COGNITION AND EMOTION, 2017
http://dx.doi.org/10.1080/02699931.2017.1315929

Angry faces are tracked more easily than neutral faces during multiple identity tracking

Jie Li[a], Lauri Oksama[b], Lauri Nummenmaa[c,d,e] and Jukka Hyönä[e]

[a]Section of Applied Psychology, Beijing Sport University, Beijing, People's Republic of China; [b]Department of Behavioral Sciences, National Defense University, Järvenpää, Finland; [c]School of Science, Aalto University, Finland; [d]Turku PET Centre, Turku University Hospital, Turku, Finland; [e]Department of Psychology, University of Turku, Turku, Finland

AQ1

ABSTRACT

We investigated whether and how emotional facial expressions affect sustained attention in face tracking. In a multiple-identity and object tracking paradigm, participants tracked multiple target faces that continuously moved around together with several distractor faces, and subsequently reported where each target face had moved to. The emotional expression (angry, happy, and neutral) of the target and distractor faces was manipulated. Tracking performance was better when the target faces were angry rather than neutral, whereas angry distractor faces did not affect tracking. The effect persisted when the angry faces were presented upside-down and when surface features of the faces were irrelevant to the ongoing task. There was only suggestive and weak evidence for a facilitatory effect of happy targets and a distraction effect of happy distractors in comparison to neutral faces. The results show that angry expressions on the target faces can facilitate sustained attention on the targets via increased vigilance, yet this effect likely depends on both emotional information and visual features of the angry faces.

ARTICLE HISTORY
Received 28 January 2016
Revised 27 March 2017
Accepted 29 March 2017

KEYWORDS
Multiple-identity tracking; attention; emotion; facial expression

Emotional stimuli convey information concerning an individual's survival and well-being. Individuals need to appraise the affective content of the stimuli to reveal their appetitive or aversive properties and to subsequently initiate approach or avoidance behaviour (Öhman, Soares, Juth, Lindström, & Esteves, 2012). For adaptive reasons, emotion and attention systems are closely intertwined (for a review, see Yiend, 2010). A large body of behavioural and neuroimaging research suggests that emotional stimuli can be processed rapidly and efficiently, and they are especially likely to capture attention when competing with non-emotional stimuli (for a review, see Vuilleumier, 2005).

Positive and negative emotional stimuli may interact differently with attention. Several studies have demonstrated that negative stimuli (especially threat- and fear-related) are detected faster and they are also more distracting than positive and neutral stimuli (e.g. Eastwood, Smilek, & Merikle, 2001; Öhman, Flykt, & Esteves, 2001), likely due to the immediate relevance of such stimuli to survival throughout the evolutionary history. Furthermore, neurophysiological research has revealed that negative vs. positive stimuli are processed faster, and that they also trigger larger and earlier responses in event-related potentials (ERPs), suggesting an early attentional bias towards threat (Feldmann-Wüstefeld, Schmidt-Daffy, & Schubö, 2011). On the other hand, several studies have also found an attentional bias towards positive stimuli, particularly toward happy faces, over neutral and negative stimuli (e.g. Calvo, Nummenmaa, & Avero, 2008). Attention capture by happy expressions likely occurs because they unambiguously communicate prosocial intent (Nummenmaa & Calvo, 2015). Due to the importance of social contacts to humans, happy faces may have evolved to be visually distinct from other expressions (Becker, Anderson, Mortensen, Neufeld, & Neel, 2011), and their sensory saliency may explain why

CONTACT Jie Li lijie.psy@gmail.com

they capture attention reflexively (see a review in Nummenmaa & Calvo, 2015). Consequently, happy faces are easily recognised in both foveal and extrafoveal vision (Calvo, Fernández-Martín, & Nummenmaa, 2014). In sum, the influence of emotional stimuli on attention depends on a variety of factors including emotional valence and visual saliency of the stimuli (Calvo & Nummenmaa, 2008), but also the context in which the stimuli are presented (Frischen, Eastwood, & Smilek, 2008).

Yet up to date, most of the studies examining the relationship between emotion and attention have focused on transient attentional capture by emotional pictures. However, emotional events (such as encounters with predators or other people) often last for long periods of time, during which the relevant stimuli also change location from time to time. Dealing with visual environments where task-relevant objects continuously move about in the visual environment requires sustained attention (Scholl, 2009). Sustained attention is usually effortful, enabling us to continuously process the visual objects after they have been intentionally or automatically selected (Pashler, 1998). Although sustained attention is subject to top-down control, it can also be affected by the properties of the objects in a bottom-up fashion (Li, Oksama, & Hyönä, 2016; Liu & Chen, 2012). Certain object properties may facilitate continuous attention engagement with the object. For instance, when emotional targets such as predators are present, it would be imperative that attention is sustained on them with the expense of non-emotional ones. Up to date, however, the effects of emotional information on sustained attention during dynamic object tracking remain poorly understood.

Here we adopted the multiple-identity tracking (MIT) paradigm, which has been successfully used for studying sustained attentive tracking of multiple moving objects with distinct identities (Horowitz et al., 2007; Oksama & Hyönä, 2004, 2008). Participants view a set of moving objects, of which a subset is designated as the targets whose identities and positions they should track. Successful tracking relies on sustained attention devoted on the targets, in order to know where each target moves to from moment to moment, while ignoring the distractors that are irrelevant to the ongoing tracking task. This type of tracking is based on interactions between the frontoparietal attention and working memory circuits as well as inferior temporal cortices supporting object recognition (Nummenmaa, Oksama, Glerean, & Hyönä, in press), thus it is well suited for studying the bottom-up stimulus influences on attentional tracking.

In our variant, several faces of different identities were moving around on the screen. Participants tracked a subset of the faces that were previously designated as targets, and then reported where each face had moved to. Importantly, we manipulated the emotional expressions (angry vs. happy) of the target and distractor faces. The emotional expressions were however irrelevant to the tracking task, as participants were only required to track the identities and locations of the target faces. This setting allowed us to investigate whether properties of the targets and distractors can be processed during dynamic tracking, which has been under dispute, and how task-irrelevant emotional information spontaneously interacts with sustained attention: If irrelevant emotional information is not processed at all as implied by some earlier research on tracking (Pylyshyn, 1989; Scholl, Pylyshyn, & Franconeri, 1999), emotional expressions on the faces would not affect the tracking performance. On the contrary, if the emotional expressions of the target and distractor faces are processed spontaneously during tracking, they may either facilitate or disrupt the sustained attention on the targets depending on the emotional expressions and their saliency. For instance, angry expressions on target faces may facilitate the sustained attention on targets as they may induce continued vigilance on the targets, whereas happy expressions on distractor faces may disrupt the sustained attention on targets as they are salient even in the peripheral vision so that the distractors may occasionally win the competition for attention against the targets.

Experiment 1

In Experiment 1, we investigated the effect of angry expressions on sustained attention. Participants tracked either angry or neutral target faces among neutral or angry distractors. We examined whether and how the tracking performance is influenced by the emotional expression of the target and distractor faces.

Method

Participants

Twenty-six students from Zhejiang University (9 females, 17 males) with normal or corrected to normal vision volunteered for the experiment. The age range of the participants was between 18 and 24 years. The study was approved by the institutional review board, and the participants provided informed consent.

Stimuli and apparatus

The stimuli were 48 digitised colour photographs selected from the Karolinska Directed Emotional Faces (KDEF; Goeleven, De Raedt, Leyman, & Verschuere, 2008; Lundquist, Flykt, & Öhman, 1998). The photos were of 24 Caucasian amateur actors (12 females – KDEF numbers 01, 02, 05, 09, 11, 13, 14, 19, 20, 26 29, 33; and 12 males – KDEF numbers 03, 05, 06, 08, 10, 11, 12, 13, 17, 23, 29, 31), gazing directly at the viewer, each posing an angry and a neutral expression. Nonfacial areas (e.g. hair, neck, etc.) in each photo were removed by applying an ellipsoidal mask. Each face subtended a visual angle of 3.0° × 4.0° at a 57-cm viewing distance. The faces were shown on a black background. The display area subtended 30.3° × 21.1°.

