Testing the social processing mode hypothesis with the *Attentional Response to Distal versus Proximal Emotional Information* (ARDPEI) task. Evidence for moderations of engagement and disengagement processes

Dirk Wentura[1], Siddhima Gupta[1,] and Benedikt Emanuel Wirth[2]

[1]Saarland University, Saarbrücken, Germany

[2]German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany

Dirk Wentura: https://orcid.org/ 0000-0001-9907-498X

Siddhima Gupta https://orcid.org/xxxxxxxxxxxxx

Benedikt Emanuel Wirth: https://orcid.org/0000-0001-8039-7316

Author Note

This research was supported by a grant from the German Research Foundation to Dirk Wentura and Benedikt Wirth (WE 2284/16-1; WI 5758/1-1). Correspondence concerning this article should be addressed to Dirk Wentura, Campus A2 4, D-66123 Saarbrücken, Germany. E-mail: wentura@mx.uni-saarland.de. Phone: (+49) 681-302 4642. Fax: (+49) 681-302 4049. ORCID: 0000-0001-9907-498X

Abstract

Attentional biases for angry versus neutral faces were assessed with the Attentional Response to Distal versus Proximal Emotional Information task (ARDPEI task; Grafton & MacLeod, 2014), which is suited to disentangle attentional capture and attentional dwelling processes. Participants performed a task involving classification of either socially meaningful targets (schematic faces; i.e., social target mode) or meaningless targets (scrambled schematic faces; i.e., non-social target mode). Engagement towards angry faces (relative to neutral faces) positively correlated with trait anxiety, despite the brief cue-target onset asynchrony (CTOA) of 100 ms. (Grafton and MacLeod found this relationship only for a CTOA of 500 ms.) Interestingly, the correlation was restricted to the non-social target mode. In contrast to recent studies with the dot-probe task, there was no attentional capture by angry faces in the social target mode. In the social target mode, increased attentional dwelling on neutral faces (relative to angry faces) was found.

Key words: Attentional bias; dot-probe task; engagement; disengagement; angry faces; social processing mode

Testing the social processing mode hypothesis with the *Attentional Response to Distal versus Proximal Emotional Information* (ARDPEI) task.

Evidence for moderations of engagement and disengagement processes

Attentional biases toward threats are a longstanding topic in psychological research. The field is characterized by two incompatible perspectives. Experimental psychopathology assumes that only highly anxious individuals will show attentional biases toward threat (e.g., Mogg & Bradley, 1998; Williams et al., 1988). In a meta-analysis, Bar-Haim et al. (2007) found results that corroborate this perspective: Biases were found in participants with high trait anxiety, but not in those with low trait anxiety. On the other hand, basic research on the interplay of affective and cognitive processes assumes that such biases are present in general, that is, in non-selected samples. For instance, Cooper and Langton (2006) found a bias toward angry faces in a sample not selected for anxiety. In a recent series of experiments using the dot-probe task, Wirth and Wentura (2018a, 2019, 2023) found a bias toward angry faces in the general population. However, this bias depended on what the authors termed 'social processing mode'. In short, it refers to whether participants' attention toward threat is tested with a task that necessitates social processing of stimuli. We will elaborate on this below.

Given these results, a potential link between the two perspectives is the assumption that trait anxious individuals are permanently in this 'social processing mode' and, hence, show attentional biases towards threat even in experimental settings in which this mode is not triggered by an experimental manipulation.

With this basic idea in mind we set up an experiment that uses a further “flagship” paradigm of the field (beyond dot-probe), which broadly supported the individual differences perspective (i.e., that trait anxiety moderate an attentional bias towards threat), but which was yet not tested with the experimental manipulation to induce a social or non-social processing mode: the Attentional Response to Distal vs. Proximal Emotional Information (ARDPEI) paradigm, introduced by Grafton and MacLeod (2014).

The ARDPEI paradigm

Grafton and MacLeod (2014) developed the Attentional Response to Distal vs. Proximal Emotional Information (ARDPEI) paradigm. They aimed to determine the nature of the attentional bias to which individuals with elevated anxiety are prone. Is it an enhanced attentional capture by threat stimuli? Or is it enhanced dwelling on these stimuli that characterizes highly anxious individuals?

In this paradigm, participants must intentionally focus on a red square that appears for 1,000 milliseconds to the right or left of the center (see Figure 1). Then, a line (the anchor probe) is presented within this square for 150 ms; participants are expected to memorize the orientation of this line (i.e., horizontal versus vertical). Next, an angry or neutral face and an abstract image are presented simultaneously on either side of the center for 100 or 500 ms. Next, the target – again, a horizontal or vertical line – is presented to the left or right of the center with equal frequency. Thus the cue-target onset asynchrony (CTOA) was either 100 or 500 ms in the study by Grafton and MacLeod (2014). Participants must identify whether the target line's orientation matches that of the anchor probe.

Assessing engagement

Attention engagement is assessed in trials in which the face appears on the opposite side of the anchor probe (distal trials). If the face does *not* capture attention, responses will be fast in trials in which the target appears on the same side as the anchor probe (since attention is still there) and therefore on the opposite side of the face (invalid trials).[1] Responses will be slow in trials where the target appears on the side opposite the anchor probe (since attention

[1] The terms "valid" and "invalid" are always used to characterize the match or non-match of face and target locations.

is not there) and therefore on the same side as the face (valid trials). Any trial in which the face captures attention will reduce the average reaction time (RT) for valid trials because attention will already be where the target appears. If angry faces are more likely to capture attention than neutral faces, then the reduction in mean RT for valid trials will be greater for angry faces. Thus, an engagement bias (EB) can be calculated using the following formula (see Grafton & MacLeod, 2014):

$$\text{Engagement Bias (EB)} = (RT_{\text{valid, neutral}} - RT_{\text{invalid, neutral}}) - (RT_{\text{valid, angry}} - RT_{\text{invalid, angry}})$$

The EB will take on positive values if the capture rate for angry faces is higher than for neutral faces. It will take on negative values if the capture rate for neutral faces is higher. Note that this formula only refers to distal trials, that is, trials in which the face is presented at the opposite location from the anchor probe.

