E3	1.05^{**} (.3)	.97 (.6)	.87 (.5)	1.13^{*} (.5)	1.23^{**} (.4)	Emotion
E4a	1.58^{**} (.4)	1.34^{**} (.4)	2.08^{***} (.5)	1.53^{**} (.5)	1.36 (.7)	Age
E4b	2.62^{***} (.4)	2.41^{***} (.4)	2.99^{***} (.5)	—	2.46^{***} (.5)	Gender

Note. Main effects of target emotion were controlled before computing flanker-related compatibility effects in Experiments 1, 2, and 3 (in Experiments 4a and 4b, *response compatibility* refers to age and gender, respectively, so flanker-related compatibility effects are not confounded with target emotion).
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

None of the target-related compatibility effects reached significance, neither for the RT nor for the error data (see Table 4, Figure 6, and Figure 7; Bayes factors for each type of target-related compatibility effect are shown in Table A3). Nor did we find significant differences between the emotional and neutral target compatibility effects, all $F < 1$ for both RTs and errors. Overall compatibility effects (averaging across all types of targets) are identical to what was reported in the previous paragraph for flanker-related compatibility effects.

Discussion

We did not find any reliable evidence for biased attention allocation to emotional faces in any of the analyses. The only hint at biased processing emerged in the analysis of distraction effects for the contrast between angry and neutral faces. However, this finding was enigmatic because the direction of the effect was opposite for RT and error data. Another reason for doubting the reliability of this finding is that the overall analysis across all flanker conditions did not indicate significant differences for distraction effects between angry, fearful, and neutral flanker faces. Bayes factor analyses did not provide positive support for either global or specific flanker distraction effects (all $BF_{10} < 1.20$) but likewise did not provide strong support for the corresponding null hypotheses (all $BF_{01} < 2.92$).

We also did not find any indication of an attentional bias toward emotional faces in the analyses of flanker- and target-related compatibility effects, which did not differ between emotional and

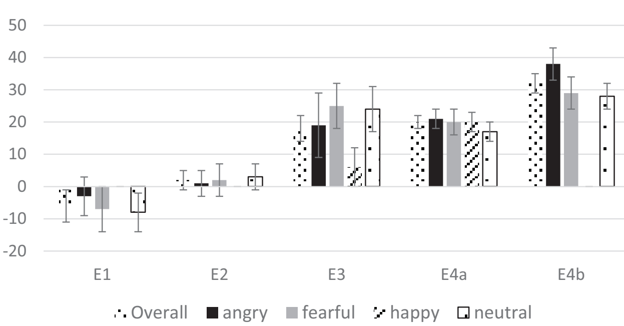

Figure 4. Flanker-related compatibility effects for the reaction time (RT) data, computed as the difference in average RTs (in ms) between compatible and incompatible trials within the respective flanker condition (positive values indicate facilitation in compatible compared to incompatible trials). Main effects of target emotion were controlled before computing flanker-related compatibility effects in Experiments 1, 2, and 3 (in Experiments 4a and 4b, response compatibility refers to age and gender, respectively, so flanker-related compatibility effects are not confounded with target emotion).

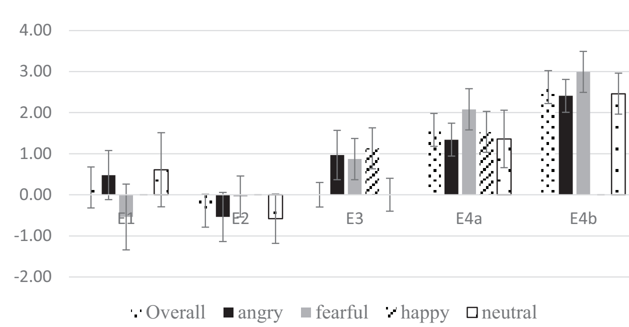

Figure 5. Flanker-related compatibility effects for the %error data, computed as the difference in average %errors between compatible and incompatible trials within the respective flanker condition (positive values indicate facilitation in compatible compared to incompatible trials). Main effects of target emotion were controlled before computing flanker-related compatibility effects in Experiments 1, 2, and 3 (in Experiments 4a and 4b, response compatibility refers to age and gender, respectively, so flanker-related compatibility effects are not confounded with target emotion).

Table 4
Target-Related Compatibility Effects for the Reaction Time (RT) and Error Data, Computed as the Difference in Average RTs (in Ms) and %errors Between Compatible and Incompatible Trials Within the Respective Target Condition (Positive Values Indicate Facilitation in Compatible Compared to Incompatible Trials)

Experiment	Overall	Angry	Fearful	Happy	Neutral	Compatibility relation
RT (SE_{diff})						
E1	−6 (5)	−7 (7)	−10 (10)	—	−0 (8)	Emotion
E2	3 (3)	7 (5)	5 (7)	—	−3 (3)	Emotion
E3	18*** (4)	10 (12)	17 (8)	6 (7)	40*** (7)	Emotion
E4a	20*** (2)	22*** (3)	24*** (4)	13** (4)	19*** (3)	Age
E4b	32*** (3)	29*** (5)	32*** (4)	—	34*** (4)	Gender
ERR% (SE_{diff})						
E1	.18 (.5)	1.02 (.9)	−.10 (1.1)	—	−.37 (.9)	Emotion
E2	−.60 (.4)	.25 (.9)	−2.13** (.7)	—	.06 (.6)	Emotion
E3	1.05** (.3)	.88 (.6)	1.02 (.5)	1.15* (.6)	1.14 (.6)	Emotion
E4a	1.58** (.4)	1.50* (.6)	1.70** (.6)	1.53** (.5)	1.58*** (.4)	Age
E4b	2.62*** (.4)	3.18*** (.6)	2.22*** (.5)	—	2.46*** (.5)	Gender

* $p < .05$. ** $p < .01$. *** $p < .001$.

neutral stimuli. An unexpected and puzzling finding in this regard, however, is the complete absence of overall compatibility effects. A possible reason for this lack of a significant compatibility effect might be the complexity of the categorization task that required participants to discriminate between three categories with three different keys, which is uncommon for flanker studies that typically use only a two-alternative categorization. This type of task also produces a large number of conflicting trials (in two out of three trials flankers are incompatible with the target), which should generally reduce compatibility effects due to cognitive control processes that lead to a focusing on the task-relevant information and that become stronger after incompatible trials (e.g., Gratton effect; Gratton, Coles, & Donchin, 1992) or in contexts with a high percentage of incompatible trials (e.g., Logan & Zbrodoff, 1979).

Experiment 2

The second experiment is a conceptual replication of the first study in which we changed the categorization task to a simple binary decision between emotional and neutral faces that is more typical to flanker tasks that have been used in the literature. This task produces equal numbers of compatible and incompatible trials which allows us to test whether this asymmetry might be responsible for the lack of compatibility effects that we encountered in the previous experiment.

Method

Sample. Forty-one participants (29 female) took part in this study and were recruited similarly to Experiment 1. One participant was excluded due to a large number of errors (25%) that was more than three interquartile ranges above the third quartile of the distribution of error rates, leaving an effective sample of 40 participants (28 female).

