



**HAL**  
open science

# At Face Value: Categorization Goals Modulate Vigilance for Angry Faces

Lotte F. van Dillen, Daniël Lakens, Kees Van den Bos

► **To cite this version:**

Lotte F. van Dillen, Daniël Lakens, Kees Van den Bos. At Face Value: Categorization Goals Modulate Vigilance for Angry Faces. *Journal of Experimental Social Psychology*, 2010, 47 (1), pp.235. 10.1016/j.jesp.2010.10.002 . hal-00918796

**HAL Id: hal-00918796**

**<https://hal.science/hal-00918796v1>**

Submitted on 15 Dec 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Accepted Manuscript

At Face Value: Categorization Goals Modulate Vigilance for Angry Faces

Lotte F. Van Dillen, Daniël Lakens, Kees van den Bos

PII: S0022-1031(10)00219-2  
DOI: doi: [10.1016/j.jesp.2010.10.002](https://doi.org/10.1016/j.jesp.2010.10.002)  
Reference: YJESP 2545

To appear in: *Journal of Experimental Social Psychology*

Received date: 2 November 2009  
Revised date: 11 August 2010



Please cite this article as: Van Dillen, L.F., Lakens, D. & den Bos, K., At Face Value: Categorization Goals Modulate Vigilance for Angry Faces, *Journal of Experimental Social Psychology* (2010), doi: [10.1016/j.jesp.2010.10.002](https://doi.org/10.1016/j.jesp.2010.10.002)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Running head: Categorizing Angry Faces

At Face Value:

Categorization Goals Modulate Vigilance for Angry Faces

Lotte F. Van Dillen, Daniël Lakens, and Kees van den Bos

Utrecht University

Word count: 5000

Author Notes:

Part of this research was supported by a VICI innovational research grant from the Netherlands Organization for Scientific Research (NWO, 453.03.603) awarded to Kees van den Bos.

We thank Lasana Harris and Kaska Kubacka for their valuable comments on an earlier draft of this manuscript.

Address correspondence to Lotte F. van Dillen, Institute for Psychological Research, Social and Organizational Psychology Unit, The Netherlands, P.O. Box 9555, 2300 RB Leiden, the Netherlands, E-mail: [Dillenfvan@fsw.leidenuniv.nl](mailto:Dillenfvan@fsw.leidenuniv.nl), Phone: +31 71 5273831, Fax: +31 71 5273619.

## Abstract

The present research demonstrates that the attention bias to angry faces is modulated by how people categorize these faces. Since facial expressions contain psychologically meaningful information for social categorizations (i.e., gender, personality) but not for nonsocial categorizations (i.e., eye-color), angry facial expressions should especially capture attention during social categorization tasks. Indeed, in three studies, participants were slower to name the gender of angry compared to happy or neutral faces, but not their color (blue or green; Study 1) or eye-color (blue or brown; Study 2). Furthermore, when different eye-colors were linked to a personality trait (introversion, extraversion) versus sensitivity to light frequencies (high, low), angry faces only slowed down categorizations when eye-color was indicative of a social characteristic (Study 3). Thus, vigilance for angry facial expressions is contingent on people's categorization goals, supporting the perspective that even basic attentional processes are moderated by social influences.

**KEYWORDS:** Social categorization, Automatic vigilance, Facial Expression

Word count: 147

## At Face Value:

## Categorization Goals Modulate Vigilance for Angry Faces

Angry facial expressions draw attention more readily than happy or neutral expressions (Eastwood, Smilek, & Merikle, 2003; Fox, Russo, & Dutton, 2002; Holmes, Bradley, Nielsen, & Mogg, 2009). Various attentional paradigms have thus demonstrated the enhanced ability of threatening stimuli to guide attention to themselves or their ability to hold attention, once captured. For example, people are more easily distracted by an angry face than by a happy or neutral face (Erthal, De Oliviera, Mocaiber, Pereira, Machado-Pinheiro, & Volchan, 2005; Van Honk, Tuiten, De Haan, Van den Hout, & Stam, 2001), and people take longer to count the features of angry faces than of happy or neutral faces (Eastwood et al., 2003). To explain this vigilance for angry faces, theorists have proposed that natural selection has equipped people to efficiently screen the environment for potential threats (Anderson & Phelps, 2001; Bradley, 2009; Öhman, 2007).

Although it may be adaptive to prioritize attention to threatening information, people may not invariably respond to angry faces in the environment. For instance, at a crowded concert, you may fail to notice the angry face of a person who's toes you just stepped on because you are too busy finding your way to the bar. Similarly, over dinner, you may not pick up on your spouse's grumpy mood because you are still contemplating over a problem at work. We therefore propose that whereas angry faces can capture and/or hold attention, angry expressions need not always bias the way faces are processed. Instead, vigilance for angry faces may be subject to contextual variations in accord with people's current activities.

