



UvA-DARE (Digital Academic Repository)

Mimicry of ingroup and outgroup emotional expressions

Sachisthal, M.S.M.; Sauter, D.A.; Fischer, A.H.

DOI

[10.1080/23743603.2017.1298355](https://doi.org/10.1080/23743603.2017.1298355)

Publication date

2016

Document Version

Final published version

Published in

Comprehensive Results in Social Psychology

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Sachisthal, M. S. M., Sauter, D. A., & Fischer, A. H. (2016). Mimicry of ingroup and outgroup emotional expressions. *Comprehensive Results in Social Psychology*, 1(1-3), 86-105. <https://doi.org/10.1080/23743603.2017.1298355>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

ARTICLE



Mimicry of ingroup and outgroup emotional expressions

Maijen S. M. Sachisthal, Disa A. Sauter and Agneta H. Fischer

Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands

ABSTRACT

Research into emotional mimicry has found that the extent to which we mimic others depends on group membership and the emotion at hand. Particularly, negative emotions are mimicked more when expressed by the perceiver's ingroup. It is, however, still debated what process underlies emotional mimicry and whether previous findings of enhanced mimicry of negative emotions expressed by ingroup members are robust. We therefore first aimed to replicate Study 2 of van der Schalk, Fischer et al. (2011), specifically testing the finding of differences in emotional mimicry for models from different ethnic groups. Moreover, we extended the study by (1) including nonverbal emotional vocalizations and (2) including all negative emotions that were previously studied in a group mimicry context, that is, anger, fear, and sadness, in addition to happiness. We test two alternative explanations of emotional mimicry: whether emotional mimicry is a matched-motor response or whether emotional mimicry is influenced by meaning and context as proposed by the Emotion Mimicry in Context view. The results do not replicate the findings of van der Schalk, Fischer et al. (2011). For the facial and vocalization stimuli, we did not find emotional mimicry effects for anger, fear, or sadness, neither did we find effects of group membership. We only found emotional mimicry effects for happiness (action units 6 and 12) in the facial study. We discuss various explanations for the lack of findings, with the within-subjects design as most likely explanation.

ARTICLE HISTORY

Received 19 December 2016
Accepted 23 December 2016

KEYWORDS

Emotional mimicry;
Matched-Motor Hypothesis;
emotional vocalizations;
emotion contagion

Introduction

When someone close to you describes what a great experience they had last weekend, you might catch yourself smiling along with them, while you would similarly share their expression of sadness if they told you about something dreadful that had happened. These are examples of emotional mimicry, which is the tendency to imitate emotions signaled by nonverbal behaviors, such as facial expressions (i.e. facial mimicry), postures, or vocalizations (screams, laughter, or similar) of an interaction partner (Hess, Philippot, & Blairy, 1999). Mimicry is time-locked, typically occurring within a second (Dimberg & Thunberg, 1998), and the emotional expression of the mimicker depends on that of the expresser (Hess & Fischer, 2014). Mimicry is assumed to be an automatic process, occurring independently of the interaction partner, or the relationship between expresser (mimickee) and perceiver (mimicker). One explanation for this mimicry process, which has been applied to a range

CONTACT Agneta H. Fischer a.h.fischer@uva.nl

© 2017 University of Amsterdam. Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

of social behaviors, has been referred to as Behavior Matching (Chartrand & Bargh, 1999) or the Matched-Motor Hypothesis (Hess & Fischer, 2013). This account argues that the activation of motor areas in the brain is caused by what observers perceive. Applying this view to the mimicry of emotions, the context and kind of emotion expressed should not influence mimicry, and so all emotions should be mimicked to the same extent across all contexts.