The stimuli were presented on a 19 in. CRT monitor (Dell Trinitron P992) with a screen resolution of 1024 × 768 pixels and a 100-Hz refresh rate. Participants were seated approximately 57 cm from the monitor. The experiment was programmed in Matlab (The Mathworks) using the Psychophysics Toolbox routines (Brainard, 1997).

The speed of motion was determined for each participant by the QUEST routine (Watson & Pelli, 1983), in which the speed was adjusted via a staircase procedure so that the average identity-tracking accuracy of each participant was around 75%. To this end, a QUEST block that contained 56 trials was performed by each participant prior to the actual experiment. Thirty-two additional facial images with neutral expression (16 females, 16 males) not used in the main experiment were used in the QUEST block. The images were from the Center for Vital Longevity Face Database modified by Ebner (Ebner, 2008; Minear & Park, 2004). Analogously to the facial images used in the experiment proper, nonfacial areas in each image were removed by applying the same ellipsoidal mask; each face subtended a visual angle of 3.0° × 4.0° at a 57-cm viewing distance. All the parameters for the stimulus presentation in the QUEST block were the same as in the experiment proper.

To eliminate possible influences of low-level image properties on the tracking performance, we assessed the mean and standard deviation of luminance level, root mean square contrast, and colour saturation for the red, green, and blue channels for the stimuli using Matlab 7.11. A set of ANOVAs revealed no differences between the three groups of angry, neutral, and happy (used in subsequent experiments) facial images in any of these measures (all $ps > .10$).

In addition, because the images used in the present experiments were faces of Caucasian people, whereas the participants were Chinese people, we conducted a control experiment to validate that Chinese observers can correctly recognise the facial expressions of these Caucasian actors. All the 72 facial images (24 actors with 3 expressions each) used in the tracking experiments were displayed on screen one by one. The images were presented either upright or upside-down, as both the upright and inverted emotional faces were used in the subsequent tracking experiments. The display size of the images was the same as in the tracking experiments. Participants were required to make a three-alternative forced choice regarding the expression of each face, by categorising it either as angry, neutral, or happy, by clicking one of the three labels on the screen under the image. Twenty Chinese students from Zhejiang University (10 males, 10 females) and 25 Chinese students from Beijing Sport University (17 females, 8 males) were tested for the recognition of the emotion on upright faces and inverted faces respectively. Most of the participants did not participated in the tracking experiments, except that seven participants from Beijing Sport University took part in the tracking experiments in some other days as a separate experiment. The results showed that the performance was practically at ceiling for recognising the emotion on upright faces. Mean recognition accuracies for the angry, neutral, and happy expression were 94.4%, 90.4%, and 99.8%, respectively, There was some decrease in the recognition performance for inverted expressions in comparison with upright expressions, yet the accuracy was still fairly high, close to 90% for the angry and neutral expressions and over 95% for the happy expression (Table 1).

Table 1. The accuracy and reaction time for recognising angry, happy, and neutral emotional expressions on upright and inverted faces.

	Upright faces		Inverted faces	
Expression	Accuracy (%)	RT (ms)	Accuracy (%)	RT (ms)
Angry	94.4 (1.5)	2062 (142)	89.5 (2.1)	1927 (144)
Happy	99.8 (0.2)	1576 (75)	96.5 (1.4)	1571 (132)
Neutral	90.4 (2.2)	1979 (119)	89.0 (2.8)	1790 (96)

The difference was marginally significant between recognising inverted and upright angry expressions [$t(43) = -1.782$, $p = .082$], significant between inverted and upright happy expressions [$t(43) = -2.267$, $p = .032$], and non-significant between inverted and upright neutral expressions [$t(43) = -0.373$, $p = .711$]. For all the three types of expressions, there was no difference in response time between the inverted and the upright conditions [$t(43) = -0.670$, $p = .506$; $t(43) = -0.026$, $p = .979$; $t(43) = 1.223$, $p = .229$, for angry, happy, and neutral expressions, respectively]. The results indicate that Chinese observers were able to recognise the facial expressions used in the present study. The recognition for the emotion on the upright faces was at ceiling level, while the recognition for the inverted faces decreased a little yet still close to ceiling.

Procedure and design

At the beginning of each trial, six faces of different identities were presented at random locations in the display area. Immediately upon the appearance of the faces, red circles flashed around 3 of the faces for 8 times, indicating that they were the targets to be tracked. The circles flashed for 3.2 s, and then remained around the 3 target faces for 2 s without flashing, during which all the faces remained still at their initial locations. Next, the circles disappeared, and all faces moved for a randomly determined duration between 3.0 and 8.0 s, which is a typical tracking duration in dynamic tracking studies (e.g. Cohen, Pinto, Howe, & Horowitz, 2011; Oksama & Hyönä, 2008, 2016). The faces did not overlap with each other during the motion. As soon as the motion stopped, each face was occluded by a dark grey oval. Subsequently, the three target faces appeared at the centre of the screen one by one. The participants were required to report where each target stopped by clicking the corresponding location (Figure 1). Participants were free to move their eyes during the experiments.

The experimental design was a 2 (target expression: angry, neutral) × 2 (distractor expression: angry, neutral) within-participants design. The expressions of the target and distractor faces were manipulated orthogonally; the expression of all the target faces on a trial was identical, and so was the expression of the distractor faces. All the faces shown during a trial were of same gender. Trial order was fully randomised. There were 24 trials in each of the 4 conditions thus totalling 96 trials.

Results and discussion

Two dependent variables – location accuracy and identity accuracy – were calculated. Location accuracy refers to the observers' ability to distinguish the target locations from the distractor locations, whereas identity accuracy refers to the participants' ability to know which specific target is where. Click on a location of any target face is considered a correct response in the calculation of the location accuracy, while only the click on the location of the probed face is considered a correct response in the calculation of the identity accuracy. With this definition, the identity accuracy is always lower than (or at most equal to) the location accuracy (Pinto, Howe, Cohen, & Horowitz, 2010). The two measures are related to each other, yet also dissociable, in that people may track the locations of objects without clearly knowing the specific content of each object. Both the location accuracy and the identity accuracy depend on observers' sustained attention on targets. Continuous attention to the locations of the targets is sufficient for achieving high location accuracy, whereas continuous attention to both the locations and identities of each target enables the observer to achieve high identity accuracy. Typically, people are able to track the locations of about four moving targets, but only about two or three identity-location bindings (Horowitz et al., 2007; Pylyshyn, 2004). For instance, in the seminal study for multiple object (location) tracking, Pylyshyn and Storm (1988) showed that participants' performance level was about 90% when tracking 4 target locations and about 85% when tracking 5 target locations. Pinto et al. (2010) showed that when participants were required to track both the identity and the location of 4 targets, their identity accuracy was around 60–80% while the location accuracy was higher than 90%.

When computing the location accuracy, we first counted in each trial how many locations clicked by the participant were indeed locations of targets, regardless of whether the identity-location bindings were correct. By dividing this number by the total numbers of targets in the trial, we obtained the location accuracy for each trial. Since there were always 3 targets in the experiment, the accuracy in each trial could be 0, 0.333, 0.667, or 1. For instance, if the participant clicked 2 target locations along with 1 distractor

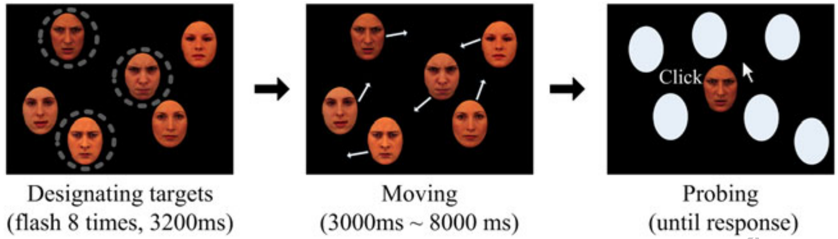

Figure 1. Trial structure in Experiment 1. First, targets are designated by flashing circles around them. In this example trial, the target objects are angry faces and the distractor objects are neutral faces. Then all the objects start moving, and participants track where each of the target face moves to. As soon as the motion stops, each face is occluded by a grey oval, and the three target faces appears at the centre as probes one by one. Participants report the final location of each probed face by clicking the oval covering that location.

location in a trial, the location accuracy in the trial was 0.667 (i.e. 2/3). Subsequently, by averaging the location accuracy across all the trials in each condition, we obtained the location accuracy under each condition. The computing process was similar for identity accuracy. We first counted how many correct responses for identity-location binding were made in each trial. That is, there were three probes appeared in turn in each trial, and participants responded to each probe by clicking the location where he/she thought the probed target finally located; a response was considered as correct only when the clicked location was indeed the location of the probed target but not the location of any other targets or distractors. Then, by dividing the correct responses for identity-location binding by the total numbers of targets in the trial, we obtained the identity accuracy for each trial, which also could be 0, 0.333, 0.667%, or 1. Subsequently, by averaging the identity accuracy across all the trials in each condition, we obtained the identity accuracy under each condition.