The type of stimulus-driven perceptual processing that underlies vigilance for angry faces is known as bottom-up attention (Desimone & Duncan, 1995; Yantis, 2000). Bottom-up attention involves information selection on the basis of salient stimulus features (i.e., novelty, predictability, and threat) that are likely to be important for adaptive behavior (Bradley, 2009). Attention can also be controlled by top-down processes (Corbetta & Shulman, 2002; Kiefer & Martens, 2010; Knudsen, 2007; Lavie, 2005), which direct attention in accord with people's goals and actions. Especially when salient features are irrelevant for task performance, top-down attention control may attenuate the processing of these features (Dreisbach & Haider, 2009; Folk, Remington, & Johnston, 1992; Knudsen, 2007).

In support of this notion, several studies suggest that top-down control processes moderate the role of bottom-up attention filters in processing salient, but task-irrelevant information (Chong et al., 2008; Dreisbach & Haider, 2009; Klauer & Musch, 2002; Spruyt, De Houwer, Hermans, & Eelen, 2007; Van Dillen & Koole, 2009). For example, loading people's mental capacity with a focal task decreases attentional interference of angry faces (Erthal et al., 2005; Holmes, Vuilleumier & Elmer, 2003; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002). In one neuroimaging study by Anderson, Christoff, Panitz, De Rosa, and Gabrieli (2003), participants focused on either faces or houses presented in a single overlapping display. The results revealed that neural responses to disgust and fear expressions were reduced or less specific when participants were focusing on the house rather than the face.

Based on the findings briefly reviewed here, we propose that responses to angry faces can be controlled in a top-down fashion in accordance with contextual variations in

task representation. More specifically, we hypothesize that categorization goals moderate vigilance for angry faces by directing the way in which people process these stimuli. Categorization goals guide information processing by directing attention toward the critical response discriminating features of the stimuli (Dreisbach & Haider, 2009). If salient stimulus features are part of the mental representation of a category, the attention system will be more susceptible to these features (Dreisbach & Haider, 2009; Folk, 1994; Koivisto & Revonsuo, 2007). Accordingly, we propose that people should respond to angry versus happy or neutral faces especially when these facial features are meaningful for the activated category representation.

To investigate this hypothesis, we compared peoples' responses to angry and happy or neutral faces during a task that required people to categorize faces on social dimensions (i.e. gender or personality) for which expression valence is a psychologically meaningful feature, versus a task that required people to categorize faces on strictly physical aspects (i.e. stimulus color) for which expression valence conveys no meaningful information.

Facial expressions play a pivotal role in social categorization processes as they communicate information about the person's intentions (Hess, Sabourin, & Kleck, 2007; Hugenberg & Bodenhausen, 2003, 2004). Through social learning and/or stereotyping, people may come to associate certain emotional expressions with certain social categories. For example, people more likely perceive a happy face as female and an angry face as male (Becker, 2007). Similarly, racial prejudice is more strongly associated with a tendency to categorize hostile (but not happy) racially ambiguous faces as African American (Hugenberg & Bodenhausen, 2004). Mental representations of nonsocial

categories, such as identifying blue from green eyes, have no intrinsic psychological meaning, in that they do not allow any inferences or expectations about people's motivations. For such categories, the information contained in facial expressions has little value. Facial expressions should therefore less likely be incorporated in the mental representations of these categories.

To summarize, vigilance mechanisms should be in place when categorization instructions trigger the processing of facial expressions, such as during social categorization tasks that allow for psychological inferences about individuals belonging to these categories. Angry facial expressions should less likely bias responses, however, during categorization tasks that solely rely on physical features, and for which facial expressions have no informative value.

#### *The present research*

We designed the present experiments to investigate the above-mentioned predictions using a speeded categorization task (see Van Dillen & Koole, 2009, for a similar paradigm). In this task, participants categorize, as quickly as possible, a series of faces. Typically, people are thought to display vigilance for angry faces when they are slower to categorize angry faces compared to happy (Study 1) or neutral (Studies 2 and 3) faces, because attention is automatically drawn to and/or held longer by the angry facial expression at the cost of task relevant features (Van Dillen & Koole, 2009; Van Honk et al., 2001). We argue, however, that when the categorization goal does not require facial expressions to be processed, this relative slow-down should not occur.

We manipulated category type in all three studies, such that it would trigger the processing of facial expressions or not. In Studies 1 and 2 participants categorized the



faces either as belonging to men or women or as being presented in blue or green (color-categorization task of Study 1) or as having blue or brown eyes (color-categorization task of Study 2). Facial expressions are irrelevant to the gender-naming task, in the sense that facial expressions are not truly indicative of gender. However, because facial expressions are part of people's mental representation of gender (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Hess, Adams, & Kleck, 2004), people may continue to process expression valence when they categorize the faces on the basis of gender. In contrast, when people categorize the faces as being either blue or green, or as having either blue or brown eyes, expression valence should have no informative value. Thus, we expected people to take longer to name the gender of angry compared to happy or neutral faces, because of vigilance for threatening faces (Bradley, 2009; Öhman, 2007), whereas we did not expect strong response time differences when people categorized the faces as being blue or green (Study 1) or having blue or brown eyes (Study 2).