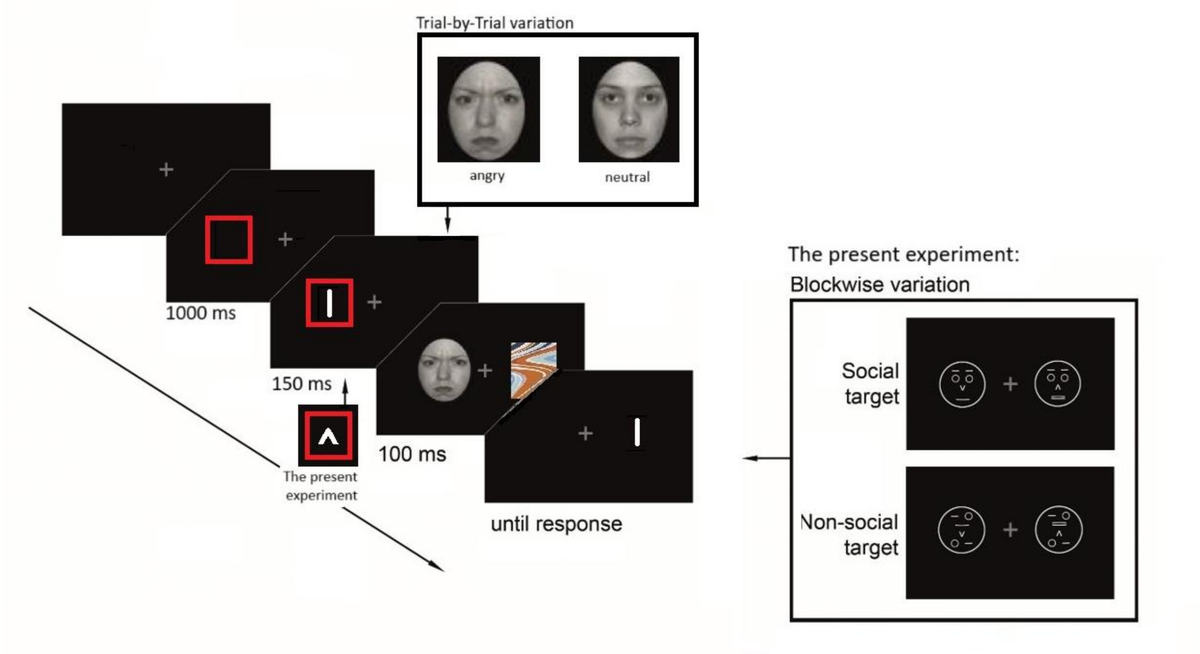

Figure 1. The trial sequence of the ARDPEI task (Grafton & MacLeod, 2014); see text for further explanation. Here a proximal invalid trial is shown: the face appeared at the location of the initial attention; the target is on the other side of the face. In the present version, two stimuli appear on the target screen. Participants have to search for the target (i.e., the face with the open mouth in the social target condition and the stimulus with the double horizontal line in the non-social condition), and to categorize the angle within the target (called "nose" in the social condition) as to whether it matches the anchor probe or not.

Assessing disengagement

Problems with disengaging attention are assessed using trials in which the face appears on the same side as the anchor probe (proximal trials). Valid trials (i.e., trials with targets appearing at the same location as the face) are generally faster because the target appears at the initial location of attention. In contrast, invalid trials (i.e., trials with targets appearing at the opposite location from the face) are generally slower because attention must shift to the opposite location. Attention dwelling on a face stimulus prolongs this shift of attention. Thus, a disengagement bias (EB) can be calculated using the following formula (see Grafton & MacLeod, 2014):

$$\text{Disengagement Bias (DB)} = (RT_{\text{invalid, angry}} - RT_{\text{valid, angry}}) - (RT_{\text{invalid, neutral}} - RT_{\text{valid, neutral}})$$

The DB will take on positive values if disengaging from angry faces is more difficult than disengaging from neutral faces. It will take on negative values if the opposite is true. Note that this formula only refers to proximal trials, that is, trials in which the face is presented at the same location as the anchor probe.

Grafton and MacLeod (2014) recruited participants with low and high trait anxiety (i.e., participants from the lower and upper thirds of a larger sample of first-year students) and found the following results. For the 100 ms CTOA condition, they found a numerically positive and nearly equal engagement bias for both groups, which, however, did not significantly differ from zero. The disengagement index (again at 100 ms CTOA) differed significantly between the two groups, with a positive mean for high-trait anxious participants (i.e., these participants dwelled more on angry than neutral faces), and a negative mean for low-anxious participants (i.e., these participants dwelled more on neutral than angry faces).

As mentioned, Grafton and MacLeod (2014) also used a 500-ms CTOA. They found that trait anxiety moderated engagement bias, with high-anxious participants having a positive mean (i.e., more capture by angry than neutral faces) and low-anxious participants having a negative mean (i.e., more capture by neutral than angry faces).

Meanwhile, this paradigm was applied to other research questions that require studying engagement and disengagement. Published research covers ruminative brooding (Grafton et al., 2016; Southworth et al., 2017), intrusive re-experiencing of negative thoughts (Dondzilo et al., 2022), social anxiety (Grafton & MacLeod, 2016), and ideal body image types (Dondzilo et al., 2018, 2021).

As already mentioned above, we wanted to introduce a further experimental manipulation – that is, the social processing mode manipulation – into the ARDPEI paradigm. Therefore, we will give a brief summary of previous research that has used this manipulation.

Dot-probe research and the social processing mode

In the conventional dot-probe task, participants are instructed to categorize targets that appear randomly on one side of the center of the screen. For example, they must categorize two vertically or horizontally arranged dots accordingly. The target screen is preceded by a cue screen composed of two faces, one on each side of the center: one angry face and one neutral face. The locations of the angry face and the target either match or do not match; the probability of a match is $p = .5$. Trials in which the target is presented at the location of the angry face are called valid trials, while trials in which the target is presented at the location of the neutral face are called invalid trials. The cue-target onset asynchrony (CTOA) is typically in the range of 100 to 500 milliseconds. In very short CTOA conditions (i.e., 100 ms) especially, faster responses in the valid condition than in the invalid condition are dominantly interpreted as attentional capture by the angry face.

Given this basic setting, Wirth and Wentura (2018a) built upon the concept of contingent capture (Folk et al., 1992), which is well-known in the field of basic research on attention. For instance, a salient red color cue – that is, the cue screen consists of four stimuli in different locations, three gray and one red – only captures attention when participants must search for a red *target* among gray distractors. However, if the target screen contains only a single gray target stimulus without competitors (i.e., an "abrupt onset" target), no color cueing will be observed. Thus, attentional capture is contingent on appropriate goal settings.

Wirth and Wentura (2018a) wondered (p. 977): "… we know from spatial-cueing studies that attentional control settings can affect which stimuli capture attention, and we know that in dot-probe studies – which can be regarded as variants of spatial cueing – we often find an attentional bias towards threat stimuli in anxious participants but only rarely in non-anxious participants. Is it therefore possible that this attentional bias depends on an attentional control setting that is only occasionally tuned to threat in non-anxious individuals (but permanently in anxious individuals)?"