Procedure. The second experiment was similar to the first one except for the following changes. The emotion identification task of the previous experiment was replaced with a binary decision between emotional (angry, fearful) and neutral faces (assignment of response keys to categories was counterbalanced). To have equal numbers of neutral and emotional targets in the task, neutral

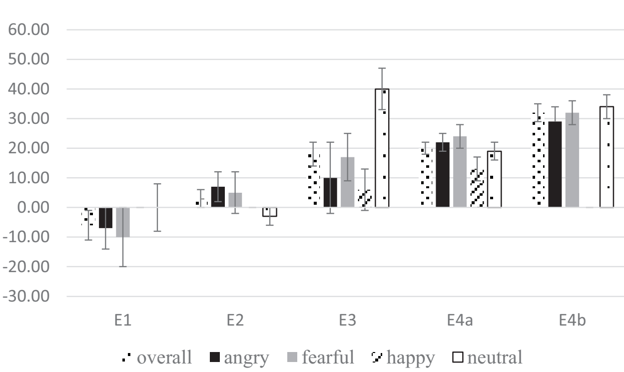

Figure 6. Target-related compatibility effects for the RT data, computed as the difference in average RTs (in ms) between compatible and incompatible trials within the respective target condition (positive values indicate facilitation in compatible compared to incompatible trials).

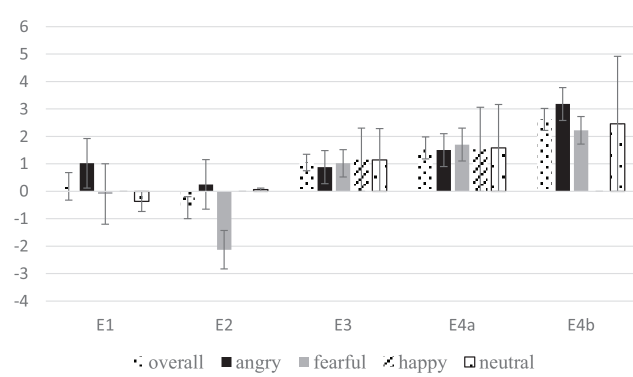

Figure 7. Target-related compatibility effects for the %error data, computed as the difference in average %errors between compatible and incompatible trials within the respective target condition (positive values indicate facilitation in compatible compared to incompatible trials).

faces were presented twice as much as the angry and fearful faces. This increases the number of experimental trials to 240 instead of 180 in Experiment 1. In all other respects, the procedure and materials were identical as in Experiment 1.

Results

After exclusion of erroneous trials (5.6%), outlier RTs (5.1%) were identified and eliminated (RTs that were more than 1.5 interquartile ranges above the third quartile of the intraindividual RT distribution [outliers according to Tukey, 1977] or that were below 200 ms). We then computed flanker distraction effects, flanker-related compatibility effects, and target-related compatibility effects in the RT and error data, and we tested for differences in the magnitude of these effects between the different types of stimuli. Mean RTs and error frequencies for Experiment 2 are shown in Table 1.

Flanker distraction effects. Flanker distraction effects for the emotional flankers based on RTs/errors are shown in Table 2 (see also Figure 2 and Figure 3). An ANOVA with flanker type as a factor missed significance, $F(2, 38) = 3.13$, $p = .06$, for RTs, and $F(2, 38) = 1.76$, $p = .19$, for errors, indicating that emotional and neutral flankers did not differ reliably in their general distractive potential. Specific flanker distraction effects for each emotion were computed by contrasting the respective emotional flanker condition to the neutral flanker condition. These analyses revealed a significantly weaker distraction effect for fearful compared to neutral faces for the RT data. No other comparisons were significant. Finally, there was no significant difference in flanker distraction between the emotional (angry and fearful) versus neutral flanker faces.

Orthogonal flanker distraction effects. The pattern of results was similar for compatibility-balanced flanker distraction effects (controlled for target main effects), with fearful faces showing that distraction was significantly weaker for the fearful compared to the neutral flanker faces, $t(39) = -2.49$, $p < .05$, for the RT data. This difference was not found in the error data, nor did angry and neutral face flankers differ with regard to distraction in either the RT or error data, all $|t| < 1.92$.

Flanker-related compatibility effects. Flanker-related response compatibility effects (controlled for target main effects; see Table 3, Figure 4, and Figure 5) were nonsignificant overall, $F < 1$, for RTs and for errors. None of the flanker-specific compatibility effects was significant.

Target-related compatibility effects. Overall, target-related compatibility effects did not differ systematically between emotional and neutral targets, $F < 1$, for RTs, $F(2, 38) = 3.25$, $p = .05$, for errors. Pairwise comparisons showed that compatibility effects were smaller for fearful than for neutral targets in the error data (see Table 4, Figure 6, and Figure 7). None of the other effects reached significance.

Discussion

Results were highly similar to Exp. 1. Again, we found no evidence of attentional biases toward emotional faces, neither with regard to distraction nor compatibility effects. The compatibility effects for fearful flankers (in the RT data) and targets (in the error data) were smaller compared to the neutral flankers/targets, but these two effects have opposite meanings, and should not be interpreted because overall analyses did not reveal significant differences between emotional and neutral stimuli regarding the strength of compatibility effects in the first place. Most of the Bayes factors indicated substantial support for the null hypotheses ($BF_{01} > 3$, see Tables A1–A3 in the Appendix).

In sum then, the first two experiments did not reveal any hint of attentional biases toward emotional faces. A surprising finding of these two studies was that although emotion was a task-relevant feature (target faces had to be categorized according to their emotion), we did not find a significant overall compatibility effect in any of the two studies, neither in the RT nor in the error data. Apparently, flankers were not even processed up to the point where emotional expressions are identified and translated into response tendencies. Participants may have succeeded in focusing their attention entirely on the targets, which was made possible by the fixed position of the targets across trials, which allowed for a focusing on the central area of the screen where the target stimuli were presented, and to filter out any distracting information from the flankers. Although such a fixed spatial arrangement is also typical for standard flanker paradigms in cognitive psychology that nevertheless show highly reliable flanker interference effects, a crucial difference of the current study using natural faces as stimuli is that face stimuli cover a much larger spatial area than do simple stimuli that are used in nonemotional tasks (e.g., arrows, letters, words). In such a setting, it is much easier to limit the breadth of the attentional spotlight so that flanking face stimuli are excluded from processing.

To address this problem, we conducted a series of further experiments in which the spatial position of the target stimulus was variable and unpredictable, and thus rendered the processing of the flankers obligatory.

Experiment 3

In Experiment 3, only two horizontally aligned face stimuli were presented, one of which was male, and the other female (left/right position of the male and female faces varied randomly across trials). The target face had to be identified on the basis of its gender. This task made it impossible to specify the target position in advance. In addition, the spatial position of the pair of faces on the screen varied across trials, which also made it impossible to focus attention on a narrow spatial region of the display. These changes should make sure that flankers are attended to a certain extent, by making it impossible to focus only on the target stimuli and to fully ignore the flankers. In addition, we added happiness as a third emotion to the design. Because some studies revealed stronger attention allocation to positive stimuli (e.g., Juth, Lundqvist, Karlsson, & Öhman, 2005), we wanted to investigate attentional biases for both poles of the valence spectrum. This also allowed us to compare attentional biases for positive and negative facial expressions of emotions. Finally, we increased the number of experimental trials to increase the power of detecting attentional biases.

Method

Sample. Forty-three participants (26 female) were recruited at the campus of the Friedrich Schiller University Jena and received chocolate and money for their participation. Two participants were excluded due to a large number of errors (>25%) that were more than three interquartile ranges above the third quartile of the distribution of error rates, leaving an effective sample of 41 participants (26 female).