To provide an even more stringent test of our hypothesis, in our third study, all participants categorized faces on their eye-color. Importantly, we manipulated whether eye-color indicated a personality trait or not. Using this approach, participants always attended to the same feature, regardless of the activated category type. Accordingly, any effects of category type cannot be attributed to different visual processing styles (i.e. integrating various features versus focusing on one feature). Rather, any effects of expression valence on task performance should arise because participants have learned to associate a physical feature (eye-color) with a social category (extraversion) for which expression valence is psychologically meaningful.

Participants read a bogus scientific text in which half of the participants were informed that eye-color indicated the personality trait of extraversion, whereas the other half of the participants were informed that eye-color indicated the light frequency that is absorbed by the eye. In the task that followed, participants used this knowledge to categorize angry and neutral blue-eyed and brown-eyed faces. We predicted that participants would be slower to categorize angry compared to neutral faces when the eye-color represented social information (i.e., introverted or extraverted people), because expression valence is a psychologically meaningful feature for the category of extraversion. We did not expect such a slow-down when the eye-color represented merely physical information (i.e. absorbing either high or low frequency light), because in this case expression valence has no informative value.

## Study 1

### *Method*

#### *Participants and Design*

Forty-five paid volunteers at the VU University Amsterdam (25 women, 20 men, average age 20) took part in the experiment. The experimental design was a 2 (target expression: happy versus angry; within participants) x 2 (target gender: male versus female; within participants) x 2 (category type: gender versus eye-color; between participants) design. The main dependent variable consisted of participants' response times (in ms) in the categorization task.

#### *Procedure and Equipment*

Upon arrival in the laboratory, participants were led to individual cubicles with a personal computer. The experimenter explained that all instructions would be

administered via a computer-program and left. After a brief introduction, participants proceeded with a speeded categorization task in which they categorized pictures of either blue or green male or female faces displaying either a happy or an angry expression. The faces were drawn from the The Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt, & Öhman, 1998). We selected pictures of ten individuals (five men and five women) facing directly into the camera and displaying either a happy or angry expression. Of these faces, we created a blue and a green version by overlaying a semi-transparent filter, which matched the contour of the faces (for blue: RGB = 0,0,255; and for green: RGB = 0,255,0; transparency 30%; see Appendix 1). Accordingly, the total set consisted of forty pictures; five male and five female faces, displaying either a happy or angry expression, with either a blue or green filter. Importantly, to ascertain that participants could not rely on single features to categorize the gender of the faces such as for example hair length, or facial hair, we included male and female faces with both long and short hair and all male faces were shaven.

The categorization task consisted of 40 trials. Before the 40 experimental trials, participants first received four practice trials to become familiar with the task. Each trial was announced by a row of four asterisks (\*\*\*\*), which remained in the center of the screen for one second. During each trial, a picture of either a blue or green angry or happy male or female face appeared on screen for 2 seconds. In the gender-naming condition ( $N = 22$ ), the participants had to decide as quickly as possible, by making a keyboard response (the a and b key; counterbalanced between participants), whether the face on the screen was male or female. In the color-naming condition ( $N = 23$ ), the participants had to decide as quickly as possible whether the face on the screen was blue

or green. The computer recorded participants' responses and response times to the categorization task unobtrusively. At the end of the experimental trials, participants were thanked, debriefed, and paid by the experimenter.

### *Results*

Incorrect responses to the categorization task represented only 4% of the trials, and were excluded from subsequent analyses. Descriptives analyses revealed no outliers ( $> 3$  SD's) or skewness of the response time data.

To analyze participants' performance on the categorization task, we conducted a 2 (target expression) x 2 (target gender) x 2 (category type) analysis of variance of participants' response times. We found a marginally significant main effect of target expression,  $F(1, 43) = 3.60, p = .065, \eta_p^2 = .08$ . Moreover, we found the predicted interaction effect of category type and target expression,  $F(1, 43) = 21.85, p < .001, \eta_p^2 = .33$ . Importantly, there was no main effect of category type on participants' response times,  $F(1, 43) < 1$ , indicating that any effects of target expression could not be attributed to differences in viewing times between the gender and the color categorization condition. Also, there were no effects for target gender on performance<sup>1</sup>.

More focused comparisons only yielded a simple main effect of target expression for the participants who judged the gender of the targets,  $F(1, 43) = 9.27, p = .006, \eta_p^2 = .30$ . These participants responded more slowly to angry faces ( $M = 908$  ms,  $SD = 150$  ms) than to happy faces ( $M = 839$  ms,  $SD = 100$  ms). We found no effect of target expression among the participants who judged the color of the faces,  $F < 1$ . In the color-naming condition, participants responded equally quickly to angry and happy faces ( $M = 848$  ms,  $SD = 141$  ms and  $M = 863$  ms,  $SD = 156$  ms, respectively).

### *Discussion*

As expected, judging the targets on the basis of color rather than gender eliminated the greater attentional interference of angry relative to happy faces. Because participants always judged the same targets, the effects could not be attributed to differential target characteristics. Moreover, participants took as long to categorize the faces on the basis of gender, as they did to categorize the faces on the basis of color, suggesting that the effects of category type could not be explained by differences in viewing time. This noted, another way to explain the absence of attentional interference in the color naming task is that participants in this condition may have directed their gaze away from emotional features of the target, for example by looking at the contour of the faces, or by looking through ones eyelashes (Dunning & Hajcak, 2009; Van Reekum et al, 2007). This alternative explanation was addressed in Study 2.