In the present experiment, the location accuracies in all conditions were close to ceiling (over 98%). This was as expected, since three target locations are within the capacity for location tracking, which is typically four or five locations (Horowitz et al., 2007; Pylyshyn & Storm, 1988). Thus, we focused on the analyses of identity accuracy.

We performed a 2 (target expression: angry, neutral) × 2 (distractor expression: angry, neutral) ANOVA on the identity-tracking accuracy. The main effect of the target expression was significant [$F(1, 25) = 29.954$, $p < .001$, $\eta_p^2 = .545$], as the accuracy for the target faces with an angry expression was higher than that for the targets with a neutral expression [86.0% vs. 79.1%]. The main effect of the distractor expression was not significant [$F(1, 25) = 0.022$, $p = .884$, $\eta_p^2 = .001$], neither was the interaction between the target expression and distractor expression [$F(1, 25) = 0.041$, $p = .842$, $\eta_p^2 = .002$] (see Figure 2(a)).

The results indicate that the emotional value of the targets can impact attentive tracking, even though it is irrelevant to the tracking task. When the target faces displayed an angry rather than a neutral expression, they were tracked more accurately, indicating that an angry expression can facilitate the sustained attention on the target faces.

Experiment 2

Positive emotional stimuli signal different information concerning individuals' survival and well-being than negative emotional stimuli. Previous research on attention capture by emotional stimuli has found that positive and negative emotional stimuli, particularly facial expressions, may differently influence attentional allocation (Becker et al., 2011; Calvo et al., 2008). For instance, by briefly presenting faces in a flanker paradigm, Fenske and Eastwood (2003) showed that an angry expression on the central target face constricts attention on the target more effectively than a neutral expression, whereas a happy expression may dilate the focus of attention.

In Experiment 2 we investigated how happy expressions affect sustained attention in dynamic circumstances. We were interested in whether a happy expression on the target faces improves the tracking performance, and whether a happy expression on the distractor faces impairs the tracking performance.

The experiment was identical to Experiment 1 except that the angry faces used in Experiment 1 were substituted by happy faces expressed by the same actors/actresses. A new group of 26 students from Zhejiang University (14 females, 12 males) with normal or corrected to normal vision volunteered for

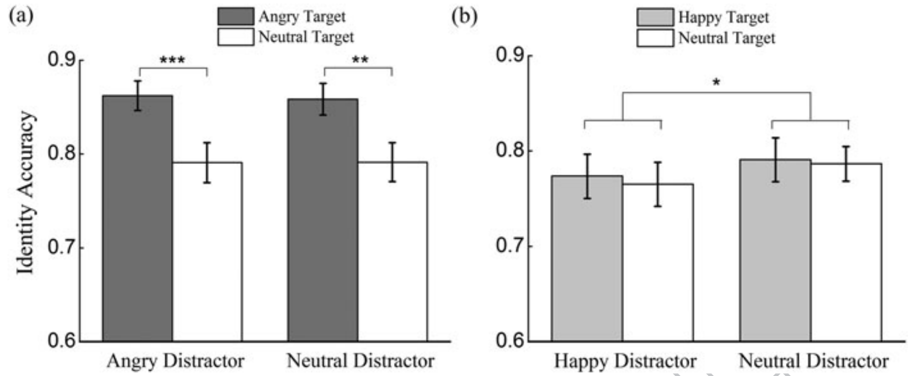

Figure 2. Identity accuracy in Experiment 1(a) and Experiment 2(b). Error bars represent the standard error of means. *indicates $p < .05$. **indicates $p < .01$. ***indicates $p < .001$.

the experiment. The age range of the participants was between 18 and 24. All participants provided informed consent.

Results and discussion

The location accuracies in all the conditions were close to ceiling (over 98%), thus we again focused the analysis on identity accuracy. As in Experiment 1, we performed a 2 (target expression: happy, neutral) × 2 (distractor expression: happy, neutral) ANOVA on the identity accuracy. The main effect of the target expression was not significant [$F(1, 25) = 0.217$, $p = .646$, $\eta_p^2 = .009$], whereas the main effect of the distractor expression was significant [$F(1, 25) = 5.157$, $p = .032$, $\eta_p^2 = .171$]. The accuracy was lower when the distractor faces showed a happy expression than a neutral expression [76.9% vs. 78.8%]. The interaction between target expression and distractor expression was not significant [$F(1, 25) = 0.040$, $p = .843$, $\eta_p^2 = .002$] (see Figure 2(b)).

Experiment 2 showed that a happy facial expression also influences attentive tracking, but in a different way from an angry expression. A happy expression on the target faces did not facilitate sustained attention on the targets. In contrast, when the distractor faces showed a happy expression, the tracking performance for the target faces was less accurate, compared with the case when the distractor faces showed a neutral expression. The result indicates that happy expression may disrupt the sustained attention on the targets by periodically drawing the observer's attention to the distractor faces.

Because Experiments 1 and 2 were run using an identical design but a different sample of participants, we next conducted an aggregate analysis across experiments to directly compare the effects of happy and angry expressions on sustained attention. We collapsed the data over the experiments, and ran a 2 × 2 × 2 ANOVA with the emotional valence of the facial expressions (happy, angry) as the between-participants factor and the target type (emotional, neutral) and the distractor type (emotional, neural) as the within-participants variables. The results showed a significant main effect of target type [$F(1, 50) = 16.411$, $p < .001$, $\eta_p^2 = .247$], and a significant interaction between emotional valence and target type [$F(1, 50) = 11.332$, $p = .001$, $\eta_p^2 = .187$]. The main effect of emotional valence approached significance [$F(1, 50) = 3.399$, $p = .071$, $\eta_p^2 = .064$]. None of the other effects approached significance ($ps > .137$). A follow-up analysis showed that the Emotional Valence × Target Type interaction was due to the effect of target type being significant when the emotion was angry [$F(1, 50) = 27.508$, $p < .001$, $\eta_p^2 = .355$] but not when the emotion was happy [$F(1, 50) = 0.234$, $p = .630$, $\eta_p^2 = .005$]. That is, in comparison with a neutral expression the accuracy was higher when the targets showed an angry expression, but not a happy expression, consistent with the results in Experiments 1 and 2. The distracting effect of happy distractors in Experiment 2 did not show up in this omnibus ANOVA.

Experiment 3a

In Experiments 1 and 2, only angry or happy faces were presented along with neutral faces. This may have led to corresponding changes in participants' mood, which may be responsible for the effects in the tracking performance. To rule out this possibility, we conducted Experiment 3, in which the trials with angry and happy faces were intermixed. This would eliminate the possible influence of mood, as well as strategy, general arousal, and repetition effects in experiments using a block design.

In addition, it was possible that the discriminability of the faces influenced participants' tracking performance. Although the angry, happy, and neutral faces used in the present experiments were posed by the same group of actors/actresses, the emotional faces may be more easily discriminated from each other while the neutral faces may be less distinctive. The higher discriminability may help participants to distinguish the three target faces during tracking and then report which is where, resulting in higher identity accuracy for emotional faces than neutral faces. To examine this possibility, we measured in Experiment 3 the discriminability of the faces during tracking, and analysed whether the emotional expression still has an effect on the tracking performance after the effect of face discriminability is partialed out.