However, previous studies have found differences in the extent to which we mimic others' emotions depending on attitudes, liking, or group membership (see Hess & Fischer, 2013 for an overview). In particular, facial displays of negative emotions have been shown to be mimicked more when expressed by the perceiver's ingroup compared to outgroup expressions (Bourgeois & Hess, 2008; van der Schalk et al., 2011a), suggesting that emotional mimicry not only depends on what we objectively see but also on the impact and meaning of what we see. In other words, a frown is not just a frown, but may signal hostile intentions. This has been proposed in the Emotion Mimicry in Context view (Hess & Fischer, 2013), which argues that emotional mimicry is different from behavioral mimicry, because there is no one-to-one relation between specific nonverbal displays and emotions. Thus, the inference and – by implication – the mimicry of emotions depends on the relationship, context, and meaning of the nonverbal signal. In this view, displays of emotions are considered intrinsically meaningful because they provide information about the expresser's dispositions and intentions (Hess, Blair, & Kleck, 2000; Knutson, 1996). The core assumption is that we only mimic emotional displays that are affiliative, that is, implying connection or approach, and thus require no or minimal social costs.

We aimed to replicate one of the studies that have provided support for this view, because the results are crucial for the explanation of emotional mimicry. Across two studies, van der Schalk, Fischer, et al. (2011) found that facial anger and fear displays were mimicked to a greater extent when displayed by an ingroup (Caucasian) member than an outgroup (non-Caucasian) member, demonstrating enhanced ingroup mimicry. In the current study, we replicated Study 2 of van der Schalk, Fischer, et al. (2011), because only small effect sizes were found and because this provides a rigorous test of the two alternative explanations of emotional mimicry, that is, whether emotional mimicry is a matched-motor response or if it is influenced by meaning and context. We aim to further test these explanations by extending the study by (1) including nonverbal emotional vocalizations (e.g. laughter) and (2) including all negative emotions that were previously studied in a group context, that is, anger, fear, and sadness, in addition to happiness. More mimicry of ingroup emotional displays compared to outgroup displays and the occurrence of "mimicry" in response to nonverbal emotional vocalizations would be interpreted as evidence for the Emotion Mimicry in Context view. If no such differences will be found, this would lend support to the Matched-Motor Hypothesis, because context such as group membership and the different emotions should not influence mimicry behavior according to the Matched-Motor Hypothesis.

Facial displays versus vocalizations

Some research supporting the contextual view of emotional mimicry has examined emotional mimicry across communicative channels. Nonverbal emotional vocalizations, that is, human vocal sounds that do not involve words, can be interpreted as

expressions of emotions, such as screams, laughter, or sighs (Laukka et al., 2013). More specifically, such emotional vocalizations have been shown to lead to emotion-congruent facial expressions (Hawk, Fischer, & van Kleef, 2012), suggesting emotional mimicry across channels. Participants were either asked to reproduce emotional vocalizations or to merely listen to them. Even when merely listening to nonverbal emotional vocalizations, individuals facially expressed the corresponding emotion. Hearing someone laugh made individuals smile, while hearing someone cry led participants to produce sad facial displays. In a similar vein, facial, face-voice combinations and bodily expressions of emotions have been shown to all result in similar, emotion-congruent facial muscle activity (Magnée, Stekelenburg, Kemner, & de Gelder, 2007). These findings indicate that individuals react to an emotional display by showing the same emotional display across channels, rather than merely imitating it, as could be argued to be the case for facial emotional mimicry. Further evidence for cross-modal mimicry comes from neuroscientific research; a functional Magnetic Resonance Imaging (fMRI) study by Warren and colleagues (2006), for example, showed that during the perception of affective nonverbal vocalizations, the same premotor regions were active that are used in the production of facial emotional displays. Moreover, suppressing sensorimotor cortices through the use of TMS has been found to disrupt participants' ability to discriminate between nonverbal affect vocalizations (Banissy et al., 2010). Summarizing, these findings suggest that some of the same brain areas are involved in processing emotional cues independent of the modality of the expressed emotion, which implies that observers may "mimic" the meaning of the signal, rather than its objective features.