Procedure. The third experiment was designed to enforce flanker processing by requiring an additional classification to identify the target, that is, participants had to classify the emotion of either the male or the female face on the screen (target gender was counterbalanced across blocks within participants). Every screen contained two faces of different gender that were horizontally aligned and separated by 50 pixels (see Figure 1). The location of the stimuli on the screen was determined randomly with the limitation that the center of the two-image-array had to be within the central 400 × 200 (w × h) pixels rectangle. The uncertainty of the location was ought to further promote flanker processing by preventing that attention was focused on a narrow spatial area. Participants had to determine the emotion (angry, fearful, happy, neutral) of the male (female) face by pressing one of four keys (“q,” “w,” “o,” “p”). The mapping was counterbalanced. After 64 practice trials (two blocks of 32 trials), 800 experimental trials were presented in 10 blocks of 80 trials. The combination of four target emotions, four flanker emotions, and two target genders led to 32 possible trial types, each of which was repeated 25 times throughout the experiment. All trials with male (female) targets were then presented randomly in either the odd-numbered (even-numbered) blocks for one half of the participants and with a reversed sequence for the other half of the participants. Instructions to either react on the male or female target were given accordingly. All other parameters as well as data analysis were similar to Experiment 1.

Results

After exclusion of erroneous trials (5.96%), outlier RTs (5.12%) were identified and eliminated (RTs that were more than 1.5 interquartile ranges above the third quartile of the intraindividual RT distribution [outliers according to Tukey, 1977] or that were below 200 ms). We then computed flanker distraction effects, flanker-related compatibility effects, and target-related compatibility effects in the RT and error data, and we tested for differences in the magnitude of these effects between the different types of stimuli. Mean RTs and error frequencies for Exp. 3 are shown in Table 1.

Flanker distraction effects. Flanker distraction effects for the emotional flankers based on RTs/errors are shown in Table 2; see also Figure 2 and Figure 3 (Bayes factors are shown in Table A1 in the Appendix). An ANOVA with flanker type as a factor yielded a significant effect for the RT data, $F(3, 38) = 5.38$, $p = .01$, but not for the error data, $F < 1$. Specific flanker distraction effects for each emotion were computed by contrasting the respective emotional flanker condition to the neutral flanker condition. These analyses revealed significantly stronger distraction effects for all emotional compared to the neutral faces for the RT data. No significant differences were obtained between the different emotional flankers, $F(2, 39) = 1.27$, $p = .29$. None of the comparisons were significant for the error data.

Orthogonal flanker distraction effects. The pattern of results was similar for compatibility-balanced flanker distraction effects (controlled for target main effects), with emotional faces showing larger distraction effects than neutral flankers for the RT data, both overall, $t(40) = 3.59$, $p < .01$, and also for each single emotion condition, all $t(40) \geq 2.80$, $p < .01$. No differences in the strength of flanker-specific distraction effects were obtained for the error data, $F < 1$.

Flanker-related compatibility effects. Flanker-related response compatibility effects (controlled for target main effects; see Table 3, Figure 4, and Figure 5; for Bayes factors, see Table A2 in the Appendix) were significant overall, for RTs and for errors. All flanker-specific compatibility effects were positive. Significant flanker-specific compatibility effects obtained for fearful and neutral faces, in the RT data, and for happy and neutral faces, in the error data. A comparison of flanker-specific compatibility effects revealed no significant differences between the four types of flankers, neither for the RT data, $F(3, 38) = 1.90$, $p = .15$, nor for the error data, $F < 1$. A direct comparison of flanker-compatibility effects for emotional versus neutral flankers likewise did not reveal any significant differences, $F < 1$ for both RT and error data.

Target-related compatibility effects. Target-related compatibility effects were also significant overall (see Table 4, Figure 6, and Figure 7; for Bayes factors, see Table A3 in the Appendix), for both RT and error data, with all target-specific compatibility effects being positive. A comparison of target-specific compatibility effects revealed significant differences between the four types of targets for the RT data, $F(3, 38) = 3.55$, $p < .05$, but not for the error data, $F < 1$. The main effect of target type for the RT data was entirely due to the fact that compatibility effects were significantly larger for the neutral compared to the emotional targets, $F(1, 40) = 7.14$, $p < .05$, indicating that emotional targets drew more attention than neutral targets, leaving less room for flanker interference. Differences in compatibility effects between the emotional targets were not significant, $F(2, 39) < 1$.

Discussion

Experiment 3 revealed evidence for emotional biases in attention allocation. First and foremost, flanker-specific distraction effects were stronger for emotional compared to neutral face flankers. This result is corroborated by Bayes factor analyses, which provided strong support for the H_1 for each test (all $BF_{10} > 17$). The strength of this difference was similar for all emotional faces and did not differ significantly between face flankers showing angry, fearful, or happy expressions. Additional evidence for an emotional bias was obtained in the analysis of target-specific compatibility effects, which were smaller for emotional than for neutral targets, indicating that the influence of flanker stimuli was generally weaker for emotional targets. The most probable explanation for this finding is that emotional faces attracted attention stronger than did neutral faces and thus had a higher probability of being focused first. This attentional capture effect supports selection of emotional targets by increasing the likelihood that emotional targets fell into the focus of attention before the flanker face. No evidence for an emotional bias was obtained for flanker-specific compatibility effects. Thus, although emotional flankers might be focused with a higher probability than were neutral flankers, this did not automatically result in a translation of the emotional content into a corresponding response tendency which would be necessary for a response compatibility effect to emerge. We reflect on the implications of this dissociation between flanker distraction and compatibility effects in the General Discussion. However, contrary to the previous experiments, overall response compatibility effects were robust and clearly significant in the present experiment, for both RTs and errors. The current version of the paradigm apparently succeeded in guaranteeing a high level of flanker processing, which produced compatibility effects, and also revealed evidence for an emotional bias.

Nevertheless, it must be noted that the current experiment not only made flanker processing obligatory by making target identification dependent on gender identification. At the same time, the target task also required a processing of the emotional content of the faces. It is thus unclear whether the emotional biases that were obtained resulted from obligatory flanker processing, or whether they are also dependent on the fact that emotional processing was task-relevant. For our understanding of emotional biases, it makes a huge difference whether attention to emotional faces occurs unconditionally, whenever a face is encountered and processed, or whether such an effect occurs only under conditions when emotional processing is the main goal of the perceiver. To disentangle these two possibilities, we conducted another flanker study with emotional face stimuli in which the emotional expressions were no longer task relevant.

Experiment 4

In the final study (Experiments 4a and 4b), we rendered flanker processing obligatory by a target identification task, as in Experiment 3 (targets could not be identified on the basis of their spatial position but had to be identified either on the basis of their gender [male vs. female; Experiment 4a] or age [young vs. old; Experiment 4b]). In contrast to the previous study, however, emotional expressions of the presented face stimuli were no longer task relevant. Instead, participants had to categorize either the age (old vs. young, when target identification was based on gender) or gender (male vs. female, when target identification was based on age) of the target face. Emotion thus was completely irrelevant for the task. Like in the previous experiments, we analyzed flanker-distraction effects for emotional and neutral flanker faces, and we also investigated whether the emotional expression of the flanker or target faces moderated the strength of the response compatibility effects that were elicited by these faces. Importantly, however, these compatibility effects were no longer based on the emotional expression itself. The advantage of this type of analysis is thus that emotional biases in these effects reflect pure attentional capture and are unrelated to translating emotions into responses.