### Study 2

Study 2 was designed to replicate and extend the findings of Study 1. Rather than manipulating the color of the entire face, Study 2 only varied whether the stimulus persons had blue or brown eyes. By varying eye-color, participants have to focus on the eyes, which constitute a central role in the communication of emotions (Baron-Cohen, Wheelwright & Jolliffe, 1997; Emery, 2000). Recent eye-tracking findings, for example, suggest that for angry expressions, people mostly focus at the eyes (Aviezer, Hassin, Ryan, Grady, Susskind, Anderson, Moscovitch, Bentin, 2009). Accordingly, participants cannot adopt gaze strategies that avoid the processing of emotionally relevant facial features. If the effects of category type in Study 1 were due to a differential processing style, as the present analysis suggests, then attentional interference of angry expressions

should be greater when participants categorize the faces on the basis of gender rather than on the basis of blue versus brown eyes. On the other hand, if the results of Study 1 were simply due to emotion avoidant gaze patterns in the color condition, then categorizing on the basis of gender and eye-color should both result in attentional interference of angry faces.

### *Method*

#### *Participants and Design*

Forty-seven paid volunteers at Utrecht University (26 women, 19 men, average age 21 years) took part in the experiment. The experimental design was a 2 (target expression: neutral versus angry; within participants) x 2 (target gender: male versus female; within participants) x 2 (category: eye-color versus gender; between participants) design. The main dependent variable consisted of participants' response times (in ms) to the categorization task.

#### *Procedure and Equipment*

The procedure was similar to Study 1. Participants performed the speeded categorization task in which this time half of the participants was instructed to identify the faces as either belonging to men or women ( $N = 23$ ) whereas the remaining half was instructed to identify the faces as either having blue or brown eyes ( $N = 24$ ). The same set of faces was used as in Study 1. We varied eye-color by overlaying a semitransparent filter (transparency = 15%), such that half of the faces had blue eyes (RGB = 0,102,204), and half of the faces had brown eyes (RGB = 204,102,0; see Appendix 2). The total stimulus set contained forty pictures; five male and five female faces, displaying either a neutral or angry expression, with either blue or brown eyes.

### Results

To analyze participants' performance on the categorization task, we conducted a 2 (target expression) x 2 (target gender) x 2 (category type) General Linear Model analysis of participants' response times. Incorrect responses on the categorization task (3% of all responses) were excluded from the data. Analyses revealed no outliers ( $> 3$  SD's) or skewness of the response time data.

We found a main effect of target expression,  $F(1, 45) = 7.79, p = .007, \eta_p^2 = .15$ . Participants responded more slowly to angry faces ( $M = 666$  ms,  $SD = 160$  ms) than to neutral faces ( $M = 643$  ms,  $SD = 145$  ms). There were no effects of target gender<sup>1</sup> or category type on participants' performance. As in Study 1 we also found a two-way interaction between category type and target expression,  $F(1, 45) = 6.22, p = .016, \eta_p^2 = .14$ . More focused comparisons revealed an effect of target expression for the participants who judged the gender of the targets,  $F(1, 45) = 20.67, p < .000, \eta_p^2 = .34$ . These participants responded more slowly to faces displaying anger ( $M = 672$  ms,  $SD = 185$  ms) rather than a neutral expression ( $M = 626$  ms,  $SD = 154$  ms). We found no effect of target expression among the participants who judged the eye-color of the faces,  $F < 1$ , such that these participants responded equally fast to angry faces ( $M = 659$  ms,  $SD = 134$  ms) as to neutral faces ( $M = 661$  ms,  $SD = 137$  ms).

### Discussion

The results of Study 2 reveal that categorization goals moderate attentional vigilance for angry faces, such that people were slower to identify the gender of angry compared to neutral faces but were equally fast when they identified the eye-color of the faces. By having participants focus on the eyes in Study 2, we demonstrated that this

attenuation of vigilance for angry faces could not be explained by emotion avoidant gaze patterns during the color-categorization task.

### Study 3

Study 3 was designed to replicate and extend the findings of Studies 1 and 2. Rather than manipulating the features that participants attended to, we varied whether one particular feature (eye-color) indicated a personality trait or reflected a physical property. Previous research suggests that people tend to incorporate knowledge about facial expressions when assessing an individual on a certain personality trait (Knutson, 1996; Said, Sebe, & Todorov, 2009). Hence, we reasoned that when categorizing a personality trait, emotional expression would likely bias participants' responses. The experimental set-up was almost identical to the set-up of Study 2. However, this time participants were informed that eye-color indexed the personality trait of extraversion versus introversion, or a physical property, namely the light frequency being absorbed by the iris of the eye. Accordingly, participants always used the same feature to categorize the faces, but the feature either referred to a category for which expression valence was psychologically meaningful (the personality trait of extraversion) or for which expression valence had no informative value (the physical trait of light frequency). If the effects of category type in the previous two studies were the result of a differential processing style, as we propose, then attentional interference of angry faces should be greater when participants categorize the faces on the basis of a personality trait rather than a physical property. On the other hand, if the results of Studies 1 and 2 were mainly the result of the different features that participants had to attend to in order to categorize on the basis of gender or eye-color,



then categorizing blue versus brown eyes should both not result in attentional interference of angry faces, regardless of the meaning of the eye-color.