Specific emotions

One core assumption of the Emotion Mimicry in Context view is that emotional displays are only mimicked to the extent that they are seen as affiliative, because the mimicry of non-affiliative or hostile intentions would not serve the general function of mimicry to smoothen social interactions (e.g. Hatfield, Cacioppo, & Rapson, 1994). Happiness displays are generally seen as signaling affiliative intentions (Hess et al., 2000) and are thus almost always mimicked, independently of group membership (Bourgeois & Hess, 2008; van der Schalk, Fischer, et al., 2011). Smiling simply comes at very low social costs. The affiliative signal of negative emotions may, however, vary with context. Previous studies have found that sadness displays are only mimicked if the mimicker feels a high level of intimacy or similarity with the expresser, as shown by Bourgeois and Hess (2008). In their study, sadness was only mimicked when the participant and the model considered themselves most similar, which, in this case, meant that they shared both parts of identity that were manipulated (ethnic group membership and being a basketball player or a non-basketball player). When displaying sadness, expressers signal a need for support. Mimicking sadness displays could thus be socially costly, especially in the case of strangers, as mimicry would signal the willingness to support the expresser (Bavelas, Black, Lemery, & Mullett, 1986). In the same vein, fear can be seen as affiliative because it signals submissiveness, which, when displayed by an outgroup member can induce dominant behavior (Tiedens & Fragale, 2003). Indeed, fear has been shown to be mimicked more by ingroup members (van der Schalk, Fischer, et al., 2011), presumably because it is less costly to show a submissive signal to ingroup than outgroup members.

Anger on the other hand, is a non-affiliative emotional signal, and has also been shown to be mimicked to a greater extent when displayed by an ingroup member (van der Schalk, Fischer, et al., 2011). This may be explained by the fact that the anger (usually operationalized by merely a frown) is perceived as not directed at the observer, but as a way of sharing something with the observer. This would explain why individuals mimic ingroup anger more than outgroup anger. In short, including happiness and three negative emotions, which have previously been studied, will allow us to further investigate if and to what extent emotional meaning influences mimicry behavior.

Hypotheses

In the current study, we tested the following hypotheses. First, mimicry of facial displays of happiness will occur for both ingroup and outgroup members, whereas we expect that the facial displays of the three negative emotions will be mimicked to a greater extent when displayed by ingroup members, which would lend support to the Emotion Mimicry in Context view. Second, facial emotional displays as a response to the non-verbal emotional vocalizations will be interpreted as evidence for the Emotion Mimicry in Context view, whereas the absence of such differences will be seen as support for the Matched-Motor Hypothesis. Emotional displays as a response to nonverbal emotional vocalizations would indicate that participants respond to the meaning of the emotion, therefore supporting the Emotion Mimicry in Context view, while the absence of such a cross-modality mimicry response would indicate that facial mimicry might indeed be due to a matched-motor response.

Finally, as was done in Study 2 of van der Schalk, Fischer, et al. (2011), we included measures of self-reported emotions after each stimulus block to measure emotional contagion. The results of van der Schalk and colleagues suggested that emotional contagion did not parallel emotional mimicry, as emotional contagion was neither affected by emotion meaning nor by group membership. van der Schalk and colleagues argued that this finding might suggest that facial behavior is a more automatic response than self-reports, which may leave more room for deliberate interpretation. Importantly, these findings contradict the core assumptions of emotional contagion theory, namely that mimicry mediates subjective emotional convergence through afferent feedback (Hatfield et al., 1994).

Method

Apparatus and materials

Stimuli

Facial emotional displays. Video clips of happiness, anger, fear, and sadness displayed by Northern European (ingroup) and Mediterranean (outgroup) models were used as stimulus material. All clips were taken from the well-validated Amsterdam Dynamic Facial Expression Set (ADFES; van der Schalk, Hawk, Fischer, & Doosje, 2011). Videos of 16 different models were used. Participants saw the videos of eight members of their ingroup and eight members of their outgroup, expressing all four emotional displays. All clips were approximately 4 s long, starting with a neutral expression, and reaching the

apex of the emotional display after approximately 1 s. The stimulus set was balanced for ethnicity, gender, and emotion, resulting in 64 stimuli. Participants saw all videos, which were presented in blocks based on the specific emotion and group membership, so that all anger displays of ingroup members were presented in one block, for instance. The presentation of emotion by group membership blocks was randomized per participant, resulting in eight blocks, each consisting of eight stimuli.