Method

Participants

Twenty-six students from Beijing Sport University (17 females, 9 males) with normal or corrected to normal vision volunteered for the experiment. The age range of the participants was between 19 and 22 years. The study was approved by the institutional review board, and the participants provided informed consent.

Stimuli and apparatus

The stimuli were the same as those used in Experiments 1 and 2. All the three types of faces (angry, happy, and neutral) were used in Experiment 3. The stimuli were presented on a 17 in. CRT monitor (SyncMaster 785MB) with a screen resolution of 1024 × 768 pixels and an 85-Hz refresh rate. Participants were seated approximately 51 cm from the monitor. The size of the facial images and the display area were the same as those in Experiments 1 and 2 in terms of visual angle. The speed of motion was fixed at 4.5°/s for all the participants.

Procedure and design

The procedure of Experiment 3 was similar to Experiments 1 and 2, except for the following differences. In half of the trials the probe was one of the targets whereas in the other half it was a new face showing the same expression as the targets, yet not presented during tracking. At the response stage, there was a button labelled "Not a target" at the upper right corner of the screen. If the participant judged that the probe was a new face, he/she clicked the button "Not a target". If the participant judged that the probe was one of the targets, he/she reported the final location of that target by clicking the disk covering that location – procedure identical to that used in Experiments 1 and 2. The experimental design was a 3 (target expression: angry, happy, and neutral) × 3 (distractor expression: angry, happy, and neutral) × 2 (probe type: target face, new face) within-participants design. Trial order was fully randomised. There were 144 trials in total.

To perform this task, the participants needed to discriminate the three target faces, compare the probe with the targets, and then judge whether the probe matched one of the targets and where the corresponding target was located. The inclusion of a new face as the probe allowed us to use the method derived from the signal-detection theory (SDT) (Green & Swets, 1966; Stanislaw & Todorov, 1999) to calculate participants' sensitivity (d') and response bias (β) for the faces while performing the tracking task. The sensitivity measure reflects how well the participants could discriminate the faces. When computing the SDT measures, the trials in which the probe was one of the targets were treated as signal present, while those in which probe was not a target but a new face were treated as signal absent. For the signal present trials, if the participant clicked "Not a target", the trial was considered a "miss"; if the participant clicked one of the object locations (no matter which one), the trial was considered a "hit" in that it showed that the participant recognised that the probed target was present during tracking. For the signal absent trials, if the participant clicked "Not a target", the trial was considered a "correct rejection"; otherwise, if the participant clicked one of the object locations, the trial was a "false alarm".

The calculation of tracking accuracy in Experiment 3 was slightly different from that in Experiments 1 and 2. Only in the trials where the probe was one of the targets, the tracking accuracy could be calculated. These trials accounted for half of the total trials. In addition, there was only one probe in each trial and the participants made only one response. The trial was considered correct only when the participant clicked the location of the target that matched the identity of the probe. This measure corresponds to the identity accuracy in Experiments 1 and 2.

The advantage of the current task setting is that we were able to compute the sensitivity measure and the tracking performance at the same time, so that the sensitivity measure can be used to reflect participants' discriminability of the faces while they were tracking the moving faces.

Results and discussion

The identity accuracy in each condition is presented in Figure 3(a). Analogously to Experiments 1 and 2, we performed a 3 (target expression: angry, happy, neutral) × 3 (distractor expression: angry, happy, neutral) repeated measures ANOVA on the identity accuracy. The results yielded a significant main effect of target expression [$F(2, 50) = 7.407$, $p = .003$, $\eta_p^2 = .229$]. Pairwise comparisons showed that angry targets were tracked better than happy and neutral targets (63.6% vs. 59.5% and 53.5%, $p = .045$ and .001), while the difference between happy and neutral targets was close to significant ($p = .071$). The main effect of distractor expression was not significant [$F(2, 50) = 1.839$, $p = .170$, $\eta_p^2 = .069$], neither was the interaction [$F(4, 100) = 1.497$, $p = .209$, $\eta_p^2 = .056$].

In order to measure and partial out the possible effect of face discriminability, we computed the indexes of sensitivity (d') and response bias (β) for each condition (see Table 2), and then conducted 3 (target expression: angry, happy, neutral) × 3 (distractor expression: angry, happy, neutral) repeated measures ANOVAs on d' and β. For d', there was a significant main effect of target expression [$F(2,50) = 13.275$, $p < .001$, $\eta_p^2 = .347$], showing that the discriminability for angry targets was higher than happy or neutral targets (2.7 vs. 2.2 and 2.0, $ps < .001$), while there was no difference between happy and neutral targets ($p = .292$). The main effect of distractor expression was not significant [$F(2,50) = 0.836$, $p = .439$, $\eta_p^2 = .032$], neither was the interaction [$F(4,100) = 1.459$, $p = .220$, $\eta_p^2 = .055$]. For β, neither the main effects of target expression and distractor expression nor the interaction were significant [$F(2,50) = 1.177$, $p = .317$, $\eta_p^2 = .045$; $F(2,50) = 1.050$, $p = .358$, $\eta_p^2 = .040$; $F(4,100) = 2.403$, $p = .055$, $\eta_p^2 = .088$, respectively]. The results indicate that the angry targets were more discriminable than the neutral targets during tracking, while the discriminability of the happy targets was similar to that of the neutral targets.

Subsequently, we used mixed-effect logit models (Arnon, 2010; Jaeger, 2008) to examine whether after partialing out the effect of face discriminability the target expression and the distractor expression had an effect on the tracking performance. The mixed-effect logit model is an extension of logistic regression that includes modelling of random subject effects. The modelling is based on participants' original response data in each trial rather than the averaged data across trials, so that it can utilise the data more effectively than ANOVA and reveal effects that are more subtle (Agresti, 2007; Jaeger, 2008). By model comparisons, we were able to examine how various factors affected the probability of correct response. The analyses were conducted in R (R Core Team, 2014) by using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015).

As a preliminary step, we built a basic model with the target expression and the distractor expression as fixed effects and the subject as a random intercept, while other possible effects, such as the interaction effects and the effect of face discriminability, were not included. The dependent variable was whether the participant made a correct response in each trial. The results yielded significant effects of target expression, showing that the angry targets were tracked better than the neutral targets ($B = 0.45$, $SE = 0.120$, $p < .001$, $e^B = 1.57$), and the happy targets were also tracked better than the neutral targets ($B = 0.26$, $SE = 0.119$, $p = .028$, $e^B = 1.30$). There was also a marginally significant effect of happy distractors, showing that the performance was worse when the distractors were happy faces rather than neutral faces ($B = -0.23$, $SE = 0.120$, $p = .056$, $e^B = 0.80$). The angry distractor faces did not exert a significantly larger distraction effect than that of the neutral distractors ($B = -0.09$, $SE = 0.120$, $p = .470$, $e^B = 0.92$).