### *Method*

#### *Participants and Design*

Seventy-six paid volunteers at Utrecht University (45 women, 31 men, average age 21 years) took part in the experiment. The experimental design was a 2 (target expression: neutral versus angry; within participants) x 2 (eye-color: brown versus blue; within participants) x 2 (category type: light frequency versus extraversion; between-participants) design. The main dependent variable consisted of participants' response times to the categorization task.

#### *Procedure and Equipment*

In a brief introduction, participants were informed that the experiment was about how people incorporate new information. Participants were told to carefully read a scientific text and that the information in the text would be used in the following task. For all participants, the text described how the gene OCA2, at chromosome 15, coded for eye-color. Next, for half of the participants, the text reported how recent scientific findings suggest that this gene also relates to the personality trait of extraversion (the social category manipulation), such that blue-eyed individuals are more likely to be extraverted and brown-eyed individuals to be introverted (counterbalanced between participants). Extraversion was explained as being more responsive to sensory stimulation, more energetic, and assertive. Introversion was explained as being more inward focused, sensitive, and observing. The full description can be found in Appendix

3. We on purpose avoided linking any specific emotions to either extraversion or introversion, such that angry expressions would not directly bias participants' responses.

The other half of the participants read that the gene codes the light frequency absorbed by the eye (the nonsocial category manipulation), such that blue eyes absorb more high frequency light and brown eyes more low frequency light (counterbalanced between participants). Thus, half of the participants read that eye-color was indicative of extraversion, and half read that it was indicative of light frequency. Next, all participants were instructed to use this newly obtained information in a categorization task of angry and neutral blue-eyed and brown-eyed male and female faces. The same set of faces was used as in Study 2. The same speeded categorization task was used as in the previous studies, in which this time half of the participants categorized the faces as being either extraverted or introverted ( $N = 44$ ) or as having eyes absorbing high or low frequency light ( $N = 32$ ).

### *Results*

To analyze participants' performance on the categorization task, we conducted a 2 (target expression) x 2 (eye-color) x 2 (category type) General Linear Model analysis of participants' response times. Incorrect responses on the categorization task (2% of all responses) were excluded from the data. Response times were not skewed. To reduce the influence of outliers, we excluded reaction times (one trial) that exceeded three standard deviations from the mean.

We found a main effect of target expression,  $F(1, 74) = 7.32, p = .008, \eta_p^2 = .09$ . Participants were slower to respond to angry faces ( $M = 697$  ms,  $SD = 108$  ms) than to neutral faces ( $M = 677$  ms,  $SD = 106$  ms). There were no effects of eye-color on

performance<sup>1</sup>. We again found a two-way interaction between category type and target expression,  $F(1, 74) = 4.64$ ,  $p = .034$ ,  $\eta_p^2 = .06$ . More focused comparisons revealed an effect of target expression for participants who used eye-color to judge whether the targets were extraverted or introverted,  $F(1, 74) = 13.28$ ,  $p < .001$ ,  $\eta_p^2 = .16$ . These participants responded more slowly to faces displaying anger ( $M = 713$  ms,  $SD = 103$  ms) compared to a neutral expression ( $M = 679$  ms,  $SD = 93$  ms). We found no effect of target expression among the participants who used eye-color to judge the light frequency absorbed by the eyes,  $F < 1$ , such that these participants responded equally fast to angry faces ( $M = 675$  ms,  $SD = 114$  ms) as to neutral faces ( $M = 672$  ms,  $SD = 120$  ms).

### *Discussion*

In a third and final study, we performed a more stringent test of the idea that categorization goals moderate vigilance for angry faces. Rather than varying the features that participants needed to attend to in order to perform the categorization task, we varied the psychological meaning of a feature. As such the feature of eye color could either refer to the socially relevant personality trait of extraversion, or to the physical aspect of light frequency. When eye color acquired psychological meaning (was indicative of extraversion) participants were slower to respond to angry than to neutral faces. When eye color had no psychological meaning, but instead referred to the light frequency absorbed by the eye, participants were as fast in response to angry as to neutral faces in the categorization task. Hence, the current findings could provide insight into how stereotypes emerge. When people learn to associate physical features (such as eye-color) with a social category, these features can become psychologically meaningful, and can subsequently trigger biased processing of emotional expressions.