Method

Sample. Experiment 4a included 40 participants (20 female). Another 40 participants (22 female) took part in Experiment 4b. One participant was excluded due to a large number of errors (>24%), which was more than three interquartile ranges above the third quartile of the distribution of error rates, leaving an effective sample of 39 participants (22 female). Participants were recruited and rewarded analogously to the previous studies.

Procedure. In both experiments, two faces were presented in each trial. In Experiment 4a, the faces always differed with respect to their sex (one male, one female), and participants were asked to identify the age (young, old) of either the male or female face. In Experiment 4b, the faces always differed with respect to their age (one old, one young), and participants were asked to identify the sex (male, female) of either the young or old face. The relevant category for the target identification task (male/female in Experiment 4a; old/young in Experiment 4b) alternated blockwise within participants. Responses were given with the "*d*" and "l" button of the keyboard. Size and position of the stimuli were equal to experiment three, but faces were taken from the FACES (Ebner, Riediger, & Lindenberger, 2010) database (see Figure 1). Emotional expressions of the target and flanker faces were irrelevant for the task and were varied independently of each other and orthogonally to the age and gender of the faces. In Experiment 4a, emotional expressions of the faces were angry, fearful, happy, or neutral. In Experiment 4b, we presented faces that had angry, fearful, or neutral expressions.[4] Two practice blocks with only neutral faces of 16 trials each were presented to familiarize participants with identification and categorization tasks, and were followed by 1,280 (Experiment 4a) or 720 (Experiment 4b) experimental trials with emotional faces that were presented in 10 blocks. For Experiment 4a, the number of trials resulted from presenting 10 instances for each combination of a 2 (target age: young vs. old) × 2 (flanker age: young vs. old) × 2 (target position: right vs. left) × 4 (target emotion) × 4 (flanker emotion), five of which were presented in the blocks in which the target face was male (and the flanker face was female), the other five instances were presented in the blocks in which the target face was female (and the flanker face was male). For Experiment 4b, the number of trials resulted from presenting 10 instances for each combination of a 2 (target age: young vs. old) × 2 (flanker age: young vs. old) × 2 (target position: right vs. left) × 3 (target emotion) × 3 (flanker emotion), five of which were presented in the blocks in which the target face was male (and the flanker face was female), the other five instances were presented in the blocks in which the target face was female (and the flanker face was male).

Results

After exclusion of erroneous trials (4.50%, Experiment 4a; 4.54%, Experiment 4b), outlier RTs (4.62%, Exp. 4a; 4.89%, Exp. 4b) were identified and eliminated (RTs that were more than 1.5 interquartile ranges above the third quartile of the intraindividual RT distribution [outliers according to Tukey, 1977] or that were below 200 ms). We then computed flanker distraction effects, flanker-related compatibility effects, and target-related compatibility effects in the RT and error data, and we tested for differences in the magnitude of these effects between the different types of stimuli. Mean RTs and error frequencies for Experiments 4a and 4b are shown in Table 1.

[4] To strengthen attentional capture effects for negative expressions, for which attentional capture had been sometimes stronger than for positive expressions in previous studies, we decided to drop the happy condition. The reason was that including the happy expression might render the negative expressions more similar to the neutral expression because the happy expression is the most distinctive one. This might lead to an implicit grouping of faces into happy and not-happy ones.

Flanker distraction effects. Flanker distraction effects for the emotional flankers based on RTs/errors are shown in Table 2 (see also Figure 2 and Figure 3). An ANOVA with flanker type as a factor yielded no significant effects for either the RT data (Experiment 4a: $F < 1$; Experiment 4b: $F < 1$) or the error data (Experiment 4a: $F < 1$; Experiment 4b: $F < 1$). Specific flanker distraction effects for each emotion were computed by contrasting the respective emotional flanker condition to the neutral flanker condition. These analyses revealed no significant differences in the strength of distraction between emotional and neutral flankers for either the RT data (Experiment 4a: all $t < 1$; Experiment 4b: all $t < 1.29$, $p > .20$) or the error data (Experiment 4a: all $t < 1$; Experiment 4b: all $t < 1$).

Orthogonal flanker distraction effects. Because emotional expressions of flankers and targets are orthogonal to the task-relevant dimension, flanker interference effects in these experiments are compatibility-balanced by default.

Flanker-related compatibility effects. Flanker-related response compatibility effects were computed by contrasting compatible and incompatible trials for the task-relevant dimension within each flanker condition. For Experiments 4a and 4b, flanker emotions are fully orthogonal to all target features, thus these analyses require no controlling for any kind of target main effects. Compatibility effects were highly significant overall, for RTs—Experiment 4a: $t(39) = 12.75$, $p < .001$; Experiment 4b: $t(38) = 10.04$, $p < .001$—and for errors—Experiment 4a: $t(39) = 3.76$, $p < .001$; Experiment 4b: $t(38) = 7.27$, $p < .001$. All flanker-specific compatibility effects were significant for RTs—Experiment 4a: all $t(39) > 4.98$, $p < .001$; Experiment 4b: all $t(38) > 5.65$, $p < .001$—and also for errors—Experiment 4a: $t(39) > 1.98$, $p < .055$; Experiment 4b: all $t(38) > 5.29$, $p < .001$. A comparison of flanker-specific compatibility effects revealed no significant differences between the different flankers, neither for the RT data—Experiment 4a: $F < 1$; Experiment 4b: $F(2, 37) = 1.65$, $p > .20$—nor for the error data—Experiment 4a: $F < 1$; Experiment 4b: $F < 1$. A direct comparison of flanker-compatibility effects for emotional versus neutral flankers likewise did not reveal any significant differences for both RT—Experiment 4a: $F(1, 39) = 1.82$, $p = .19$; Experiment 4b: $F(1, 38) = 1.79$, $p = .18$—and error data—Experiment 4a: $F < 1$; Experiment 4b: $F < 1$.

Target-related compatibility effects. Like in the previous analyses, target-related compatibility effects were also significant overall (see Table 4, Figure 6, and Figure 7), for both RT (Experiment 4a: $t(39) = 12.75$, $p < .001$; Experiment 4b: $t(38) = 10.04$, $p < .001$) and error data (Experiment 4a: $t(39) = 3.76$, $p < .001$; Experiment 4b: $t(38) = 7.27$, $p < .001$). A comparison of target-specific compatibility effects revealed no significant differences between the different types of targets, neither for the RT data (Experiment 4a: $F < 1$; Experiment 4b: $F < 1$), nor for the error data (Experiment 4a: $F < 1$; Experiment 4b: $F(2, 37) = 1.30$, $p > .28$).

Discussion

We found no evidence whatsoever for attentional biases toward emotional stimuli in the present experiments. Distraction and compatibility effects were indistinguishable for emotional and neutral stimuli for both Experiment 4a and 4b, with not the slightest hint at emotional-neutral attentional asymmetries regarding flanker and target stimuli alike. Accordingly, Bayes factors for emotional distraction effects provide substantial support for the H_0 (BF_{01}'s ranging between 3 and 6, see Table A1 in the Appendix). This finding is all the more remarkable since like in the previous experiment, flanker processing was obligatory in the present experiments as well (the spatial position of the target was variable and unpredictable across trials, and targets had to be identified based on their gender or age). Accordingly, we found robust response compatibility effects for the flanker stimuli, indicating that flankers were indeed processed.