In the next step, we added face discriminability to the model as a fixed effect, to examine whether face discriminability affected the tracking performance, and whether expression of the faces still affected the performance after the effect of face discriminability was partialed out. The results showed that the fitness of this model was significantly better than the basic model reported above ($\chi^2(1) = 30.529$, $p < .001$), indicating that taking face discriminability into account did improve the explanatory power. There was a significant effect of face discriminability, showing that as the discriminability increased, the probability of correct tracking increased ($B = 0.30$, $SE = 0.055$, $p < .001$, $e^B = 1.35$). More importantly,

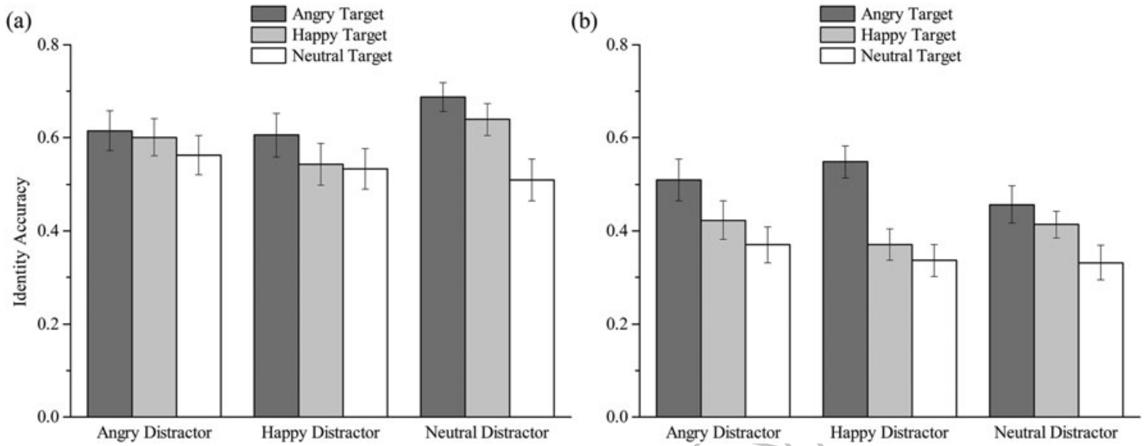

Figure 3. Identity accuracy in Experiments 3(a) and (b). Error bars represent the standard error of means.

there was still a significant effect of target expression, showing that the angry targets were tracked better than the neutral targets ($B = 0.27$, SE = 0.125, $p = .034$, $e^B = 1.30$). Thus, the results reveal that an angry expression improves sustained attention on the target faces even after the effect of face discriminability is partialed out. The magnitude of the effect appeared smaller than that in the basic model, suggesting that the effect resulted partly from the higher face discriminability and partly from the angry expression. The effect of happy targets was close to significant ($B = 0.21$, SE = 0.120, $p = .073$, $e^B =$ 1.24). The effects for the distractor expression were not significant, yet there was a trend suggesting that happy distractor faces leading to a poorer tracking performance than neutral distractor faces ($B = -0.18$, SE = 0.121, $p = .139$, $e^B = 0.84$). The effect of angry distractor faces was negligible and non-significant ($B =$ -0.07, SE = 0.121, $p = .566$, $e^B = 0.93$) (see Table A1 in the Appendix for a full report).

We also tested other possible models, such as further adding the interaction effects and the by-subject random slope to the model. The results of Chi square tests showed that, despite being more complicated, the fitness of these models was not better than the model above ($ps > .114$). Thus, the model above can be taken as the best model for the participants' response. It demonstrates that compared with a neutral expression, an angry expression leads to a higher probability of correctly tracking the target faces; a happy expression on target faces also leads to a facilitatory effect, and a happy expression on distractors exerts a detrimental effect on the tracking performance, yet these latter effects are smaller and less robust.

Experiment 3b

Previous research suggests that the impact of emotional facial expressions on attention can arise from the emotional content of the expressions and/or the visual features of the expressions (Juth, Lundqvist, Karlsson, & Öhman, 2005; Nummenmaa & Calvo, 2015). To disentangle the possible roles of the two factors, studies have compared the recognition and detection of upright vs. inverted faces (Calvo & Nummenmaa, 2008; Eimer & Holmes, 2002; Fox & Damjanovic, 2006; Horstmann & Bauland, 2006). In inverted emotional faces the visual features are preserved, while inversion impairs at least configurational or holistic processing of the emotional expressions (Eimer & Holmes, 2002; but see Nummenmaa & Calvo, 2015 for an alternative explanation for feature-based recognition). Thus, in Experiment 3b, we used inverted faces as stimuli in an attempt to examine the role of emotional content and visual features of facial expressions on sustained attention. Experiment 3b was identical to Experiment 3a, except that the faces were presented upside-down. A new group of 26 students from Beijing Sport University (13 females, 13 males) with normal or corrected to normal vision volunteered for the experiment. The age range of the participants was between 18 and 22 years.

Results and discussion

Condition-wise accuracy scores are shown in Figure 3 (b). Analogously to Experiment 3a, we performed a 3 (target expression: angry, happy, neutral) × 3 (distractor expression: angry, happy, neutral) repeated measures

Table 2. Sensitivity (d') and response bias (β) in each condition in Experiment 3(a).

	d'			β		
Distractor expression	Angry	Happy	Neutral	Angry	Happy	Neutral
Target expression						
Angry	3.0 (0.24)	2.4 (0.18)	2.6 (0.23)	4.8 (1.08)	2.6 (0.79)	2.9 (0.92)
Happy	2.0 (0.15)	2.1 (0.22)	2.4 (0.17)	2.3 (0.81)	1.7 (0.61)	3.3 (0.92)
Neutral	1.9 (0.28)	2.1 (0.24)	2.1 (0.18)	1.6 (0.50)	2.9 (0.89)	3.6 (0.94)

Note: Standard errors are in parentheses.

ANOVA on the identity accuracy. The results yielded a significant main effect of target expression [$F(2, 50) = 16.257$, $p < .001$, $\eta_p^2 = .394$]. Pairwise comparisons showed that angry targets were tracked better than happy and neutral targets (50.5% vs. 40.2% and 34.6%, $p = .002$ and $p < .001$), while happy targets were tracked better than neutral targets ($p = .020$). The main effect of distractor expression was not significant [$F(2, 50) = 0.893$, $p = .416$, $\eta_p^2 = .035$], neither was the interaction between target expression and distractor expression [$F(4, 100) = 1.328$, $p = .265$, $\eta_p^2 = .050$]. Thus, angry expressions still facilitate attentional tracking even when presented upside-down. The overall performance in Experiment 3b was quite low for some participants in some conditions, presumably because tracking identities of inverted faces was much more difficult than that of upright faces. When performing the SDT analysis as in Experiment 3a, d' computed in Experiment 3b was around or even below 0 in some occasions. The reliability of the ANOVAs on d' and subsequent mixed-effect models including d' may be discounted under such circumstances, and thus we did not conduct further analyses.

Experiment 3b showed that the inverted angry and happy faces were tracked better than the inverted neutral faces, similar to the result pattern of tracking upright faces in Experiment 3a. Such results are consistent with the performance in the control experiment for emotion recognition, which showed that the recognition of emotion was accurate and fast for both upright and inverted faces (see the section of *Stimuli and apparatus* in Experiment 1). It is possible that some visual features of the emotional faces (e.g. frowning eyebrows of the angry faces) facilitate sustained attention on the target faces both when presented upright and upside-down. Yet, it is also possible that the emotional content of the inverted faces can be effectively processed when being continuously presented in the tracking task (Derntl, Seidel, Kainz, & Carbon, 2009), which in turn facilitates sustained attention. This issue is discussed in more detail in the General Discussion.

Experiment 4

Experiments 1–3 showed that emotional expressions influence identity-location binding in multiple-identity-tracking tasks. In Experiment 4, we examined whether the emotional expressions can also impact the performance of tracking only the locations of the faces. We adopted a modified version of the multiple-object location tracking task, in which participants did not need to discriminate the target faces and bind each identity with its location, but rather just track the locations of the target faces and then report their final locations at the end of motion. Faces with angry, happy, and neutral expressions served as the targets and distractors. The hypothesis was that the locations of the angry targets would be tracked better than those of the neutral targets.

Method

Participants

Thirty students from Beijing Sport University (16 females, 14 males) with normal or corrected to normal vision volunteered for the experiment. The age range of the participants was between 20 and 29 years. The study was approved by the institutional review board, and the participants provided informed consent.

Stimuli and apparatus

The stimuli and apparatus were the same as in Experiment 3. Participants were seated approximately 51 cm from the monitor. Each face subtended a visual angle of 2.6° × 3.4°. The whole display area subtended a visual angle of 28.5° × 19.4°. The speed of motion was fixed at 6.1°/s for all the participants.