Nonverbal emotional vocalizations. Nonverbal vocalizations of the same four emotions (anger, fear, happiness, and sadness) of either eight Dutch or eight Namibian individuals were presented to each participant (Dutch: Sauter, Crasborn, & Haun, 2010; Namibian: Sauter, Eisner, Ekman, & Scott, 2010). We used the nonverbal emotional vocalizations of Namibian individuals, as there is currently no validated stimulus material of Moroccan producers available. Furthermore, Sauter (2013) found that listeners were not able to accurately infer the ethnic origin of nonverbal emotional vocalizations, and so the Namibian stimuli could reasonably be used as a different ethnic outgroup. All participants heard each stimulus once. The entire stimulus set was normalized for peak amplitude and digitized at 41 kHz. The stimulus set was balanced for ethnicity, gender, and emotion, resulting in a total of 64 stimuli. The stimuli were presented in blocks of emotion by group membership, resulting in eight auditory blocks, which were presented in a random order.

Measures

Identification. An adapted version of the overlap of self ingroup and outgroup (OSIO) questionnaire was included (Aron, Aron, & Smollan, 1992; Schubert & Otten, 2002). As in the study of van der Schalk, Fischer, et al. (2011), participants indicated how much overlap they perceived between themselves and the groups “native Dutch” (OSIO-ingroup) and “immigrants” (OSIO-outgroup) on a scale from 1 (“no overlap”) to 7 (“complete inclusion”). We used this questionnaire to exclude participants who saw too much overlap and thus for whom the distinction between ingroup and outgroup would not apply (see section on “Exclusion criteria”).

FaceReader. To determine the expressed emotions of the participants, we used FaceReader 6.0 (Noldus, 2014), an automated facial coding software package. After finding the participant’s face in the video, FaceReader creates an active appearance model of the face (Cootes & Taylor, 2004), which is then used to compute scores for the intensity of facial action units (AUs) (i.e. facial muscle movements) on a scale from 0 to 1 (Lewinski, Den Uyl, & Butler, 2014). FaceReader reports these variables at each time point of the video, and these data can then be averaged to determine the participant’s reaction to a stimulus within a given time window. As FaceReader uses a model-based approach, its performance is largely unaffected by factors such as lighting, facial orientation, and background variation (Van Kuilenburg, Wiering, & den Uyl, 2005). This makes FaceReader a good tool for the current study, as no calibration is needed, thus reducing the risk that participants notice being filmed.

The software has been validated using the ADFES stimulus set with an 88% basic emotion accuracy score, which was highest for happiness (96%) and lowest for anger (76%) and had an accuracy of 91% for Dutch faces and 86% for faces of Turkish/

Moroccan origin (Lewinski et al., 2014). FaceReader also allows for the classification of AUs separately, and AU1 (inner brow raise), AU2 (outer brow raise), AU4 (brow lowerer), AU5 (upper lid raise), AU6 (cheek raise), AU9 (nose wrinkle), AU15 (lip corner depressor), and AU25 (lips part) have been found to be the best classifiers, whereas reasonable classifiers were AU12 (lip corner puller), AU14 (dimpler), AU17 (chin raiser), and AU26 (jaw drop).

The specific AUs were recorded and used in our analysis. AUs are facial AUs, that is, facial muscle movements, with some AUs or the combination of AUs being characteristic for an emotion (e.g. Duclos et al., 1989; Friesen & Ekman, 1983). AU4, for example, describes the lowering and drawing together of the brows, which requires the activity of the corrugator supercilii, is typically present in anger displays, whereas AU12, describing the pulling up of the lip corners, which involves the zygomaticus major, is a typical AU used in happiness displays (Hess & Blairy, 2001). For the current study, we used the following AUs as they are characteristic of the emotions included here: AU4 (anger), AU5 (fear), both AU6 and AU12 (happiness), and AU15 (sadness). If participants express the same facial action as the model, this would be interpreted as mimicry (in the case of the facial stimuli) and a convergent emotional display (in the case of vocal stimuli).