The null findings that obtained in Experiments 4a and 4b also have important implications for the interpretation of the findings that did obtain in Experiment 3. A possible explanation for these effects might have been that Experiment 3 was much more complex due to the 2-step process that was needed to identify and categorize the target, and thus more resource demanding than the previous experiments. Although it is by no means clear why an increase in task demands should lead to stronger flanker effects—to the contrary, load theory (Lavie & Tsal, 1994) suggests that a high cognitive load depletes the resources that are required for distractor processing—the possibility remains that the difference in findings is due to differences in task complexity. Such an alternative explanation, however, is made even less plausible by the results of Experiments 4a and 4b that employed a similar two-step task but did not yield any evidence for emotional biases.

The crucial difference between the present experiments (Experiments 4a and 4b) and the previous experiment (Experiment 3) is that the categorization task was unrelated to emotion (target faces had to be categorized according to their age or gender), rendering emotional expressions task-irrelevant. Thus, although flanker and target processing was obligatory, we did not find any evidence for automatic attentional biases toward emotional stimuli if emotional expressions were unrelated to responses. This finding is in line with findings reported by Barratt and Bundesen (2012, Experiment 2), who also did not find evidence of attentional capture of emotional faces.

General Discussion

The present study contains a series of five flanker experiments that were designed to test whether emotional faces capture attention to a greater extent than neutral faces. When addressing this question in a flanker task, it is important to distinguish between different types of attentional effects and attentional asymmetries that can be computed in this paradigm. We computed stimulus-specific flanker interference effects as a pure measure of attentional capture that is independent of response compatibility. In addition, we also computed stimulus specific response-compatibility effects both for flankers and for targets. Comparing interference and response compatibility effects between emotional and neutral stimuli (or between positive and negative stimuli) yields indices of attentional bias.

Across experiments, a fairly consistent pattern of findings emerged for all types of attentional indices: Across all indices we did not find systematic emotional-neutral or positive-negative asymmetries for Experiments 1, 2, 4a, and 4b. The only study that revealed an emotional-neutral asymmetry (but no positive-negative asymmetry) was Experiment 3. This latter study stands out procedurally because it combined obligatory flanker processing with emotion relevance in the main task. That is, when processing of the emotional content of facial expressions is required to translate stimuli into responses, and when processing of all stimuli in a visual display is made obligatory by an additional target identification task, we do find stronger attentional capture for emotional than for neutral stimuli, and we find this asymmetry consistently across all indicators of attention allocation.

If, however, either of these two features is lacking, then attention allocation is not biased toward faces with emotional expressions. We started with standard flanker tasks in Experiments 1 and 2 in which target faces were identified by their spatial position, which was fixed in the center of the display. Although emotional processing was required by the task in these experiments (target faces had to be categorized based on their emotional expressions), we did not find evidence for a systematic or reliable attentional bias toward emotional faces. Similarly, if flanker processing was made obligatory in Experiments 4a and 4b, but emotional expressions were no longer task relevant, we again found no evidence of an attentional bias toward emotional faces. Our findings thus suggest that for emotional biases to occur in a flanker paradigm, two conditions have to be met, which is task relevance of emotions and obligatory processing of all presented stimuli. This pattern renders emotional/neutral asymmetries in attentional capture highly conditional, depending on task features and processing goals. According to Moors and De Houwer (2006), these findings should thus be classified as conditionally automatic. Our findings thus add to the previous literature in suggesting that attentional bias toward emotional faces is not an unconditionally automatic phenomenon.

Interestingly, Experiment 3 was the only experiment that yielded evidence for robust effects of emotional distraction and of compatibility effects for emotional stimuli, indicating a common source for these two types of effects: Task relevance of emotional features, and obligatory processing of flanker stimuli apparently boost both types of effects. A closer look at the pattern of findings in Experiment 3, however, also reveals a dissociation between emotional distraction (i.e., distraction effects were stronger for emotional compared to neutral flankers) and (nonexistent) emotional modulations of response compatibility (i.e., significant compatibility effects of equal strengths were obtained for emotional and neutral stimuli). This implies that despite sharing a common source, distraction and response compatibility also reflect different processing components. The fact that robust response compatibility effects were obtained for both neutral and emotional stimuli indicates that both types of stimuli drew attention and influenced processes of response selection in a similar way (if anything, compatibility effects were even larger for neutral than for emotional stimuli). The fact that nevertheless, distraction effects were markedly stronger for emotional compared to neutral stimuli in the same experiment indicates that distraction reflects another component of emotional processing that is unrelated to attentional capture and processing fluency, which is best characterized as the "interrupt" function of emotion and emotional stimuli (Simon, 1967; see also McKenna & Sharma, 2004). Emotional interrupt delays responding for emotional compared to neutral trials and thus increases emotional distraction effects. Emotional interrupt does not have an effect on response compatibility effects, however, because it delays responding in both compatible and incompatible trials and is thus not reflected in stronger response compatibility effects for emotional stimuli. This dissociation further highlights the necessity to carefully distinguish between emotional distraction and emotional response compatibility effects in the flanker paradigm, since these effects reflect different components of attention and processing.

The overall pattern of findings that obtained in our study is consistent with what was found with regard to attentional biases for faces with emotional/neutral expressions with other attention paradigms. As our short review of the findings that were gathered with these paradigms revealed (see introduction), there is no reliable evidence for attentional biases for emotional faces in paradigms that do not require emotion processing (dot probe task, flanker interference for tasks that are unrelated to emotion), but there is evidence for emotional influences on attention if the task requires processing of emotional features for all stimuli that are presented in a display (search task).

In evaluating our findings in the context of previous flanker studies using facial expressions of emotion as stimuli, it should be highlighted that our study focused on fairly large pictures of natural faces, which endows our study with a high ecological validity. We were interested in attention to emotion in real-life contexts that only very rarely include schematic faces and also typically does not contain interactions with people whose faces occupy only an extremely small area of the visual field. Our findings and conclusions are thus limited to these conditions, and cannot be generalized to other settings, in which schematic faces are processed or stimuli are presented that are so small that multiple stimuli can be presented simultaneously within the foveal area of vision.

Thus, findings might differ for studies using schematic and natural facial expressions. In this case, however, we would argue that although schematic faces may allow for better control or manipulation of low-level visual features (brightness, contrast, complexity), they may not easily allow for conclusions that generalize to the processing of natural emotional expressions that occur in real life. Besides these differences in physical features or attributes, natural and schematic faces also differ with regard to their complexity. This structural difference could be of major importance when investigating effects of automatic attentional capture. According to the perceptual underload hypothesis, the entire visual field is automatically analyzed as long as there is some spare capacity (see Lavie & Tsal, 1994; Reiner & Morrison, 1983). In situations with natural faces, however, automatic processing of emotional expressions may suffer due to the high levels of complexity that leaves no room for an automatic processing of emotional facial expressions. Although this may explain why attentional effects may be weaker or sometimes even nonexistent in experiments that make use of natural face stimuli, we would argue that nevertheless results of these studies should be considered to be more conclusive with regard to characterizing emotion processing in real social interactions.