Procedure and design

The procedure was similar to that in the experiments above (see Figure 4). At the beginning of each trial,

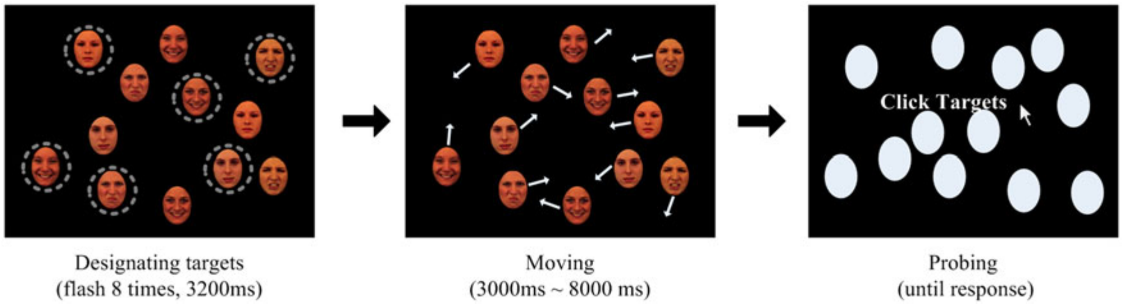

Figure 4. Trial structure in Experiment 4. Participants were required to track the locations of six target faces (two angry, two happy, two neutral) among six distractors that were the same as the targets, and then click the final locations of all the target faces without regard to the identity of each one. Objects are not drawn to scale.

12 faces were presented at random locations. Immediately upon the appearance of the faces, red circles flashed around six of the faces for eight times, indicating that they were the targets to be tracked. The circles flashed for 3.2 s, and then remained around the 6 target faces for 2 s without flashing, during which all the faces remained still at their initial locations. Next, the circles disappeared, and all faces moved for a randomly determined duration between 3.0 and 8.0 s. As soon as the motion stopped, each face was occluded by a dark grey oval. The faces could overlap during motion, but not at the beginning and the end of motion. The participants were required to track the changing locations of the targets, and report the final locations of all the targets by clicking the corresponding ovals. Participants were free to move their eyes during the experiment.

In each trial, the six target faces were of the same gender but different identities. Importantly, the six target faces were showing different emotional expressions: two of them were angry, two were happy, and the other two were neutral. The six distractor faces were exactly the same as the six target faces, so that the participants could not judge which faces were the targets by their appearance and could not recover a target after losing track of it. There were 80 experiment trials, preceded by 8 practice trial. In half of the trials, all the faces were upright, while in the other half, all the faces were inverted. The trials were randomly mixed.

Results and discussion

We counted how many locations of the angry, happy, and neutral targets were correctly reported in each trial, and then computed the location accuracy for each target type (see Table 3). A 3 (facial expression: angry, happy, neutral) × 2 (face orientation: inverted, upright) ANOVA on the location accuracy revealed a significant main effect of the facial expression [$F(2, 58) = 5.568$, $p = .006$, $\eta_p^2 = .161$]. Pairwise comparisons showed that the angry targets were tracked better than the happy and neutral targets (72.2% vs. 69.2% and 69.2%, $p = .007$ and $.009$, respectively), while there was no difference between the happy and neutral targets ($p = .984$). The main effect of face orientation was not significant [$F(1, 29) = 0.022$, $p = .883$, $\eta_p^2 = .001$], neither was the interaction between facial expression and face orientation [$F(2, 58) = 0.495$, $p = .612$, $\eta_p^2 = .017$].

We next examined whether the locations of the angry and happy distractors would be mistaken as target locations more frequently than those of the neutral distractors. We counted how many locations of the angry, happy, and neutral distractors were clicked by mistake in each trial, and then computed the error rate for each distractor type (see Table 3). A 3 (facial expression: angry, happy, neutral) × 2 (face orientation: inverted, upright) ANOVA on the error rate showed that neither the main effect of facial expression, the main effect of face orientation, nor the interaction was significant [$F(1, 29) = 0.022$, $p = .883$, $\eta_p^2 = .001$; $F(2, 58) = 0.199$, $p = .820$, $\eta_p^2 = .007$; $F(2, 58) = 0.716$, $p = .493$, $\eta_p^2 = .024$].

Table 3. Results of Experiment 4: The percentages of target faces with different emotional expressions being correctly tracked, and the percentages of distractor faces with different emotional expressions being mistaken as targets.

	Angry	Happy	Neutral
Targets correctly tracked (%)			
Inverted faces	71.9 (1.7)	69.1 (1.4)	69.6 (1.8)
Upright faces	72.5 (1.5)	69.2 (1.6)	68.8 (1.8)
Distractors mistaken as targets (%)			
Inverted faces	29.2 (1.6)	29.5 (1.8)	30.7 (1.4)
Upright faces	29.6 (1.4)	30.5 (1.7)	29.4 (2.2)

Note: Standard errors are in parentheses.

The results of Experiment 4 demonstrated that the angry but not happy expressions facilitated attentional tracking of the target locations, both when the faces were presented upright and upside-down. Yet, neither the angry nor the happy expressions on distractor faces interfered more with the tracking of the target locations than the neutral distractors.

Summary of Experiments 1–4: A local meta-analysis

To summarise our findings, we computed a local meta-analysis for the target and distractor effects of emotional faces across Experiments 1–4 (see Figure 5). Corresponding weighted effect sizes (*r*) were computed and subjected to meta-analysis with unbiased estimates of correlation coefficients and a restricted maximum likelihood estimator, yielding mean and 95% confidence intervals (CIs) for the effect sizes. Such a model assumes that effect sizes depend on study parameters, allowing for an estimation of both within- and between-studies variances. The results demonstrate an overall facilitatory effect of angry targets, with a mean effect size of 0.585 with a confidence interval (0.443–0.727) not overlapping zero, while the effect sizes of happy targets ($r = 0.182$, 95% CI = [−0.003, 0.368]) as well as angry ($r = 0.101$, 95% CI = [−0.072, 0.275]) and happy ($r = -0.167$, 95% CI = [−0.372, 0.037]) distractors had a CI that was overlapping with zero.

General discussion

Our main finding was that attentive tracking of multiple moving faces varies as a function of the emotional expression of the faces. In comparison with neutral expressions, angry expressions on target faces led to higher tracking performance, whereas angry expressions on distractor faces did not interfere with tracking. Such a facilitatory effect of the angry expression was found consistently when the faces were presented both upright and upside-down, and the effect held after the effect of face discriminability had been partialed out and when only target locations, but not identities were required to be tracked. On the other hand, the effect of the happy expression was less robust. In the identity-tracking task, there was a trend toward improved tracking of happy over neutral targets; moreover, happy distractors exerted a larger distraction effect than neutral distractors. However, in the location tracking task happy expressions did not influence tracking at all.

Taken together, the results indicate that emotional information can be processed spontaneously during attentive tracking, even if it is irrelevant to the tracking task. This is at odds with earlier research on dynamic tracking which suggested that observers are “blind” to the properties of the moving objects except their locations (Pylyshyn, 1989; Scholl et al., 1999). However, more recent studies accord with the present findings (Papenmeier, Meyerhoff, Jahn, & Huff, 2014). For instance, target colour (Makovski & Jiang, 2009a, 2009b; Papenmeier et al., 2014) and even complex

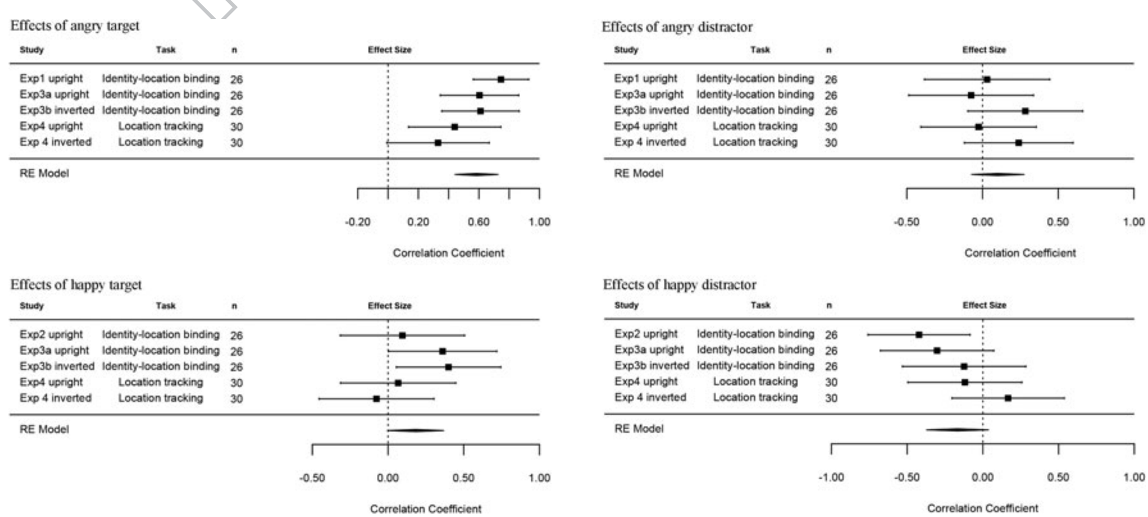

Figure 5. Local meta-analyses of the effects of emotional expressions across the experiments.

information such as targets' facial attractiveness and social identity can influence the tracking performance (Allen & Gabbert, 2013; Li et al., 2016; Liu & Chen, 2012). Thus, object properties other than mere spatio-temporal information can be processed during dynamic tracking. The present study extended previous research by showing that task-irrelevant emotional information can also be processed during tracking and affect tracking performance.