Self-reported emotions. Following van der Schalk, Fischer, et al. (2011), we used an adapted version of the discrete emotion scale (DES; Izard, Dougherty, Bloxom, & Kotsch, 1974) to measure self-reported emotions. We used the same items as van der Schalk and colleagues, but additionally included the three sadness items of the DES. Specific emotions are measured with different items per emotion category (Dutch terms in brackets); happiness is measured with the items “amused (geamuseerd),” “happy (blij),” “cheerful (vrolijk);” anger items are “angry (boos),” “irritated (geïrriteerd),” and “mad (kwaad);” fear is measured with the items “anxious (bezord),” “fearful (angstig),” and “nervous (zenuwachtig);” sadness with “sad (verdrietig),” “worried (ongerust),” and “down (teneergeslagen);” and finally “aversion (afkeer)” and “contempt (minachting).” Participants indicated to which extent they felt each of these emotions on a scale from 1 (“not at all”) to 5 (“very intense”). In the van der Schalk, Fischer, et al. study (2011; Study 2), reliabilities of the emotion categories ranged from $\alpha = .87$ to $\alpha = .93$.

Interpersonal Reactivity Index (IRI). To create a break between the main study and the extension study, we used the IRI questionnaire, which measures cognitive and emotional components of empathy and was developed by Davis (1980). The IRI is made up of four subscales, two of which measure cognitive components of empathy, namely perspective taking (PT) and fantasy, while the other two components measure emotional components of empathy, namely empathic concern (EC) and personal distress (PD). Previous research has shown that emotional mimicry can be influenced by emotional empathy, that is, people who are emotionally empathetic mimic others’ emotions to a greater extent (e.g. Sonnby-Borgström, 2002), and that emotional mimicry can induce PT (Stel & Vonk, 2009). We thus decided to include the items of these three subscales (EC, PD, PT) in the current study. We thus used 21 items of the scale, which are answered on a five-point Likert scale ranging from 1 (“Does not describe me well”) to 5 (“Describes me well”). We conducted an exploratory analysis to examine the relation between the IRI and mimicry.

Procedure

After signing the informed consent, participants were seated in front of a computer, where the experiment was administered individually. The study was presented as a study on the recognition of emotional expressions of native Dutch models compared with immigrant models, so that the ethnicity of the models would become salient. Each participant was presented with the video clips (and emotional vocalizations), which were presented consecutively with an intertrial period between stimuli of 1 s within each stimulus block. All participants were first presented with the facial emotion part and then with the nonverbal emotional vocalization of the same emotion. This was to ensure that the replication part of the current study on facial expressions was not influenced by possible previous mimicry to the nonverbal emotional vocalizations.

After each stimulus block, participants were asked to indicate the intensity of the emotions they felt at that moment. Between the main and extension study, they were asked to fill in a questionnaire, the IRI. Next, participants heard the vocalization of each emotion in eight stimulus blocks. Just as in the first part of the study, participants were asked to indicate the intensity of their presently felt emotions on a scale from 1 to 5 after each stimulus block.

Having completed both parts of the study, participants were asked to fill in some additional questionnaires, including the OSIO and questions regarding their demographics. During the stimulus presentations, participants' faces were filmed by the computer's webcam, without them being aware of it. The recordings were used for the FaceReader software to determine the participants' facial emotions as a response to the models' emotional expressions. The facial expressions of the participants were coded per stimulus block lasting from first stimulus onset until last stimulus offset within each block. Lastly, participants were debriefed and asked for consent for the filmed material to be used for further analysis.

Exit interview

In order to make sure that participants were unaware of the goal of the study, we asked them about their assumptions regarding the goals and the nature of the study. More specifically, we excluded participants if they explicitly referred to "mimicry," or "my facial expressions," because the awareness that the focus is on their own faces may contribute to the regulation of their facial expression. Further, participants were asked to indicate where they and their parents were born to make sure that they were indeed Dutch.

Sample and design

Sixty students (see section on "Power analysis") were recruited at the University of Amsterdam to participate in the study for either one participant point or €10 compensation. They were all presented with the same emotional stimuli (Within Subjects (WS): emotion: anger, fear, happiness, sadness) expressed by both ingroup and outgroup members (WS: group membership: ingroup, outgroup).