In a series of experiments, Wirth and Wentura (2018a, 2019, 2023) tested this basic assumption. In short, they found that an attentional bias toward angry faces occurs when participants are required to process social information as opposed to non-social information. Their task differs from conventional emotional dot probe tasks in that both the cue and target displays present stimuli to the left and right of the fixation cross. In the social mode, two schematic faces were presented on the target screen: one with a single line for the mouth and one with a double line ("open mouth"). Participants were instructed to categorize whether the nose of the "open mouth" face pointed up (^) or down (v). In the non-social mode, participants saw scrambled versions of the two schematic faces. That is, the elements that made up a face were randomly rearranged to create an abstract, meaningless pattern. Participants were instructed to categorize the arrow-like element of the pattern, which included a double horizontal line instead of a single horizontal line. The cueing effect for angry faces only occurred in the social target block, not the non-social one. This finding was interpreted to mean that participants were actively processing social information in the social block, resulting in a bias toward angry faces.

Wirth and Wentura (2018a, 2019) used a CTOA of 100 ms to increase the probability that attentional capture by the angry face is the decisive process behind the reported effects. Long CTOAs (as 500 ms) are more ambiguous since these delays allow for several shifts of attention (see also Cooper & Langton, 2006), moreover gaze movements are possible within this time range, with consequences for the measured effects (Petrova et al., 2013).

Nevertheless, one might argue that attention goes to one of the cue faces on a random base and that it depends on the valence of the attended face whether attention can be easily shift from this location in case that the target appears at the invalid position. It might be the case that attention cannot be easily disengaged from angry faces (Fox et al., 2001). Recently, Wirth and Wentura (2023) conducted an EEG experiment to disentangle between capture and disengagement in the dot-probe experiments using the N2pc event-related potential as an index of attentional shifts. Indeed, a N2pc component locked to angry face cues was found; it was – fitting the social processing mode hypothesis – significantly larger in the social target condition than in the non-social target condition. These results suggest that the dot-probe effects found by the authors indicates initial attention toward angry faces.

The ARDPEI paradigm is a different tool to disentangle engagement and disengagement processes. However, the ARDPEI paradigm cannot simply be understood as a means of breaking down the processes driving dot-probe effects into their components (i.e., engagement and disengagement). It is a paradigm on its own. For example, the dot-probe task is characterized by a competition between two types of cue stimuli that should be – in the ideal case – different on only one specific feature, for example, the negative valence in the case of angry versus neutral faces. In the ARDPEI paradigm this competition is indirect: which of the two stimuli – that is, angry or neutral face – causes stronger processes in contexts that are, on the one hand, absolutely equal for the two types, but, on the other hand, not well defined. This holds especially for the distal trials (i.e., trials where the initial attention is on the abstract image and the face is presented on the opposite side): The attentional capture potential of angry and neutral faces is tested against unknown attention-holding properties of the abstract images. Nevertheless, it is worth to explore whether the social processing mode hypothesis can be successfully transferred from the dot-probe paradigm to the ARDPEI task.

Overview

In our experiment we use the same trial structure as the one in Grafton and MacLeod (2014). The most important difference is the introduction of a search target display as in Wirth and Wentura (2019, 2023) and the factor target type (social versus non-social; see Figure 1). Accordingly, now the anchor probe is an angle (shaped like the nose of the schematic faces) either facing up or down (∧ or ∨). In accordance with our earlier research, we used only the 100 ms CTOA condition.

If the activation of a social processing mode affects attentional engagement with angry faces, we expect to find a significant difference in the engagement bias between the social and the non-social block of the experiment. Conversely, if the activation of a social processing mode affects attentional disengagement from angry faces, we expect to find a significant difference in the disengagement bias between the social and the non-social block of the experiment. It is also possible, that the social processing mode affects both engagement and disengagement (which would be reflected by a difference between social and non-social trial's for both bias indices).

As in our earlier research, our main focus were on general effects, that is, not on individual differences. Actually, we routinely assessed trait anxiety in our previous experiments (Wirth & Wentura, 2018a, 2019, 2023), but never found significant moderations of dot-probe effects by trait anxiety (but see Wirth & Wentura, 2018b). Nevertheless, we added the trait anxiety scale of the STAI (Spielberger et al., 1983) in our ARDPEI experiment to assess individual differences in trait anxiety and to directly relate our study to Grafton and MacLeod (2014).

Method

Transparency and Openness

We report how we determined our sample size, any data exclusions, all manipulations, and all measures in this study. The data and the program code for all experiments are available on the Open Science Framework (OSF). These files can be accessed via the following link: xxxxxxxxx We pre-registered our experiments on aspredicted.org. The preregistration can be accessed via the following link: https://aspredicted.org/9ZL_9SR). The study was approved by the Ethics Committee of the Faculty of Human and Business Sciences, Saarland University, Saarbrücken, Germany.

Participants

The final analysis included $N = 124$ participants of which 36 were women, 84 were men, and 3 were diverse; their ages ranged from 18-35 years ($M = 29.26$ years, $SD = 4.2$ years). A total of 160 participants were recruited via the online recruitment platform Prolific (prolific.co, Prolific Academic Ltd. London, England). Thirty-six participants were excluded from the final analysis because their overall accuracy fell below 70% (see preregistration) All participants provided informed consent prior to testing.

The focus of power planning was on the two bias indices (engagement and disengagement) that we would compare between the social and social target trials. We expected significant biases only in the social condition and not in the non-social condition. According to the social-processing hypothesis, we expected to find neither significant allocation scores nor significant disengagement scores on trials with socially meaningless targets. However, on trials with socially meaningful targets, we expect to find significant capture scores and/or disengagement scores. The average effect size of the bias towards angry faces in our previous studies was $dz = 0.30$. According to GPower (Faul et al, 2007), to detect an effect of size $d_Z = 0.25$ (that is, an effect that is a bit attenuated relative to our former results), given an error of α =.05 with a power of 1-β = .8, we needed a sample size of 128 participants. We slightly deviated from our preregistration (“If we have to exclude any participants …, we will recruit new participants until the sample size of N = 128 is reached.”) by stopping at N=124. With N=124, still an effect of size $d_Z = 0.254$ can be detected with a power of 1-β = .8 (α =.05).

Design

We used a 2 (target type: social vs. non-social) × 2 (initial attention focus: proximal vs. distal to face cue) × 2 (cue emotion: angry vs. neutral) × 2 (cue validity: valid vs. invalid face cue) within-participants design.