We also show that the task-irrelevant emotional information is processed even when the task puts a heavy load on observers' attention and working memory. In the MIT task, observers have to maintain the target identities and locations in working memory, and constantly refresh the identity-location binding (Oksama & Hyönä, 2008). The capacity of such dynamic identity tracking is typically around two or three objects (Horowitz et al., 2007). When only locations need to be tracked, the capacity is typically four to five object locations (Pylyshyn & Storm, 1988). In the present experiments, the task load was three object identities (Experiments 1–3) or six object locations (Experiment 4), thus around or slightly beyond the typical capacity limit. The irrelevant emotional information of the faces could still be extracted under such circumstances and affected the performance of attentional tracking. In future studies, it would be interesting to vary the number of targets according to each observer's capacity, so as to examine the relationship between working memory load and the processing of irrelevant emotional information.

As noted above, we found that target faces with angry expressions were tracked better than those with neutral expressions. The result may indicate that the presence of angry target expressions induces continuous vigilance for threat, and hence facilitates sustained attention for tracking their identities and/or locations. Such an effect is consistent with the "anger superiority" effect on emotion and attention (Fox & Damjanovic, 2006; Pinkham, Griffin, Baron, Sasson, & Gur, 2010). Previous research has mostly focused on short-lived attentional effects of emotional information, such as rapid attentional engagement and disengagement. Using visual search paradigms and dot-probe paradigms, studies have found that angry faces are more easily detected and more likely draw attention than neutral faces (Eastwood et al., 2001; Schmidt-Daffy, 2011). The present study qualified these findings by showing that the angry expressions engage also sustained attention on targets. Such "sustained anger superiority" may have great adaptive value from the perspective of evolution: considering that the angry expression signals threat towards the observer, it is critical not only to quickly detect and direct attention to the angry faces, but also to stay vigilant for them as long as they continue to pose a potential threat to the observer.

When the angry faces served as distractors in the attentive tracking task, however, they affected tracking similarly to neutral distractors. This contradicts with the previous studies showing both an attraction effect and a distraction effect from angry faces using visual search paradigms (for a review, see Nummenmaa & Calvo, 2015). For instance, Hahn, Carlson, Singer, and Gronlund (2006) found that participants showed a more effective search when the target face was angry rather than happy or neutral (attraction effect), and a less effective search when the distractor faces were angry rather than happy or neutral (distraction effect). By using ERPs, Burra, Barras, Coll, and Kerzel (2016) found that angry distractors in a neutral crowd triggered an N2pc (an electrophysiological marker of attentional selectivity 200–300 ms after stimuli onset). Together with the present results, these data suggest that the distraction effect of the angry faces on attention is short lived. That is, attention may be quickly drawn to the angry faces in the early stages of scene processing, yet once the faces are clearly identified as distractors, attention can be disengaged from them and not being repeatedly captured by the angry distractors in the long term.

Visual search tasks measure principally attention allocation immediately upon the face onset (Frischen et al., 2008). In such tasks, observers do not know where the targets will appear, thus they need to be vigilant for all the possible locations and objects for a short period. The angry faces may be more likely to win the competition for attention during this period because of their adaptive value (Desimone & Duncan, 1995; Yiend, 2010), even if they afterwards turn out to be distractors. Such a misdirection of attention may result in slow detection of the actual targets, and hence produce a distraction effect of angry distractors in visual search. On the contrary, the presently employed dynamic tracking task measures attention allocation in later stages, after the angry distractors have been identified as distractors. The angry distractors may quickly capture attention at an early stage of the tracking task (i.e. the target designation stage). Yet, the distractor faces are clearly marked as distractors for several seconds before the tracking starts, allowing time for withdrawing attention from the distractors and directing it to targets. In addition, research has shown that during tracking the distractors are usually ignored or suppressed (Pylyshyn, Haladjian, King, & Reilly, 2008). Accordingly, our results show that observers are able to ignore the angry faces as efficiently as the neutral face after they are identified as distractors, and hence there is no distraction effect from the angry distractors on sustained attention.

Considering together the target effect and the (lack of) distractor effect of angry faces in the present study and the literature, a tentative model can be proposed to account for the attentional processing of the angry faces across the processing timeline. The first step is a quick direction of attention to the angry faces when they appear, regardless of their relevance to the task at hand. The second step is to evaluate the relevance of the angry faces to the observer, and decide whether to continue their attentional processing. Although quickly detecting angry faces in the environment is adaptive (Öhman et al., 2012), continuous engaging attention with them may not always be necessary and beneficial (Cooper & Langton, 2006; Yiend, 2010). Thus, the third step is to maintain vigilance on the angry faces if they are evaluated as relevant to the observer and need to be continuously monitored, otherwise attention will be withdrawn and the angry faces will be ignored just as any other irrelevant objects. Systematic research on the entire process of attention–emotion interaction will be valuable in future studies.

Interestingly, both the upright and inverted angry targets were tracked better than neutral targets (Experiments 3 and 4). Presenting emotional faces upside-down has been used as a method for keeping the visual features of the faces intact while diminishing the configuration-based emotional information (Calvo & Nummenmaa, 2008; Eimer & Holmes, 2002; Horstmann & Bauland, 2006). In some visual search studies, the effects of emotional expressions on attention disappeared when the faces were presented upside-down (e.g. Burra et al., 2016; Fox & Damjanovic, 2006), which is contrary to the present results (for a review, see Nummenmaa & Calvo, 2015). We suggest that there may be two possible, yet not mutually exclusive explanations for the facilitatory effect of the inverted angry expressions on the tracking performance. First, some visual features of the angry expression (e.g. frowning eyebrows) facilitate sustained attention on the target faces when being presented both upright and upside-down. Second, emotional content of the inverted faces can be effectively processed in the tracking experiment, which in turn facilitates sustained attention on the faces. The latter possibility is supported in Experiment 3 by the high recognition rate for the angry expression of the inverted faces that was achieved fast.

In visual search or dot-probe paradigms there may be too little time for the inverted emotional expressions to be effectively processed and affect transient attention allocation (Burra et al., 2016). Yet, as the observers in the dynamic tracking task are allowed more time to continuously process the targets, the angry emotion of the inverted target faces may be processed to a larger extent (Derntl et al., 2009) and hence exert a substantial impact on attention. It is plausible that both the emotion-related visual features and the emotional content of the angry faces affected sustained attention in the present experiments, and in the real world they are likely to work in tandem (Horstmann & Bauland, 2006; Nummenmaa & Calvo, 2015). Some emotion-related visual features may have evolved to be easily recognised so as to convey the emotional content (Calvo & Nummenmaa, 2008).