Furthermore, no additional task may be needed to render flanker processing obligatory if very small target and flanker stimuli are presented that fall into within the foveal area (regardless of whether these are natural or schematic stimuli). In this regard, it should be noted, however, that we also did not obtain evidence for emotional/neutral asymmetries with regard to the target-related response compatibility effects in Experiments 1 and 2, although target processing is of course obligatory in every task. Thus, strong competition for attention between multiple stimuli might be a necessary condition for obtaining emotional/neutral asymmetries in automatic attention allocation.

A final point that we want to discuss with regard to the absence of significant emotional biases in (most of) our experiments regards the possible lack of reliability of the effect measures. Following recommendations by Parsons, Kruijt, and Fox (2018), we calculated bootstrap-based split-half reliabilities with the help of the R-package splithalf (Parsons, 2018) for all effect measures. Similar to many other response time-based effects (e.g., priming effects), these reliabilities were small and mostly close to zero. Importantly, however, we want to argue that these low reliabilities do not discredit the results that we reported regarding average effects and their significance. The reliability coefficient is an index of the proportion of variance in a variable that is due to systematic individual differences (e.g., Lord & Novick, 1968). Even in case of zero reliability, indicating that all between persons variance is unsystematic or random, the mean of the effect variable can still be highly significant. Exactly this pattern became evident in Experiment 3 for which highly significant distraction and compatibility effects were obtained despite zero reliabilities. In this case, the true effect is of similar strength for all individuals, and differences in the strength of the effect are just unsystematic variance. This situation is not an unlikely scenario for general psychology experiments like ours that employ a sample of normal participants without clinical symptoms who can be assumed to respond similarly to the experimental manipulations, so that variance in the effect variables reflects just random noise. Thus, although the low reliabilities clearly indicate that our measures cannot and should not be used as diagnostic instruments to assess systematic interindividual differences in the strength of the effects (at least not in a homogeneous sample of participants), they definitely do not rule out the possibility to detect general tendencies in emotional processing (this is also the reason why reliabilities are typically not reported in general psychology experiments).

Our finding that attentional biases in processing emotional faces is a strongly conditional phenomenon adds to the literature on the context dependence of social information processing (e.g., Blair, 2002; Casper, Rothermund, & Wentura, 2010, 2011; de Gelder, Meeren, Righart, Van den Stock, van de Riet, & Tamietto, 2006; Gawronski & Cesario, 2013; Lavie, Ro, & Russell, 2003; Müller & Rothermund, 2012; Righart & de Gelder, 2008; Van den Stock & de Gelder, 2014; Van den Stock, Vandenbulcke, Sinke, Goebel, & DeGelder, 2014; Wittenbrink, Judd, & Park, 2001). Contexts differ with regard to all kinds of attributes, including goals, risks, and opportunities, norms, expectations, familiarity, closeness, tasks, rules, standards, social composition, and also physical features (e.g., loudness, brightness, temperature). All of these variables may have implications with regard to attention allocation, but the following may have particular relevance with regard to the processing biases for emotional facial expressions: Emotional processing goals are the default in some situations (e.g., therapy, mother-child interactions, romantic encounters) but not others (e.g., at work, in sports). Contexts also differ with regard to the number of social interaction partners involved, which has immediate implications for selective attention (e.g., persons compete for attention if there is more than one interaction partner, but not if there is only one). Finally, contexts also differ with regard to the frequency, base rates, and composition of emotions that are typically expressed and encountered (e.g., within the family vs. in the supermarket). Some of these variables have already been shown to moderate emotional biases in the flanker task. For instance, Schulte Holthausen, Regenbogen, Turetsky, Schneider, and Habel (2016) showed a context-dependency of flanker effects in terms of emotion category composition. They found no attentional capture of emotional or neutral flanker faces when using fearful, sad, angry, happy, and neutral faces but could show significant attentional capture effects with a more balanced emotion distribution using each one of positive (happy), negative (fearful) and neutral emotion. Based on our findings, we conclude that rather than looking for general emotional biases in attention allocation, future studies should systematically investigate the moderating effects of contextual factors in order to come to an adequate understanding of when and which emotional features become prioritized. Previous findings on affective processing biases already highlight the enormous flexibility and context-dependency of affective processing in general (e.g., counterregulation, Rothermund, Voss, & Wentura, 2008; Schwager & Rothermund, 2013, 2014; Wentura, Voss, & Rothermund, 2009; control-dependency; Rothermund, 2011) and of individual differences (Bar-Haim et al., 2007). We thus propose that similar factors are also likely to affect attention allocation to emotional faces (e.g., Puls & Rothermund, 2018). As has become evident in our study, the flanker paradigm is a highly versatile tool that allows researchers to systematically vary different processing parameters and task requirements, and it contains various options for investigating attentional biases for flankers and targets. We thus recommend this paradigm as a particularly well-suited tool for future research on the context-dependency of affective processing biases.

References

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General, 133,* 323–338. http://dx.doi.org/10.1037/0096-3445.133.3.323

Bargh, J. A., & Gollwitzer, P. M. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behavior. In W. D. Spaulding (Ed.), *Nebraska symposium on motivation: Vol. 41. Integrative views of motivation, cognition, and emotion* (pp. 71–124). Lincoln: University of Nebraska Press.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133,* 1–24. http://dx.doi.org/10.1037/0033-2909.133.1.1

Barratt, D., & Bundesen, C. (2012). Attentional capture by emotional faces is contingent on attentional control settings. *Cognition and Emotion, 26,* 1223–1237. http://dx.doi.org/10.1080/02699931.2011.645279

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242–261. http://dx.doi.org/10.1207/S15327957PSPR0603_8

Casper, C., Rothermund, K., & Wentura, D. (2010). Automatic stereotype activation is context dependent. *Social Psychology, 41,* 131–136. http://dx.doi.org/10.1027/1864-9335/a000019

Casper, C., Rothermund, K., & Wentura, D. (2011). The activation of specific facets of age stereotypes depends on individuating information. *Social Cognition, 29,* 393–414. http://dx.doi.org/10.1521/soco.2011.29.4.393

Chen, S., Yao, N., Qian, M., & Lin, M. (2016). Attentional biases in high social anxiety using a flanker task. *Journal of Behavior Therapy and Experimental Psychiatry, 51,* 27–34. http://dx.doi.org/10.1016/j.jbtep.2015.12.002

Coelho, C. M., Cloete, S., & Wallis, G. (2010). The face-in-the-crowd effect: When angry faces are just cross(es). *Journal of Vision, 10*(1), 14.

de Gelder, B., Meeren, H. K., Righart, R., van den Stock, J., van de Riet, W. A., & Tamietto, M. (2006). Beyond the face: Exploring rapid influences of context on face processing. *Progress in Brain Research, 155,* 37–48. http://dx.doi.org/10.1016/S0079-6123(06)55003-4

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. http://dx.doi.org/10.1037/0033-295X.96.3.433

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics, 63,* 1004–1013. http://dx.doi.org/10.3758/BF03194519

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42,* 351–362. http://dx.doi.org/10.3758/BRM.42.1.351

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143–149. http://dx.doi.org/10.3758/BF03203267

Fenske, M. J., & Eastwood, J. D. (2003). Modulation of focused attention by faces expressing emotion: Evidence from flanker tasks. *Emotion, 3,* 327–343. http://dx.doi.org/10.1037/1528-3542.3.4.327