Materials

We used faces of eight female and eight male individuals from the Chicago Face data base (Ma et al., 2015). We used both their neutral and closed mouth angry expressions. Dot-probe experiments have shown that exposed teeth are a strong perceptual confound in happy expressions (Wirth & Wentura, 2018b), so we employed only angry faces with closed mouths in the present study. All stimuli were cropped into a standard oval shape concealing hair and external features and were converted to greyscale. The abstract images were taken from Grafton and MacLeod (2014).

Procedure

The study was conducted online. The experimental routine was created using PsychoJS, the JavaScript counterpart to PsychoPy (Peirce et al., 2022) and hosted on Pavlovia (pavlovia.org, Open Science Tools Ltd., Nottingham, England). After agreeing to participate in the study on Prolific, participants were automatically transferred to Pavlovia and the experiment started in their Browser (participation was only allowed on a desktop computer or laptop). To adjust presentation parameters to the actual screen size, participants were asked to resize a credit card image (presented on the screen) to the size of a real credit card (or equivalent) by using the arrow buttons on their keyboard (Morys-Carter, 2021, May 18). After completing the informed consent form, they were shown a screen with instructions on the experimental procedure.

Each trial of the ARDPEI paradigm started by a fixation cross at the center of the screen for xxx ms. Then a red square (x.x × x.x cm) appeared for 1,000 milliseconds to the right or left of the center (see Figure 1); participants had to intentionally focus on the square. Then, the anchor probe (∧ or ∨; x.x × x.x cm) is presented within this square for 150 ms; participants are expected to memorize the orientation (i.e., up versus down). Next, an angry or neutral face and an abstract image were presented simultaneously on either side of the center for 100 ms. Both had a size of x.x × x.x cm and the center-to-center distance between them was xx.x cm. Next, the target was presented to the left or right of the center with equal frequency until response. The schematic faces / scrambled faces had a size of 2.8 × 2.8 cm and the center-to-center distance between them was 11.1 cm. Participants had to identify whether the orientation of the target's arrow-like element matches that of the anchor probe. In the social target type condition, two schematic faces appeared on the screen, one with a single line for the mouth and one with a double line ("open mouth") with the latter one being the target. Accordingly, participants were instructed to categorize the nose of the "open mouth" face as either matching the direction of the anchor probe (by pressing key 't') or not (by pressing key 'v'). In the non-social target type condition, two scrambles faces (see Figure 1) appeared on the screen, one with a single horizontal line and one with a double line with the latter one being the target. Accordingly, participants were instructed to indicate whether the arrow in the double-line pattern matched the direction of the anchor probe or not. The intertrial interval was xxxx ms.

The experiment comprised 512 trials and lasted approximately xx minutes. Trials were presented in two blocks consisting of 256 trials each – one with schematic faces as target and distractor stimuli and one with scrambled faces as target and distractor stimuli – in a counterbalanced order. Xxxx explanation of the number, e.g.: Anchor position, face position, target position, face emotion, and face identity were fully crossed, which resulted in N = (2×2×2×2×16 =) 256 trials for each of the two blocks. Direction of anchor probe and target "nose"/arrow were randomly selected. Within each block, a self-paced break was included after 128 trials. At the beginning of each block, participants were presented with xx training trials that were not included in data analysis. At the end of the experiment, participants completed the trait-anxiety scale of the STAI (Spielberger et al., 1983). This self-assessment scale contains 20 items (e.g., "I worry too much over something that really doesn't matter"), all rated on a 4-point scale (e.g., from "Almost Never" to "Almost Always").

Results

For the RT analysis, only trials with correct responses were considered. The average classification accuracy was 91.5% (*SD* = 6.9%). RTs below 150 ms were excluded, as were RTs more than 1.5 interquartile ranges above the third quartile of the individual participant's distribution (separately for both experimental conditions, i.e., social target vs. non-social target; Tukey, 1977). This procedure led to the exclusion of 3.1% of all trials with correct responses. Table 1 shows the mean RTs.

Table 1

Response Times (in ms; Accuracy Rates in Parentheses; in %) for Engagement and Disengagement Biases for both Social and Non-social Trials

		Attentional focus			
		Proximal		Distal	
		Cue Validity		Cue Validity	
Target Type	Emotion	Valid	Invalid	Valid	Invalid
Social	Angry	839 (.93)	873 (.90)	876 (.90)	836 (.93)
	Neutral	836 (.92)	882 (.89)	877 (.90)	836 (.93)
Non-social	Angry	825 (.93)	873 (.90)	874 (.90)	834 (.93)
	Neutral	826 (.94)	880 (.90)	876 (.91)	831 (.93)

Planned analyses

Response Times

We conducted a 2 (target type: social vs. non-social) × 2 (initial attention focus: proximal vs. distal to face cue) × 2 (cue emotion: angry vs. neutral) × cue validity (valid vs. invalid face cue) repeated-measures ANOVA with mean RTs as dependent variable. Table 2 (left column) shows the results. The analysis yielded two significant effects: The two-way interaction of initial attention focus and cue validity was significant. This interaction is rather trivial: In the proximal valid and the distal invalid conditions the target appeared at the location of the initial focus; hence, in general responses are faster (and less error prone, see below). However, this interaction is further moderated by emotion.

The meaning of this effect becomes clearer in the planned analyses of the indices. In line with Grafton and MacLeod (2014) and our preregistration, we focused our further analyses on the engagement and disengagement indices. Figure 2 shows the indices split for social and non-social target conditions. Both mean engagement indices are not significantly different from zero, both $|t|$s < 1. The disengagement index for social targets is significantly below zero, $t(123) = 2.02$, $p = .046$, $d_Z = 0.18$, whereas the index for non-social targets is not, $t(123) = 1.15$, $p = .251$, $d_Z = 0.10$. However, the two indices are not significantly different, $|t| < 1$.

Table 2

Results of the Target type × Initial Attention Focus × Emotion × Cue Validity ANOVA with mean RTs as Dependent Variable

	Effect			× Trait Anxiety		
Effekt	F(1, 123)	p	η_p^2	F(1, 122)	p	η_p^2
Attention focus (A)	< 1			2.03	0.156	0.016
Emotion (E)	1.93	0.167	0.015	1.58	0.211	0.013
Target Type (T)	< 1			2.84	0.094	0.023
Cue Validity (C)	1.18	0.279	0.010	< 1		
A × E	1.81	0.181	0.014	< 1		
A × T	2.69	0.103	0.021	< 1		
A × C	56.73	0.000	0.316	< 1		
E × T	< 1			< 1		
E × C	1.11	0.294	0.009	3.78	0.054	0.030
T × C	1.67	0.198	0.013	< 1		
A × E × T	< 1			< 1		
A × E × C	4.16	0.043	0.033	2.04	0.156	0.016
A × T × C	1.02	0.315	0.008	< 1		
E × T × C	< 1			7.14	0.009	0.055
A × E × T × C	< 1			< 1		

Note: The columns with the headers "Effect" refer to a 2 (target type: social vs. non-social) × 2 (initial attention focus: proximal vs. distal to face cue) × 2 (cue emotion: angry vs. neutral) × cue validity (valid vs. invalid face cue) repeated measures ANOVA. The columns with the header "× Trait Anxiety" refers to the same analysis with z-standardized trait anxiety added as a covariate.