While the present results consistently showed facilitatory effects of angry targets, the effect of the happy expression in the attentive tracking experiments seemed elusive. There was a significant distraction effect of happy distractor faces in Experiment 2 and a significant facilitatory effect of happy targets in Experiment 3b. There was also a trend for a distraction effect and a facilitation effect in Experiment 3a, yet neither of them was statistically significant. There was no effect of happy expressions in Experiment 4, where only target locations needed to be tracked. The local meta-analysis of the effect sizes yielded a small facilitatory effect of happy targets and a small distraction effect of happy distractors. Taken together, the results suggest that the happy expression may affect sustained attention in some circumstances (e.g. when observers are processing the identities of the faces), presumably due to the social reward value (Chakrabarti, Kent, Suckling, Bullmore, & Baron-Cohen, 2006; Phan, Wager, Taylor, & Liberzon, 2002) and the salient visual feature of happy faces (e.g. smiling mouth, Calvo & Nummenmaa, 2008); yet, such a "happiness superiority" effect on sustained attention is less robust than the "anger superiority".

Conclusions

We conclude that emotional information can be spontaneously processed during dynamic tracking and affect observer's sustained attention. Angry expressions on target faces can facilitate sustained attention on the targets via increased vigilance for them, whereas angry distractors do not impair sustained attention more strongly than neutral distractors. The effect of angry expressions persists when the faces are presented upside-down and when any surface features of the faces are irrelevant to the ongoing task. On the other hand, the effect of happy expressions on sustained attention is not as robust as the angry expressions; only a trend towards a facilitatory effect of happy targets and a distraction effect of happy distractors was observed.

Acknowledgements

We are grateful to Mowei Shen, Zaifeng Gao, and Fan Wu for helpful discussion.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Li was supported by the National Natural Science Foundation of China [grant number 31500887], and Li's post-doc year in Turku, Finland, was supported by the University of Turku. Oksama was supported by the Academy of Finland [grant number 273413]. Nummenmaa was supported by the MIND program grant, Academy of Finland [grant number 265915] and European Research Council [ERC Starting Grant #313000].

References

Agresti, A. (2007). *An introduction to categorical data analysis*. Hoboken, NJ: John Wiley & Sons.

Allen, R., & Gabbert, F. (2013). Exogenous social identity cues differentially affect the dynamic tracking of individual target faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1982–1989.

Arnon, I. (2010). Rethinking child difficulty: The effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language*, *37*(1), 27–57.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in single- and multiple-target visual search tasks. *Journal of Experimental Psychology: General*, *140*(4), 637–659.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Burra, N., Barras, C., Coll, S. Y., & Kerzel, D. (2016). Electrophysiological evidence for attentional capture by irrelevant angry facial expressions. *Biological Psychology*, *120*, 69–80.

Calvo, M. G., Fernández-Martín, A., & Nummenmaa, L. (2014). Facial expression recognition in peripheral versus central vision: Role of the eyes and the mouth. *Psychological Research*, *78*(2), 180–195.

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: Salient physical features guide effective visual search. *Journal of Experimental Psychology: General*, *137*(3), 471–494.

Calvo, M. G., Nummenmaa, L., & Avero, P. (2008). Visual search of emotional faces: Eye-movement assessment of component processes. *Experimental Psychology*, *55*(6), 359–370.

Chakrabarti, B., Kent, L., Suckling, J., Bullmore, E., & Baron-Cohen, S. (2006). Variations in the human cannabinoid receptor (CNR1) gene modulate striatal responses to happy faces. *European Journal of Neuroscience*, *23*(7), 1944–1948.

Cohen, M. A., Pinto, Y., Howe, P. D. L., & Horowitz, T. S. (2011). The what-where trade-off in multiple-identity tracking. *Attention Perception & Psychophysics*, *73*(5), 1422–1434.

Cooper, R., & Langton, S. (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.

Derntl, B., Seidel, E., Kainz, E., & Carbon, C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception*, *38*(12), 1849–1862.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics*, *63*(6), 1004–1013.

Ebner, N. C. (2008). Age of face matters: Age-group differences in ratings of young and old faces. *Behavior Research Methods*, *40*(1), 130–136.

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *NeuroReport*, *13*(4).

Feldmann-Wüstefeld, T., Schmidt-Daffy, M., & Schubö, A. (2011). Neural evidence for the threat detection advantage: Differential attention allocation to angry and happy faces. *Psychophysiology*, *48*(5), 697–707.

Fenske, M. J., & Eastwood, J. D. (2003). Modulation of focused attention by faces expressing emotion: Evidence from flanker tasks. *Emotion*, *3*(4), 327–343.

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion*, *6*(3), 534–539.

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin*, *134*(5), 662–676.

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation study. *Cognition & Emotion*, *22*(6), 1094–1118.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Krieger.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*, *54*(6), 917–924.

Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception & Psychophysics*, *69*(2), 172–184.

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion*, *6*(2), 193–207.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

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*(4), 379–395.

Li, J., Oksama, L., & Hyönä, J. (2016). How facial attractiveness affects sustained attention. *Scandinavian Journal of Psychology*, *57*(5), 383–392.

Liu, C. H., & Chen, W. (2012). Beauty is better pursued: Effects of attractiveness in multiple-face tracking. *The Quarterly Journal of Experimental Psychology*, *65*(3), 553–564.

Lundquist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces (KDEF)*. Department of Neurosciences, Karolinska Hospital, Stockholm, Sweden.

Makovski, T., & Jiang, Y. V. (2009a). Feature binding in attentive tracking of distinct objects. *Visual Cognition*, *17*(1–2), 180–194.

Makovski, T., & Jiang, Y. V. (2009b). The role of visual working memory in attentive tracking of unique objects. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1687–1697.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 630–633.

Nummenmaa, L., & Calvo, M. G. (2015). Dissociation between recognition and detection advantage for facial expressions: A meta-analysis. *Emotion*, *15*(2), 243–256.

Nummenmaa, L., Oksama, L., Glerean, E., & Hyönä, J. (in press). Cortical circuit for binding object identity and location during multiple object tracking. *Cerebral Cortex*. AQ10

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478.

Öhman, A., Soares, S. C., Juth, P., Lindström, B., & Esteves, F. (2012). Evolutionary derived modulations of attention to two common fear stimuli: Serpents and hostile humans. *Journal of Cognitive Psychology*, *24*(1), 17–32.

Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, *11*(5), 631–671.

Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive Psychology*, *56*(4), 237–283.

Oksama, L., & Hyönä, J. (2016). Position tracking and identity tracking are separate systems: Evidence from eye movements. *Cognition*, *146*, 393–409.

Papenmeier, F., Meyerhoff, H. S., Jahn, G., & Huff, M. (2014). Tracking by location and features: Object correspondence across spatiotemporal discontinuities during multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 159–171.

Pashler, H. (1998). *The psychology of attention*. Cambridge: MIT Press.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*(2), 331–348.

Pinto, Y., Howe, P., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. *Journal of Vision*, *10*(10), 4, 1–15.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65–97.

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, *11*(7), 801–822.

Pylyshyn, Z. W., Haladjian, H. H., King, C. E., & Reilly, J. E. (2008). Selective nontarget inhibition in multiple object tracking. *Visual Cognition*, *16*(8), 1011–1021.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Schmidt-Daffy, M. (2011). Modeling automatic threat detection: Development of a face-in-the-crowd task. *Emotion*, *11*(1), 153–168.

Scholl, B. J. (2009). What have we learned about attention from multiple-object tracking (and vice versa). In D. Dedrick & L. Trick (Eds.), *Computation, cognition, and pylyshyn* (pp. 49–77). Cambridge: MIT Press.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? [Abstract]. *Investigative Ophthalmology and Visual Science*, *40*(4), S797. AQ11

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120.

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition & Emotion*, *24*(1), 3–47.

Appendix

Table A1. Summary of the effects in the mixed logit model for correct response ($N = 1872$; log-likelihood = −1201.9).

	Fixed effects			
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	−0.37	0.181	−2.06	.040
Face discriminability	0.30	0.055	5.52	<.001
Target expression = Angry	0.27	0.125	2.13	.034
Target expression = Happy	0.21	0.120	1.80	.073
Distractor expression = Angry	−0.07	0.121	−0.57	.566
Distractor expression = Happy	−0.18	0.121	−1.48	.139
Random effects				
	s^2			
Intercept	0.2083			

Note: Angry and happy expressions are compared against neutral expressions.