Exclusion criteria

Participants' data were excluded on the following grounds. First, only Dutch students who identified with the Dutch identity could participate. This was operationalized as (1) both their parents having been born in the Netherlands, and (2) their scores on the OSIO toward the Dutch being higher than 3.5. Second, participants who showed awareness of the aim of the study were excluded. Participants who explicitly referred to "mimicry," their own "facial expressions," or "being filmed" were assumed to be aware of the true nature of the study and were therefore excluded.

Power analysis

A power analysis for a within-subjects MANOVA was conducted using the G*power software (Faul, Erdfelder, Lang, & Buchner, 2007). The analysis suggested that a sample size of 53 is needed for a power of 0.95, when using the following values: $\alpha = 0.05$, $f = 0.21$ (based on the overall interaction of Study 2 of van der Schalk, Fischer, et al., 2011), and a correlation among repeated measures of 0.2, which is the lowest average correlation we expect to find among measures. We planned to analyze the data using a mixed-model approach, which has not been used in previous studies; therefore, we employed the within-subjects MANOVA design as an approximate measure to determine how many participants we needed to recruit for sufficient power.

As we planned to exclude participants (see section on "Exclusion criteria") on a number of different criteria, we aimed to test 60 people to ensure a sufficiently large final sample. Given the possibility of technical issues regarding the use of FaceReader (e.g. videos of the participants not being readable by FaceReader) as an additional factor for losing data, we believed that a sample size of 60 was justifiable.

Results

Participants

Eighty-one students¹ of the University of Amsterdam were recruited (60 female, 22 male; mean, $M_{Age} = 21.78$; standard deviation, $SD = 2.41$), who participated for €10 or participant credit. The data of 15 participants could not be used due to our exclusion criteria or because of technical reasons. We excluded five participants because they suspected that they were being filmed or guessed the aim of the study correctly; we excluded two participants as their parents were not Dutch; the videos of four participants could not be used due to technical errors; one participant did not give approval to use their videos for analysis and one participant had a lower score on the ingroup score of the OSIO than 3.5. The data of 67 participants were used for analysis.

Preregistered analysis

Social category

As expected, participants reported a larger amount of overlap between themselves and the group of Dutch individuals ($M = 5.72$, $SD = .97$) than with the group of immigrants

($M = 2.76$, $SD = 1.21$), $t(66) = 16.00$, $p < .001$. As mentioned earlier, participants' ingroup score showed that one individual reported a smaller overlap than 3.5 with the ingroup and was therefore excluded from analysis.

Self-reported emotions

Facial study

To test whether the Emotion conditions elicited convergent emotions, we conducted 4 (Emotion) by 2 (Group membership) repeated measures ANOVAs per emotion measure. We used Bonferroni adjusted alpha levels of .0125 per test (.05/4). Table 1 reports all means (M) and SDs.

For the anger measure ("boos"), the main effect of Emotion was significant, $F(1.41, 93.29) = 505.75$, $p < .001$, $\eta^2_p = .94$, as was the main effect of Group membership, $F(1, 66) = 10.28$, $p = .002$, $\eta^2_p = .135$, and the interaction, $F(1.51, 99.46) = 4.51$, $p = .001$, $\eta^2_p = .12$. To disentangle the significant interaction term, we ran planned contrasts, which revealed that in the anger condition, participants reported significantly more anger ($M = 4.14$, $SD = .69$) than in the other three conditions ($M = 1.07$, $SD = .12$), $F(1, 66) = 1297.27$, $p < .001$, $\eta^2_p = .91$. Moreover, participants reported being significantly more angry after seeing angry ingroup members ($M = 4.39$, $SD = .78$) than after seeing angry outgroup members ($M = 3.90$, $SD = 1.29$), $F(1, 66) = 1297.27$, $p < .001$, $\eta^2_p = .91$.