The latter result fits to the fact that in the overall ANOVA only the triple interaction *initial attention focus* × *cue emotion* × *cue validity* was significant (see above), but not the four-way interaction. To relate analyses of indices to the ANOVA results: The mean disengagement index (collapsed over target type) is significantly below zero as well, $t(123) = 2.14$, $p = .034$, $d_Z = 0.19$; the mean engagement index is not significant, $|t| < 1$. The difference between mean disengagement index and mean engagement index is significant, $t(123) = 2.04$, $p = .043$, $d_Z = 0.18$. Note: This t-test result is equivalent to the triple interaction reported above (with $t = \sqrt{F}$).

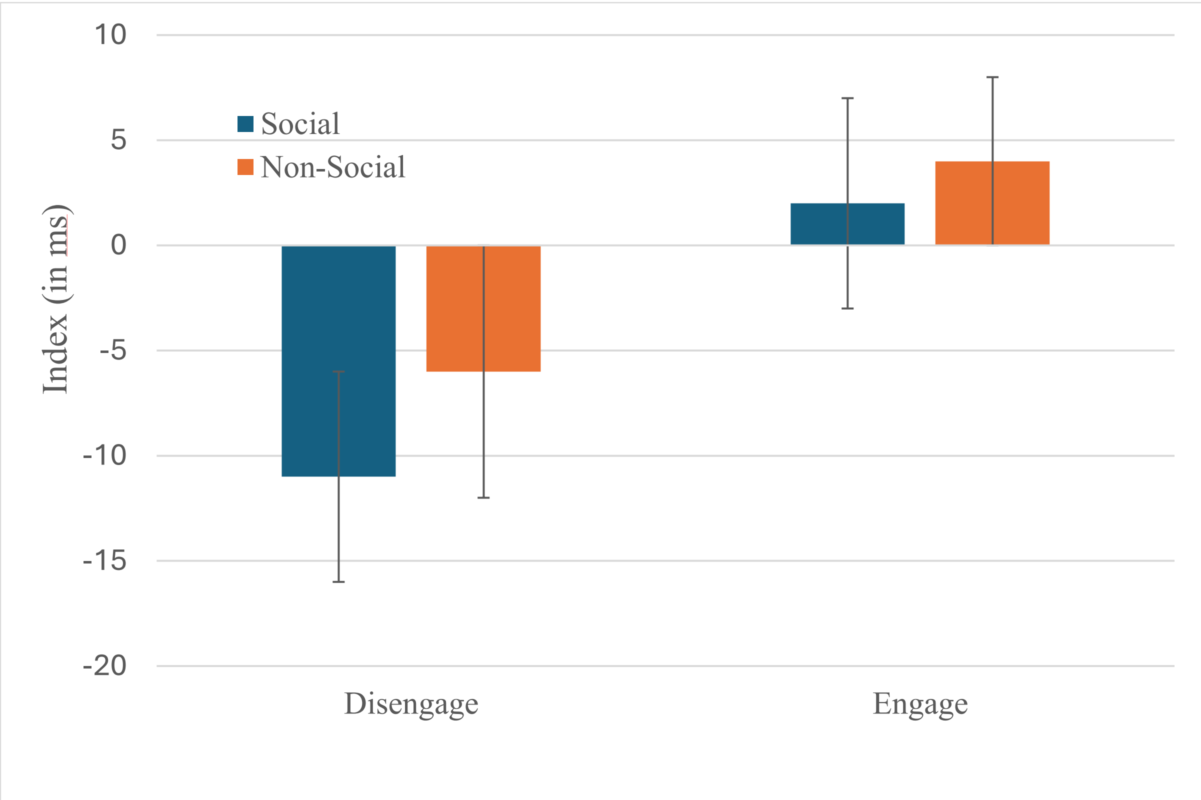

Figure 2. Mean engagement and disengagement indices as a function of target type

Trait anxiety as a covariate

As preregistered, we repeated the overall ANOVA with z-standardized trait anxiety (STAI) values as a covariate (see Table 2, right column, for the results). Given that we usually did not found moderations by trait anxiety in our dot-probe studies (but see Wirth & Wentura, 2018b), we wrote in the preregistration “we do not expect a significant moderation here", despite the fact that Grafton and MacLeod (2014) reported a trait anxiety moderation for the disengagement index. However, contrary to our expectation, the *Target Type* × C*ue Emotion* × C*ue Validity* interaction of the main analysis was significantly moderated by trait anxiety.

To elucidate this complex moderation, we conducted a 2 (target type: social vs. non-social) × 2 (index: engagement vs. disengagement) repeated-measures ANOVA with index values as dependent variable and z-standardized trait anxiety as covariate. The moderation of the triple interaction by STAI reported above returns here as the moderation of the target type main effect by STAI. Table 3 shows the correlations of the indices with STAI. We added the sum of the two indices as further variables because the correlations of STAI with the two sum indices are the adequate follow-up analyses to the moderation of the target type main effect by STAI.

Table 3

Correlations of Trait Anxiety with the Bias Indices

	Attentional bias type					
	Engagement Bias (EB)		Disengag. Bias (DB)		EB+DB	
Target Type	r	p	r	p	r	p
Social	.04	.646	-.09	.298	-.04	.641
Non-social	.25	.004	.12	.167	.28	.002

Note. EB+DB is the sum of the indices Engagement Bias and Disengagement Bias (see text for further explanation.

As can be seen, trait anxiety correlates positively with the non-social sum index, but not with the social one. That is, the higher the trait anxiety level the more is attention engaged by an angry face and the more attention dwells on an angry face (both relative to the neutral face), but only in the non-social target mode. However, looking at the correlations for engagement and disengagement separately reveals that it is especially the non-social *engagement* index that correlates with STAI.