*General Discussion*

Across three studies, the present research demonstrates that people are more likely to process angry facial expressions when categorization instructions trigger processing of facial expressions, such as during gender or personality categorization tasks, whereas angry facial expressions do not slow down responses when they are meaningless to the activated category, such as during color-categorization tasks. In Study 1, participants were slower to categorize the gender of angry than of happy faces, whereas this slowdown did not occur for participants who categorized these faces on color. Studies 2 and 3 replicated these findings, and showed that this slowdown disappeared even when participants focused on the eye-color of the faces during the eye-color-categorization task. Given that the eyes are thought to play a central role in conveying emotional information (Baron-Cohen, et al., 1997; Emery, 2000), it is unlikely that the findings of Study 1 could therefore be attributed to avoidant gazing styles in the color-naming condition. Moreover, people were slower to name the gender of angry faces compared to both happy and neutral faces, suggesting that this relative slow-down in response times could not be attributed to increased vigilance to happy faces in the color-naming condition. Finally, in Study 3, all participants focused on just one feature (eye-color). The results revealed that when participants learned that eye-color indicated the personality trait of extraversion, the slow-down to angry faces occurred, likely because expression valence is a meaningful dimension for this category. However, when participants were told eye-color reflected a strictly physical feature (the light frequency that is absorbed by the eye), no slow-down for angry faces was observed, because expression valence in this case was not related in any meaningful way to the activated category.

Conceptually the present findings are important, we think, as they provide further support for the growing notion that responses to (threatening) social information can be controlled in accord with specific task parameters (Blair, 2001; Fiske, 1998; Neumann 1984; see also Barrett, Mesquita, & Smith, 2009). Indeed, response biases are in place when categorization instructions trigger the processing of facial expressions, such as during categorization tasks of gender or personality that call for psychological inferences. However, when categorization instructions involve a processing style for which emotional expressions provide no informative value, such as during the categorization of blue versus brown eyes, or the categorization of high versus low light frequency, angry expressions no longer interfere with the performance on the focal task. Accordingly, our findings provide further evidence that the likelihood with which people respond to (threatening) information in their environment depends on their current goals (Gu & Han, 2007; Hajcak, Moser, & Simons, 2006; Van Dillen & Koole, 2009).

Earlier research (Becker et al. 2007) has shown that gender biases people's responses to angry expressions, such that people are especially slow to categorize female angry expressions. This finding was presumably observed because of a natural confound between the features that correspond to maleness and anger on the one hand, and femaleness and happiness on the other. No such interaction between gender and valence was observed in our Studies 1 and 2. Such an interaction possibly requires highly prototypical male and female faces, such as the computerized faces used by Becker and colleagues (2007), whereas we used real male and female faces that varied substantially in their gender typicality (i.e. women with short hair, males with long hair, etc.).

Possibly, using highly prototypical male and female faces would have resulted in an interaction between gender and expression.

Similarly, Study 3 did not reveal either extraversion or introversion to enhance the slow-down to angry faces. Such an interaction should be observed when introversion or extraversion is strongly associated with anger. Naturally, in our description of extraversion and introversion we purposefully avoided the mention of any specific emotions. Perhaps then, participants had no clear-cut predictions about which personality type would be more or less likely to express anger. Assessing people's lay theories about the correlation between certain social categories or traits and specific facial expressions may therefore further qualify our current findings.

In the present paper, we demonstrated how the implementation of a social versus non-social category moderates processing of angry expressions. An important question is which categorization goal reflects the default processing mode. Does implementing a non-social category inhibit further processing of angry faces, or the extraction of expression valence altogether? On the one hand, it has been proposed that in interpersonal situations, people automatically pick up on the emotions of others (Schilbach, Eickhoff, Mojzisch, & Vogeley, 2008), possibly through the automatic imitation of these expressions (e.g., Niedenthal, Winkielman, Mondillon, & Vermeulen, 2009), and that overcoming these responses involves some cognitive control process (Spengler, Brass, Kühn & Schütz-Bosbach, 2010). On the other hand, by definition, any interaction between people takes place in a certain context, and we know that context already shapes early perceptual processes, i.e. determine which features are extracted from the environment in the first place (Chong et al., 2008; Kiefer & Martens, 2010; Liu,

Slotnick, Serences & Yantis, 2003). Whereas most of this research has mainly focused on early attention to low level features such as color or movement, category type may similarly gate early attention to more complex stimuli such as emotional expressions. Future research involving more direct measures of attention processing (e.g, online neurophysiological measures) is needed to further examine this relevant issue.

Whereas the current research focused on the role of categorization processes in basic responses to angry faces, categorization processes may similarly moderate the impact of emotional cues on more complex interpersonal judgments and behavior. For example, people may be guided more by emotional cues in their decisions when they have a social rather than a non-social goal. Given that the unfolding of an emotional response begins with the attentional capture of emotional cues (Gross, 2005; Pessoa, 2008), categorization processes may control the impact of emotional cues on people's thoughts and behavior right at the pass. Furthermore, these findings mirror recent studies, which reveal the top-down moderation of automatic processes in other areas of social psychology, such as the attentional moderation of automatic imitation (Leighton, Bird, Orsini & Heyes, in press; Likowski, Mühlberger, Seibt, Paulia, & Weyers, 2007; Longo & Bertenthal, 2009).

In conclusion, the present research demonstrates that people not invariably slow down their responses when categorizing angry facial expressions, but that this depends on people's current categorization goal. Accordingly, the present findings provide further support to the emerging notion that even basic attentional processes are moderated by social factors. Thus, on the face of it, an angry expression may either stand out, or remain in the background.