For the fear measure ("angst"), the main effect of Emotion was significant, $F(2.07, 136.44) = 423.40$, $p < .001$, $\eta^2_p = .903$. The main effect of Group membership was not significant, $F(1, 66) = .85$, $p = .36$, $\eta^2_p = .013$, nor was the interaction effect, $F(2.37, 156.29) = 2.68$, $p = .03$, $\eta^2_p = .013$.

For the happiness measure ("blij"), the main effect of Emotion was significant, $F(1, 66) = 755.00$, $p < .001$, $\eta^2_p = .972$. Self-reported happiness levels were higher in the happiness condition ($M = 4.39$, $SD = .58$) than in the anger ($M = 1.00$, $SD = .00$), fear

Table 1. Means (M) and standard deviations (SDs) of self-reported emotion per emotion by group membership in facial study.

Emotion	Emotion Condition							
	Anger		Fear		Happiness		Sadness	
	M	SD	M	SD	M	SD	M	SD
Anger								
Ingroup	4.39	0.78	1.09	0.29	1.01	0.12	1.12	0.37
Outgroup	3.90	1.02	1.04	0.21	1.00	0.00	1.16	0.41
Total	4.14	0.94	1.07	5.56	1.01	0.09	1.14	0.39
Fear								
Ingroup	1.16	0.41	4.58	0.72	1.01	0.12	1.93	0.97
Outgroup	1.30	0.65	4.39	0.72	1.00	0.00	1.85	0.91
Total	1.23	0.55	4.49	0.72	1.01	0.08	1.89	0.94
Happy								
Ingroup	1.00	0.00	1.00	0.00	4.36	0.73	1.00	0.00
Outgroup	1.00	0.00	1.00	0.00	4.42	0.68	1.00	0.00
Total	1.00	0.00	1.00	0.00	4.39	0.70	1.00	0.00
Sad								
Ingroup	1.17	0.41	1.27	0.54	1.01	0.12	4.40	0.82
Outgroup	1.19	0.50	1.28	0.57	1.01	0.12	4.12	0.81
Total	1.18	0.46	1.28	0.55	1.01	0.12	4.26	0.82

Mean values in bold represent convergent emotion responses. Results are based on a sample of $N = 67$.

($M = 1.00$, $SD = .00$), and sadness conditions ($M = 1.00$, $SD = .00$). Neither the main effect of Group membership, $F(1, 66) = .38$, $p = .54$, $\eta^2_p = .001$, nor the interaction effect, $F(1, 66) = .13$, $p = .54$, $\eta^2_p = .006$, was significant.

For the sadness measure (“verdrietig”), the main effect of Emotion was significant, $F(1.96, 129.35) = 606.15$, $p < .001$, $\eta^2_p = .86$. Sadness was reported to be higher in the sadness condition ($M = 4.26$, $SD = .63$) than in the anger ($M = 1.18$, $SD = .31$), fear ($M = 1.28$, $SD = .46$), and happiness condition ($M = 1.01$, $SD = .09$). The interaction effect was not significant, $F(2.01, 132.96) = 2.01$, $p = .05$, $\eta^2_p = .043$, nor was the main effect of Group membership, $F(1, 66) = .38$, $p = .54$, $\eta^2_p = .034$.

Vocalization study

To test whether the emotion vocalizations elicited convergent emotions, we conducted 4 (Emotion) by 2 (Group membership) repeated measures ANOVAs per emotion measure. Again, Bonferroni adjusted alpha levels of .0125 per test were used (.05/4). Table 2 reports all M and SD s.

For the anger measure (“boos”), the main effect of Emotion was significant, $F(1.69, 111.60) = 229.90$, $p < .001$, $\eta^2_p = .86$, as was the main effect of Group membership, $F(1, 66) = 132.09$, $p < .001$, $\eta^2_p = .67$, and the interaction, $F(1.95, 128.52) = 144.57$, $p < .001$, $\eta^2_p = .77$. To disentangle the significant interaction term, we ran planned contrasts, which revealed that in the anger condition, participants reported significantly more anger ($M = 2.76$, $SD = .59$) than in the other three conditions ($M = 1.07$, $SD = .14$), $F(1, 66) = 556.95$, $p < .001$, $\eta^2_p = .83$. Moreover, participants reported being significantly more angry after hearing angry ingroup members ($M = 3.87$, $SD = .89$) than angry outgroup members ($M = 1.66$, $SD = .64$), $F(1, 66) = 321.69$, $p < .001$, $\eta^2_p = .80$.