A regression analysis with the non-social engagement index regressed on z-standardized trait anxiety values indicates that from trait anxiety values of ~0.5 SD above the mean on the 95% confidence band does no longer include zero, that is, starting with this trait anxiety value there is the estimation of a positive non-social engagement index.[2]

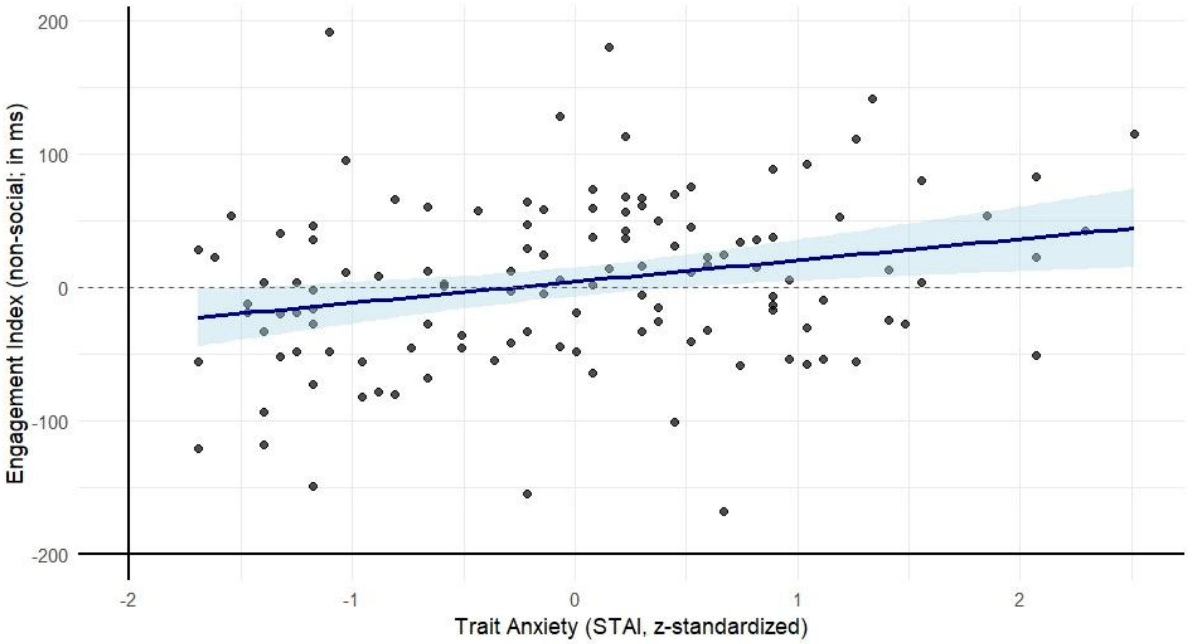

Figure 3. Scatterplot illustrating the relationship between participants' STAI-scores and their individual engagement scores (non-social condition; in ms). The solid line depicts the slope of the regression, the blue-shaded area is the 95%-confidence interval of the slope.

[2] To be more transparent with regard to the underlying analysis: Using zSTAI' (= zSTAI – 0.4842) as the predictor yields an intercept of b_0 = 12 ms (SE = 6 ms), $t(122) = 1.98$, p = .0499.

Accuracy

Results of a 2 (target type: social vs. non-social) × 2 (initial attention focus: proximal vs. distal to face cue) × 2 (cue emotion: angry vs. neutral) × cue validity (valid vs. invalid face cue) repeated-measures ANOVA with accuracy rates as dependent variable yielded are presented in Table A1 (Appendix). There was only one significant effect, that is, the rather trivial two-interaction of initial attention focus and cue validity (see above, analyses of RTs). Thus, there is no evidence for a speed-accuracy trade-off.

Discussion

In the present experiment, we combined elements of the dot-probe version of Wirth and Wentura (2018a, 2019) with Grafton and MacLeod's (2014) ARDPEI paradigm. Specifically, we introduced the target type variation of "social target" versus "non-social target," which was used by Wirth and Wentura in their dot-probe research. We found two main results: First, against expectation (based on our earlier research, Wirth & Wentura, 2018a, 2019), we found no overall evidence for a difference in engagement between angry and neutral faces. Instead, we found a difference in the disengagement index that indicates that on average neutral faces hold attention more than angry faces. Though this result was not significantly moderated by target type, it was more pronounced in the social target condition.

Second, we found that trait anxiety positively correlated especially with the engagement bias in the non-social target condition. This is an important results for several reasons. First, it shows that the distal condition in our experimental setting works, that is, engagement processes can in principle be found. Second, the direction of the correlation fits to the general expectation that high trait anxiety participants show capture by threatening stimuli. Third, the correlation is clearly constrained to the non-social block. We will discuss the results separately in their meaning for the research of Wirth and Wentura (2018a, 2019,

2023) on the one hand and Grafton and MacLeod's (2014) ARDPEI paradigm on the other hand.

Setting the results in relation to the dot-probe studies by Wirth and Wentura

At least on first sight, the present results and the ones of Wirth and Wentura (2018a, 2019, 2023) cannot easily be reconciled. Wirth and Wentura found cueing effects for angry faces, but only in the social target mode. These cueing effects were interpreted as capture effects, especially with regard to the N2pc results of Wirth and Wentura (2023). Of course, it was conceivable that the effects were based partly on attentional dwelling on angry faces. The present results do not support either interpretation. There was definitely no evidence for attentional engagement by angry faces (relative to neutral faces) in the overall sample since the mean engagement score was not significantly different from zero, neither in the social target mode, not in the non-social target mode. What was found was a significant mean disengagement effect. However, it was negative in sign, that is, attentional dwelling on *neutral* faces was larger than the one on angry faces. In the following, we will discuss both results with regard to their meaning for the dot-probe task.

The meaning of the missing engagement effect for the dot-probe

One initial response to the missing engagement effect would be to question whether our version of the ARDPEI paradigm can assess engagement processes at all at 100 ms CTOA. However, this answer can be refuted due to the correlations between the non-social engagement index and trait anxiety (see below for a discussion).

Instead, we must analyze the differences between the dot-probe and ARDPEI tasks. In the former, the two face types are in direct competition. Additionally, at the beginning of a trial, the focus of attention in the dot-probe task is halfway between the two possible target locations. There is no a priori tendency to first pay attention to one possible target location and then switch to the other if it does not contain the target. Therefore, the situation is optimally designed to determine if the *difference* in emotional expression between competing cue locations is reflected in a *difference* in focus of attention. Metaphorically speaking, one could describe it as a balance beam situation: the measuring apparatus is constructed in such a way that it is highly sensitive to differences between the two sides.

In the ARDPEI task, there is competition – or an "attention competition," if you will – between the abstract image stimulus and the face (separately for angry and neutral expressions). In the ARDPEI task, the anchor probe focuses attention on one location. For the present discussion of the "engagement" bias, this is the location where the face stimulus will not occur. Using the balance beam metaphor, one could describe the situation as follows: an unknown but potentially significant weight on one side competes with the face stimulus on the other side. Additionally, the balance beam is deliberately tilted initially to the side opposite the face, as the task suggests checking the possible target stimulus on the anchor probe side first to see if it is the target before moving on to the other side if it is not.