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion, 14,* 61–92. http://dx.doi.org/10.1080/026999300378996

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin, 134,* 662–676. http://dx.doi.org/10.1037/0033-2909.134.5.662

Garrido, M. I., Barnes, G. R., Sahani, M., & Dolan, R. J. (2012). Functional evidence for a dual route to amygdala. *Current Biology, 22,* 129–134. http://dx.doi.org/10.1016/j.cub.2011.11.056

Gawronski, B., & Cesario, J. (2013). Of mice and men: What animal research can tell us about context effects on automatic responses in humans. *Personality and Social Psychology Review, 17,* 187–215. http://dx.doi.org/10.1177/1088868313480096

Gotlib, I. H., & McCann, C. D. (1984). Construct accessibility and depression: An examination of cognitive and affective factors. *Journal of Personality and Social Psychology, 47,* 427–439. http://dx.doi.org/10.1037/0022-3514.47.2.427

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121,* 480–506. http://dx.doi.org/10.1037/0096-3445.121.4.480

Grose-Fifer, J., Rodrigues, A., Hoover, S., & Zottoli, T. (2013). Attentional capture by emotional faces in adolescence. *Advances in Cognitive Psychology, 9,* 81–91. http://dx.doi.org/10.5709/acp-0134-9

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54,* 917–924. http://dx.doi.org/10.1037/0022-3514.54.6.917

Harms, L., & Bundesen, C. (1983). Color segregation and selective attention in a nonsearch task. *Perception & Psychophysics, 33,* 11–19. http://dx.doi.org/10.3758/BF03205861

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *The Journal of Cognitive Neuroscience, 30,* 11096–11103. http://dx.doi.org/10.1523/JNEUROSCI.1026-10.2010

Hickey, C., Chelazzi, L., & Theeuwes, J. (2011). Reward has a residual impact on target selection in visual search, but not on the suppression of distractors. *Visual Cognition, 19,* 117–128. http://dx.doi.org/10.1080/13506285.2010.503946

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion, 6,* 193–207. http://dx.doi.org/10.1037/1528-3542.6.2.193

Horstmann, G., Borgstedt, K., & Heumann, M. (2006). Flanker effects with faces may depend on perceptual as well as emotional differences. *Emotion, 6,* 28–39. http://dx.doi.org/10.1037/1528-3542.6.1.28

Huang, S. L., Chang, Y. C., & Chen, Y. J. (2011). Task-irrelevant angry faces capture attention in visual search while modulated by resources. *Emotion, 11,* 544–552. http://dx.doi.org/10.1037/a0022763

Juth, P., Lundqvist, D., Karlsson, A., & Öhman, A. (2005). Looking for foes and friends: Perceptual and emotional factors when finding a face in the crowd. *Emotion, 5,* 379–395. http://dx.doi.org/10.1037/1528-3542.5.4.379

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24,* 1377–1388. http://dx.doi.org/10.1080/02699930903485076

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological Science, 14,* 510–515. http://dx.doi.org/10.1111/1467-9280.03453

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56,* 183–197. http://dx.doi.org/10.3758/BF03213897

Liu, X., Qian, M., Zhou, X., & Wang, A. (2006). Repeating the stimulus exposure to investigate what happens after initial selective attention to threatening pictures. *Personality and Individual Differences, 40,* 1007–1016. http://dx.doi.org/10.1016/j.paid.2005.10.006

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition, 7,* 166–174. http://dx.doi.org/10.3758/BF03197535

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Menlo Park, CA: Addison Wesley.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95,* 15–20. http://dx.doi.org/10.1037/0021-843X.95.1.15

McKenna, F. P., & Sharma, D. (2004). Reversing the emotional Stroop effect reveals that it is not what it seems: The role of fast and slow components. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 382–392. http://dx.doi.org/10.1037/0278-7393.30.2.382

Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception & Psychophysics, 49,* 270–288. http://dx.doi.org/10.3758/BF03214311

Moeller, B., & Frings, C. (2014). Attention meets binding: Only attended distractors are used for the retrieval of event files. *Perception & Psychophysics, 76,* 959–978. http://dx.doi.org/10.3758/s13414-014-0648-9

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132,* 297–326. http://dx.doi.org/10.1037/0033-2909.132.2.297

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393,* 467–470. http://dx.doi.org/10.1038/30976

Müller, F., & Rothermund, K. (2012). Talking loudly but lazing at work—Behavioral effects of stereotypes are context dependent. *European Journal of Social Psychology, 42,* 557–563. http://dx.doi.org/10.1002/ejsp.1869

Müller, S., Rothermund, K., & Wentura, D. (2016). Relevance drives attention: Attentional bias for gain- and loss-related stimuli is driven by delayed disengagement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 69,* 752–763. http://dx.doi.org/10.1080/17470218.2015.1049624

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80,* 381–396. http://dx.doi.org/10.1037/0022-3514.80.3.381

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108,* 483–522. http://dx.doi.org/10.1037/0033-295X.108.3.483

Parsons, S. (2018). *Splithalf: Robust estimates of split half reliability* (Version 3). http://dx.doi.org/10.6084/m9.figshare.5559175.v3

Parsons, S., Kruijt, A. W., & Fox, E. (2018). Psychological science needs a standard practice of reporting the reliability of cognitive behavioural measurements. http://dx.doi.org/10.31234/osf.io/6ka9z

Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods, 162,* 8–13. http://dx.doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*. Advance online publication. http://dx.doi.org/10.3389/neuro.11.010.2008

Puls, S., & Rothermund, K. (2018). Attending to emotional expressions: No evidence for automatic capture in the dot-probe task. *Cognition and Emotion, 32,* 450–463.

Purcell, D. G., & Stewart, A. L. (2010). Still another confounded face in the crowd. *Perception & Psychophysics, 72,* 2115–2127. http://dx.doi.org/10.3758/BF03196688

Purcell, D. G., Stewart, A. L., & Skov, R. B. (1996). It takes a confounded face to pop out of a crowd. *Perception, 25,* 1091–1108. http://dx.doi.org/10.1068/p251091

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35,* 741–756. http://dx.doi.org/10.1016/S0005-7967(97)00022-3

Reiner, M. B., & Morrison, F. J. (1983). Is semantic interference really automatic? *Bulletin of the Psychonomic Society, 21,* 271–274. http://dx.doi.org/10.3758/BF03334707

Righart, R., & De Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience, 8,* 264–272. http://dx.doi.org/10.3758/CABN.8.3.264

Rothermund, K. (2011). Counter-regulation and control-dependency: Affective processing biases in the service of action regulation. *Social Psychology, 42,* 56–66. http://dx.doi.org/10.1027/1864-9335/a000043

Rothermund, K., Voss, A., & Wentura, D. (2008). Counter-regulation in affective attentional biases: A basic mechanism that warrants flexibility in emotion and motivation. *Emotion, 8,* 34–46. http://dx.doi.org/10.1037/1528-3542.8.1.34

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality, 19,* 595–605. http://dx.doi.org/10.1002/per.554

Schulte Holthausen, B., Regenbogen, C., Turetsky, B. I., Schneider, F., & Habel, U. (2016). The influence of task-irrelevant flankers depends on the composition of emotion categories. *Frontiers in Psychology*. Advance online publication. http://dx.doi.org/10.3389/fpsyg.2016.00712