*References*

- Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*, 305-309.
- Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. E. (2003). Neural correlates of the automatic processing of threat facial signals. *The Journal of Neuroscience*, *23*, 5627-5633.
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2009). Angry, disgusted or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*, 724-732.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (1997). Is there a “language of the eyes”? Evidence from normal adults, and adults with autism or asperger syndrome. *Visual Cognition*, *4*, 11-31.
- Barrett, L. F., Mesquita, B., & Smith, E. R. (2010). The context principle. In Mesquita, B., Barrett, L. F., & Smith, E.R. (Eds.). *The Mind in Context* (pp. 1-21). New York: Guilford Press.
- Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*, 179-190.
- Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*, 1142-1163.
- Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*, 1-11.



- Chong, H., Riis, J. L., McGinnis, S. M., Williams, D. M., Holcomb, P. J., & Daffner, K. R. (2008). To ignore or explore: Top-down modulation of novelty processing. *Journal of Cognitive Neuroscience*, *20*, 120-134.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201-215.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193-222.
- Dreisbach, G., & Haider, H. (2009). How task representations guide attention: Further evidence for the shielding function of task sets. *Journal of Experimental Psychology-Learning Memory and Cognition*, *35*, 477-486.
- Dunning, J. P., & Hajcak, G. (2009). See no evil: Directing visual attention within unpleasant images modulates the electrocortical response. *Psychophysiology*, *46*, 28-33.
- Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics*, *65*, 352-358.
- Emery, N. J. (2002). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*, 581-604
- Erthal, F. S., De Oliviera, L., Mocaiber, I., Pereira, M. G., Machado-Pinheiro, W., & Volchan, E. (2005). Load-dependent modulation of affective picture processing. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 388-395.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert & S. T.

- Fiske (Eds.), *The handbook of social psychology*, Vol. 2 (4th ed., pp. 357-411).  
New York: McGraw-Hill.
- 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*, 1030-1044.
- Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion*, *16*, 355-379.
- Hajcak, G., Moser, J. S., & Simons, R. F. (2006). Attending to affect: Appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*, 517-522.
- Hess, U., Adams, R. B. Jr. & Kleck, R. E. (2004). Facial appearance, gender, and emotion expression. *Emotion*, *4*, 378-388.
- Hess, U., Sabourin, G., & Kleck, R. E. (2007). Postauricular and eyeblink startle responses to facial expressions. *Psychophysiology*, *44*, 431-435.
- Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: Evidence from event-related brain potentials. *Cognitive Brain Research*, *16*, 174-184.
- Hugenberg, K. & Bodenhausen, G. V. (2004). Category membership moderates the inhibition of social identities. *Journal of Experimental Social Psychology*, *40*, 233-238.
- Klauer, K. C., & Musch, J. (2002). Goal-dependent and goal-independent effects of irrelevant evaluations. *Personality and Social Psychology Bulletin*, *28*, 802-814.

- Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience, 30*, 57-78.
- Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20*, 165-182.
- Koivisto, M., & Revonsuo, A. (2007). How meaning shapes seeing. *Psychological Science, 18*, 845-849.
- Lavie, N., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review, 12*, 669-674.
- Liu, T., Slotnick, S. D., Serences, J. S., Yantis, S. (2003) Cortical mechanisms of feature-based attentional control. *Cerebral Cortex, 13*, 1334-1343.
- Longo, M. R., & Bertenthal, B. I. (2009). Attention modulates the specificity of automatic imitation to human actors. *Experimental Brain Research, 192*, 739-744.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces (KDEF)*. Department of Neurosciences, Karolinska Hospital, Stockholm.
- Mitchell, J. (2009). Social psychology as a natural kind. *Trends in Cognitive Science, 13*, 246-251.
- Neumann, O. (1984). *Automatic processing: A review of recent findings and a plea for an old theory*. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 255-293). Berlin, Germany, Springer.
- Niedenthal, P. M., Winkielman, P., Mondillon, L., & Vermeulen, N. (2009). Embodied emotion concepts. *Journal of Personality and Social Psychology, 96*, 1120-1136.

- Öhman, A. (2007). Has evolution primed humans to "beware the beast"? *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 16396-16397.
- Pessoa, L., (2008). On the relationship between emotion and cognition. *Nature Review of Neuroscience*, *9*, 148–158.
- Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 11458\_11463.
- Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, *9*, 260 – 264
- Schilbach, L., Eickhoff, S. B., Mojzisch, A., & Vogeley, K. (2008). What's in a smile? Neural correlates of facial embodiment during social interaction. *Social Neuroscience*, *3*, 37-50.
- Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology*, *54*, 44–53.
- Van Dillen, L. F., & Koole, S. L. (2009). How automatic is “automatic vigilance”? The role of working memory in attentional interference of negative information. *Cognition and Emotion*, *23*, 1106 - 1117.
- Van Honk, J., Tuiten, A., De Haan, E., Van den Hout, M., & Stam, H. (2001). Attentional biases for angry faces: Relationships to trait anger and anxiety. *Cognition & Emotion*, *15*, 279–297.

Yantis, S. (2000). *Goal-directed and stimulus-driven determinants of attentional control*.