Thus, we should concede that the dot-probe task and the ARDPEI test different hypotheses. The dot-probe task is a finely tuned measurement tool that assesses differences in attentional capture qualities between two well-defined stimulus types that differ in only one specific feature. However, the dot-probe task is merely a measurement tool and should not be confused with an attempt to model a real-life situation in a laboratory. The ARDPEI, on the other hand, is an attempt to do so. Does a threatening stimulus (e.g., an angry face) attract attention when the observer is focused on a different location that requires direct action? This phrasing of the research question is reminiscent of Lavie's perceptual load theory (e.g., Lavie, 1995, 2005; Lavie & Tsal, 1994), which proposes that task-irrelevant distractors are processed in "low perceptual load" environments but not in "high perceptual load" environments. This reinterpretation is especially important for ARDPEI because it suggests manipulating the complexity of the stimulus situation on the side opposite the face.

The meaning of dwelling on neutral faces for the dot-probe task

Although the dominant interpretation of their dot-probe effects was in terms of capture processes, Wirth and Wentura (2018a, 2019) discussed an alternative interpretation in terms of attentional dwelling. The disengagement effect observed in the present study contradicts this interpretation. According to this interpretation, attention will initially be randomly on either the angry or neutral face. If the target appears on the same side as the attended face, the response is quick. However, if the target appears on the opposite side, the time it takes to shift attention depends on how easily it can disengage from the attended face. According to this logic, positive dot-probe effects would be found if attention dwelled on angry faces more than neutral ones, and the reverse would be true if attention dwelled on neutral faces more than angry ones. However, Wirth and Wentura's dominant finding was a positive dot-probe effect. Therefore, the finding that attention dwells on neutral faces cannot reinterpret these positive effects.

However, if we assume that dwelling on neutral faces was present in the dot-probe experiments whenever attention was accidentally on the neutral face, we can conclude that this process dampened the overall dot-probe effect. We can reconstruct the dot-probe situation as follows: There is a capture bias for angry faces, meaning the probability of attention being captured by an angry face is greater than the probability of attention being captured by a neutral face. To illustrate this logic, let's assume p(angry) = .4 and p(neutral) = .2. Each trial with attention on the angry face contributes to a positive dot-probe effect; each trial with attention on the neutral face runs counter to it. Hence, the surplus (.4 - .2 = .2) is decisive. However, an attended face contributes to the dot-probe effect in two ways: targets that appear at the attended location can be quickly processed, but attention must move to the opposite location if the target appears there. However, the latter process depends on how easily attention can disengage from the attended face. If attention is held longer on neutral faces than on angry faces, the proportion of trials with attention on the neutral face may dampen the dot-probe effect, which is primarily due to the capture bias for angry faces. Interestingly, Experiment 2 by Wirth and Wentura (2019) yielded somewhat unexpected results that fit to these considerations. While it aligned with the general expectation that the dot-probe effect for angry faces in the social target condition would significantly exceed that of the non-social condition, the specific results were puzzling. In the social condition, the dot-probe effect was nearly non-existent (M = 1 ms), while in the non-social condition, a significant negative effect was observed. However, this pattern can easily be explained by a two-process logic. If we assume that being captured by angry faces depends on social processing, whereas dwelling on neutral faces does not, then this pattern becomes clear.

What is the meaning of increased dwelling on neutral faces? A plausible answer to this question lies in the ambiguity of neutral faces. In social contexts, a smile is the common signal to indicate an affiliative relationship. Neutral expressions do not have this quality and are often seen as slightly negative, however, with unclear emotional category (e.g., Park et al., 2015).

The meaning of the trait anxiety results for the social mode hypothesis

A final point of the present results that has to be discussed with reference to Wirth and Wentura (2018a, 2019, 2023) is the fact that trait anxiety moderated the cueing effects, but only in the *non-social* target condition. This aligns with the authors' broader idea that threat-related attentional bias depends on an attentional control setting that is occasionally activated in non-anxious individuals but is permanently engaged in anxious individuals.

Thus, the fact that only the engagement index of the non-social condition correlates with trait anxiety aligns with the idea that individual differences in attentional bias studies are not primarily differences in basic attention systems, but rather, differences in attentional control settings. Admittedly, this statement is less than perfect since a general engagement effect (not moderated by trait anxiety) was missing for the social condition. Nevertheless, the social condition factor appears to systematically alter something, as evidenced by our finding of a significant moderation of the target type (social vs. non-social) by trait anxiety for the bias indices as dependent variables.

Setting the results in relation to the dot-probe studies by Grafton and McLeod

How do our results relate to those of Grafton and MacLeod (2014)? Obviously, results are not the same. But before evaluating these differences we should remind that it was an attempt to *conceptually* replicate the former study, including some changes from the original experiment to the present one. Most dominant, we introduced the competitive target screen (i.e., a target screen with two potential targets and thus the need to search for the target); Grafton and McLeod had a single stimulus target screen. We used 100 ms CTOA throughout whereas Grafton and McLeod varied 100 ms versus 500 ms CTOA on a trial-by-trial basis.

Grafton and MacLeod (2014) summarized their results by the sentence (p. 1298): "Using this ARDPEI task, our findings demonstrate that participants with higher levels of trait anxiety, compared to participants with lower levels of trait anxiety, display both relatively facilitated attentional engagement with and impaired attentional disengagement from negative information." Very roughly, this statement can indeed be used as a summary of our (trait anxiety-related) results as well, supplemented by the addition "these findings were constrained to the non-social target condition." Indeed, we found a Target Type × Trait Anxiety interaction in the 2 (target type: social vs. non-social) × 2 (index: engagement vs. disengagement) repeated-measures ANOVA with index values as dependent variable and z-standardized trait anxiety as covariate. The follow-up analyses indicated that a compound of non-social engagement and non-social disengagement index correlated with trait anxiety.

The most important difference between the results of Grafton and MacLeod (2014) and ours is the fact that they found the moderation of engagement by trait anxiety only for their 500 CTOA condition, but not for their 100 ms CTOA condition. In contrast to this, we found the correlation with the short CTOA. What might have caused the difference?