Schupp, H. T., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion, 4,* 189–200. http://dx.doi.org/10.1037/1528-3542.4.2.189

Schwager, S., & Rothermund, K. (2013). Counter-regulation triggered by emotions: Positive/negative affective states elicit opposite valence biases in affective processing. *Cognition and Emotion, 27,* 839–855. http://dx.doi.org/10.1080/02699931.2012.750599

Schwager, S., & Rothermund, K. (2014). On the dynamics of implicit emotion regulation: Counter-regulation after remembering events of high but not of low emotional intensity. *Cognition and Emotion, 28,* 971–992. http://dx.doi.org/10.1080/02699931.2013.866074

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review, 74,* 29–39. http://dx.doi.org/10.1037/h0024127

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology, 54,* 44–53. http://dx.doi.org/10.1027/1618-3169.54.1.44

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51,* 599–606. http://dx.doi.org/10.3758/BF03211656

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The eyebrow frown: A salient social signal. *Emotion, 2,* 288–296. http://dx.doi.org/10.1037/1528-3542.2.3.288

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. http://dx.doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114,* 285–310. http://dx.doi.org/10.1037/0096-3445.114.3.285

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.

Van den Stock, J., & de Gelder, B. (2014). Face identity matching is influenced by emotions conveyed by face and body. *Frontiers in Human Neuroscience*. Advance online publication. http://dx.doi.org/10.3389/fnhum.2014.00053

Van den Stock, J., Vandenbulcke, M., Sinke, C. B., Goebel, R., & de Gelder, B. (2014). How affective information from faces and scenes interacts in the brain. *Social Cognitive and Affective Neuroscience, 9,* 1481–1488. http://dx.doi.org/10.1093/scan/nst138

Wentura, D., Müller, P., & Rothermund, K. (2014). Attentional capture by evaluative stimuli: Gain- and loss-connoting colors boost the additional-singleton effect. *Psychonomic Bulletin & Review, 21,* 701–707. http://dx.doi.org/10.3758/s13423-013-0531-z

Wentura, D., Voss, A., & Rothermund, K. (2009). Playing TETRIS for science counter-regulatory affective processing in a motivationally "hot" context. *Acta Psychologica, 131,* 171–177. http://dx.doi.org/10.1016/j.actpsy.2009.05.008

Williams, J. M. G., & Nulty, D. D. (1986). Construct accessibility, depression and the emotional stroop task: Transient mood or stable structure? *Personality and Individual Differences, 7,* 485–491. http://dx.doi.org/10.1016/0191-8869(86)90127-3

Williams, M., Moss, S., Bradshaw, J., & Mattingley, J. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition, 12,* 29–50. http://dx.doi.org/10.1080/13506280444000193

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815–827. http://dx.doi.org/10.1037/0022-3514.81.5.815

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion, 24,* 3–47. http://dx.doi.org/10.1080/02699930903205698

(*Appendix follows*)

Appendix

Bayes Factors

Table A1

Flanker Distraction Effects: Bayes Factors (BF) for the Reaction Time (RT) and Error Data (in %errors), Computed as the Difference in Average RTs and %errors Between Emotional and Neutral Flanker Conditions

	Emotional vs. neutral		Angry vs. neutral		Fearful vs. neutral		Happy vs. neutral	
Exp.	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}
RT								
E1	.774	1.292	1.194	.837	.343	2.919	—	—
E2	.772	1.296	.241	4.152	2.644	.378	—	—
E3	32.474	.031	17.066	.059	40.794	.025	1.771	.565
E4a	.171	5.861	.189	5.299	.18	5.555	.175	5.708
E4b	.433	2.309	.369	2.709	.334	2.998	—	—
ERR								
E1	1.002	.998	1.1	.909	.485	2.062	—	—
E2	.483	2.07	.202	4.946	.862	1.159	—	—
E3	.175	5.708	.169	5.929	.19	5.271	.173	5.778
E4a	.192	5.204	.203	4.933	.178	5.621	1.78	5.629
E4b	.224	4.472	.201	4.973	.235	4.264	—	—

Table A2

Flanker-Related Compatibility Effects: Bayes Factors (BF) for Reaction Time (RT) and Error Data, Based on the Difference in Average RTs (in Ms) and %errors Between Compatible and Incompatible Trials Within the Respective Flanker Condition (Positive Values Indicate Facilitation in Compatible Compared to Incompatible Trials)

	Overall		Angry		Fearful		Happy		Neutral	
Experiment (E)	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}
RT										
E1	.336	2.734	.194	5.151	.245	4.078	—	—	.368	2.715
E2	.202	2.734	.173	5.789	.183	5.461	—	—	.202	4.953
E3	3222.044	3.104e−4	.982	1.018	24.736	.04	.283	3.534	18.888	.053
E4a	3.761e+12	2.659e−13	2.199e+6	4.548e−7	1549	6.454e−4	7.583e+6	1.319e−7	10671	9.371e−5
E4b	2.854e+9	3.504e−10	3.418e+7	2.926e−8	10353	9.659e−5	—	—	5.731e+6	1.745e−7
ERR										
E1	.178	5.612	.291	4.557	.206	4.851	—	—	.211	4.742
E2	.285	3.514	.254	3.943	.171	5.843	—	—	.285	3.514
E3	9.498	.105	.624	1.604	.728	1.375	1.37	.73	6.051	.165
E4a	51.233	.02	11.718	.085	252.864	.004	6.732	.149	1	1
E4b	1.220e+6	8.198e−7	22501	4.444e−5	27517	3.634e−5	—	—	3634	2.752e−4

Note. Main effects of target emotion were controlled before computing flanker-related compatibility Bayes factors in Exp.'s 1, 2, and 3 (in Exp.'s 4a and 4b, response compatibility refers to age and gender, respectively, so flanker-related compatibility effects are not confounded with target emotion).

(*Appendix continues*)

Table A3
Target-Related Compatibility Effects: Bayes Factors (BF) for the Reaction Time (RT) RT and Error Data, Based on the Difference in Average RTs (in Ms) and %errors Between Compatible and Incompatible Trials Within the Respective Target Condition

	Overall		Angry		Fearful		Happy		Neutral	
Experiment (E)	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}
RT										
E1	.366	2.734	.275	3.636	.273	3.665	—	—	.169	5.923
E2	.27	3.708	.446	2.24	.226	4.422	—	—	.229	4.365
E3	3222.04	3.104e−4	.243	4.118	1.031	.97	.234	4.273	6538.151	1.529e−4
E4a	3.761e+12	2.695e−13	82020.05	1.219e−5	32343.4	3.092e−5	15.42	.065	2.445e+6	4.091e−7
E4b	2.854e+9	3.504e−10	8996	1.112e−4	1.372e+6	7.2894–7	—	—	6.056e+6	1.651e−7
ERR										
E1	.178	5.612	.318	3.148	.169	5.901	—	—	.181	5.516
E2	.421	2.375	.176	5.572	8.817	5.672	—	—	.171	.113
E3	9.498	.105	.462	2.163	.887	1.128	1.168	.856	1.037	.964
E4a	51.233	.02	2.032	.492	6.316	.158	6.047	.165	182.609	.005
E4b	1220e+6	8.198e−7	5773.9	1.732e−4	563.2	.002	—	—	2470.5	4.048e−4

Received November 22, 2017
Revision received September 11, 2018
Accepted September 14, 2018 ■