In S. Monsell & J. Driver (Eds.), *Attention and performance* (Vol. 18, 73-103).

Cambridge, MA: MIT Press.

ACCEPTED MANUSCRIPT

## Footnotes

<sup>1</sup> Because we did not find any effects of gender in Study 1 and 2, or eye-color in Study 3, we collapsed responses across male and female faces and across faces with blue and brown eyes.

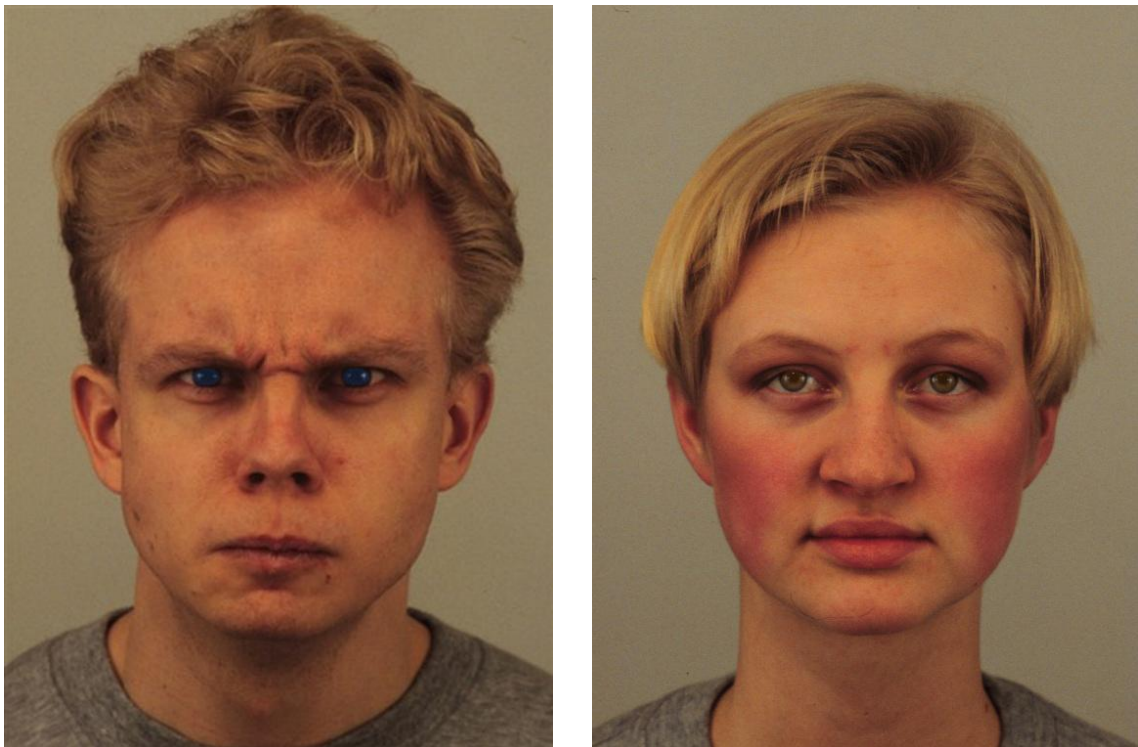
ACCEPTED MANUSCRIPT

## APPENDIX 1



Example of the faces used in Study 1: male and female faces were presented in either blue or green and displayed either an angry or a happy expression. Participants had to decide as quickly as possible whether the faces were male or female (the gender naming condition) or blue or green (the color naming condition).

## APPENDIX 2



Example of the faces used in Study 2: male and female faces were presented having either blue or brown eyes and displaying either an angry or a happy expression. Participants had to decide as quickly as possible whether the faces were male or female (the gender naming condition) or had blue or brown eyes (the color naming condition).



## APPENDIX 3

**Wat eye-color can tell us**

The **iris** is the diafragm of the human eye. The iris determines eye-color as we perceive it. Blue or brown eyes are actually blue or brown irises. In eyes of all colors, the iris contains the black pigment, eumelanin. Color variations among different irises are typically attributed to the eumelanin content within the iris. The density of cells within the stroma affects how much light is absorbed by the underlying pigment epithelium. An eye with little eumelanin is perceived as blue, whereas an eye with a lot of eumelanin is perceived as brown. The OCA2 gene on chromosome 15, explains most human eye-color variation.

Interestingly, the OCA2 gene not only relates to eye-color, but to certain personality traits as well. The OCA2 codes for the amount of eumelanin, which also determines to some extent whether someone has a more extraverted or introverted personality

Introverts display a stronger blood circulation in brain areas involved in internal processes such as planning and problem solving. Extraverted people, on the other hand, display a stronger blood circulation in areas involved in sensory processing. Introverted people are sensitive, observing, and considerate. Extraverts are energetic, enthusiastic, talkative, and assertive.

The OCA2 gene variant for **blue** eyes is commonly seen in **Extraverted** people. The gene variant for **brown** eyes is usually observed in **Introverted** people. As such, people with **blue** eyes are more likely to be **extraverted**, and people with **brown** eyes are more likely to be **introverted**.

Translation from Dutch of the bogus popular science text used to introduce the personality categorization task of Study 3.