Of course, the difference between Grafton and MacLeod (2014) and our experiment cannot directly be related to our target screen manipulation, since the correlation of the engagement score with trait anxiety was found only in the *non-social* condition; Grafton and MacLeod's version of the ARDPEI is, of course, a *non-social* condition as well. However, to create social versus non-social conditions, we always realized a target competition situation: Participants had to select the correct target. Wirth and Wentura (2018a) compared a single-target screen with a target-competition screen and found engagement by angry faces only in the competition mode. They provided a tentative post-hoc interpretation by arguing (a) with the concept of a priority map (Wolfe, 1994, 2021) whose activity peaks guide attention and (b) an integration of temporal close visual presentations, that is, integration of activity values caused by the cue screen and the target screen. When a target display with two stimuli is presented, the target should create a slightly larger peak at its location in the priority map than the distractor due to the target-relevant feature being top-down activated. In the valid cue condition, the two activation differences — the difference generated by the angry face cue (versus the neutral face cue) and the difference generated by the schematic target face — add up to determine that the target position is attended to first. In the invalid cue condition, however, the two small activation differences oppose one another, decreasing the probability that the target will be attended first. However, when only one stimulus is presented in the target display, the target itself creates a massive activation peak due to its onset characteristics. Thus, when a single onset target is used, the massive activation peak at the target location always exceeds the activation of the opposite location, regardless of whether the small activation difference transferred from the cue display favors the target location or the opposite location. Thus, its location will be attended to first, regardless of which location was cued by the angry face.

Meanwhile, in basic attention research Lamy and colleagues (Darnell & Lamy, 2022; Lamy et al., 2018) proposed the priority accumulation framework to account in general for the interplay of cued-generated priorities and target-generated priorities. The basic assumption is, as above, that attentional priorities generated by a cue do not directly guide attention if two displays are temporally very close but will add to priorities generated by the stimuli of the target screen. Given this basic assumption the influence of a cue-related priority depends on the search difficulty of the target screen. The authors provide impressive evidence for their theory by systematically varying the similarity of target-screen distractors with the target.

Thus, it might be that the target competition setting was the decisive ingredient for the ARDPEI task to trigger trait anxiety dependent engagement processes in a 100 ms CTOA condition. Why did Grafton and MacLeod (2014) found a comparable result without target competition for their 500 ms CTOA condition? We do not know but want to point out again that this rather long CTOA is somewhat problematic, since it allows for more than one shift of covert attention (Posner & Cohen, 1984). For example, this allows again explanations involving dwelling processes: Attention might slip in a portion of trials to the face (see Lachter et al., 2004), irrespective of valence. It returns fast to the location of initial attention, except for angry face trials of high trait anxious individuals.

References

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), 1-24. https://doi.org/10.1037/0033-2909.133.1.1

Cooper, R. M., & Langton, S. R. H. (2006). Attentional bias to angry faces using the dot-probe task? It depends when you look for it. *Behaviour Research and Therapy*, *44*(9), 1321-1329. https://doi.org/10.1016/j.brat.2005.10.004

Darnell, M., & Lamy, D. (2022). Spatial cueing effects do not always index attentional capture: evidence for a priority accumulation framework. *Psychological Research-Psychologische Forschung*, *86*(5), 1547-1564. https://doi.org/10.1007/s00426-021-01597-0

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030-1044.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, *130*(4), 681-700. https://doi.org/10.1037/0096-3445.130.4.681

Grafton, B., & MacLeod, C. (2014). Enhanced probing of attentional bias: The independence of anxiety-linked selectivity in attentional engagement with and disengagement from negative information. *Cognition and Emotion*, *28*(7), 1287-1302. https://doi.org/10.1080/02699931.2014.881326

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*, 880-913.

Lamy, D., Darnell, M., Levi, A., & Bublil, C. (2018). Testing the attentional dwelling hypothesis of attentional capture. *Journal of Cognition*, *1*(1), 43.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 451-468.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75-82.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*, 183-197.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122-1135.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy*, *36*, 809-848.

Morys-Carter, W. L. (2021, May 18). *ScreenScale [Computer software]*. Pavlovia. https://doi.org/10.17605/OSF.IO/8FHQK

Park, H.-B., Han, J.-E., & Hyun, J.-S. (2015). You may look unhappy unless you smile: The distinctiveness of a smiling face against faces without an explicit smile. *Acta Psychologica*, *157*, 185-194. https://doi.org/https://doi.org/10.1016/j.actpsy.2015.03.003

Peirce, J. W., Hirst, R. J., & MacAskill, M. R. (2022). *Building Experiments in PsychPy (2nd ed.)*. Sage.

Petrova, K., Wentura, D., & Bermeitinger, C. (2013). What happens during the stimulus onset asynchrony in the dot-probe task? Exploring the role of eye movements in the assessment of attentional biases. *PLoS ONE*, *8*, e76335.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556). Erlbaum.

Spielberger, C. D., Gorsuch, R., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1988). *Cognitive psychology and emotional disorders*. Wiley.

Wirth, B. E., & Wentura, D. (2018a). Attentional bias to threat in the general population is contingent on target competition, not on attentional control settings. *Quarterly Journal of Experimental Psychology*, *71*(4), 975-988.

Wirth, B. E., & Wentura, D. (2018b). Furious snarling: Teeth-exposure and anxiety-related attentional bias towards angry faces [Article]. *PLoS ONE*, *13*(11), 14, Article e0207695. https://doi.org/10.1371/journal.pone.0207695

Wirth, B. E., & Wentura, D. (2019). Attentional bias towards angry faces is moderated by the activation of a social processing mode in the general population [Article]. *Cognition & Emotion*, *33*(7), 1317-1329. https://doi.org/10.1080/02699931.2018.1561423

Wirth, B. E., & Wentura, D. (2023). Social processing modulates the initial allocation of attention towards angry faces: Evidence from the N2pc component. *Social Cognitive and Affective Neuroscience*, *18*(1). https://doi.org/10.1093/scan/nsad070

Wolfe, J. M. (1994). Guided search 2.0. A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202-238.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060-1092. https://doi.org/10.3758/s13423-020-01859-9

Appendix

Table A1

Results of the Target type × Initial Attention Focus × Emotion × Cue Validity ANOVA with Accuracy Rates as Dependent Variable

	Accuracy		
Effekt	$F(1, 123)$	p	η_p^2
Attention focus (A)	0.69	0.406	0.006
Emotion (E)	0.55	0.460	0.004
Target Type (T)	0.45	0.506	0.004
Cue Validity (C)	0.80	0.374	0.006
A × E	0.82	0.366	0.007
A × T	2.81	0.096	0.022
A × C	16.88	0.000	0.121
E × T	3.00	0.086	0.024
E × C	0.87	0.352	0.007
T × C	1.44	0.233	0.012
A × E × T	1.73	0.191	0.014
A × E × C	0.70	0.405	0.006
A × T × C	0.10	0.749	0.001
E × T × C	2.81	0.096	0.022
A × E × T × C	0.23	0.636	0.002