For the fear measure (“angst”), we found a significant main effect of Emotion, $F(1.94, 128.01) = 219.48$, $p < .001$, $\eta^2_p = .837$. The main effect of Group membership was not significant, $F(1, 66) = 4.90$, $p = .03$, $\eta^2_p = .069$. The interaction effect was significant, F

Table 2. Means (M) and standard deviations (SD s) of AU-activity per emotion by group membership in vocalization study.

Emotion	Emotion Condition							
	Anger		Fear		Happiness		Sadness	
	M	SD	M	SD	M	SD	M	SD
Anger								
Ingroup	3.87	0.89	1.09	0.42	1.00	0.00	1.01	0.12
Outgroup	1.66	0.64	1.27	0.62	1.01	0.12	1.06	0.24
Total	2.76	1.35	1.18	0.53	1.01	0.09	1.04	0.19
Fear								
Ingroup	1.06	0.24	4.51	0.68	1.00	0.00	1.57	0.78
Outgroup	1.16	0.48	3.10	1.34	1.00	0.00	2.33	1.19
Total	1.11	0.38	3.81	1.27	1.00	0.00	1.95	1.07
Happy								
Ingroup	1.00	0.00	1.00	0.00	4.07	0.94	1.01	1.22
Outgroup	1.06	0.30	1.22	0.71	4.16	0.91	1.01	1.22
Total	1.03	0.21	1.11	0.52	4.12	0.93	1.01	1.22
Sad								
Ingroup	1.03	0.17	1.25	0.59	1.00	0.00	4.46	0.77
Outgroup	1.07	0.32	1.10	0.35	1.03	0.17	3.58	1.26
Total	1.05	0.25	1.18	0.49	1.01	0.12	4.02	1.13

Mean values in bold represent convergent emotion responses. Results are based on a sample of $N = 67$.

(2.18, 143.68) = 51.34, $p < .001$, $\eta^2_p = .517$. To disentangle the significant interaction term, we ran planned contrasts, which revealed that participants reported significantly more fear in the fear condition ($M = 3.81$, $SD = .83$) than in the other three conditions ($M = 1.35$, $SD = .32$), $F(1, 66) = 551.15$, $p < .001$, $\eta^2_p = .536$. Moreover, participants reported being significantly more fearful after hearing fearful ingroup members ($M = 4.51$, $SD = .68$) than after hearing fearful outgroup members ($M = 3.10$, $SD = 1.34$), $F(1, 66) = 76.27$, $p < .001$, $\eta^2_p = .79$.

For the happiness measure (“blij”), the main effect of Emotion was significant, $F(1.36, 90.08) = 384.22$, $p < .001$, $\eta^2_p = .928$. Planned contrasts showed that happiness was reported to be higher in the happiness condition ($M = 4.12$, $SD = .79$) than in the anger ($M = 1.03$, $SD = .15$), fear ($M = 1.11$, $SD = .36$), and sadness condition ($M = 1.01$, $SD = .09$). The main effect of Group membership was not significant, $F(1, 66) = 4.90$, $p = .03$, $\eta^2_p = .069$, nor was the interaction effect, $F(2.07, 136.32) = 1.12$, $p = .20$, $\eta^2_p = .024$.

For the sadness measure (“verdrietig”), the main effect of Emotion was significant, $F(1.38, 91.02) = 386.24$, $p < .001$, $\eta^2_p = .927$, as was the main effect of Group membership, $F(1, 66) = 32.31$, $p < .001$, $\eta^2_p = .329$, and the interaction, $F(1.51, 99.38) = 9.69$, $p < .001$, $\eta^2_p = .226$. To disentangle the significant interaction term, we ran planned contrasts, which revealed that participants reported significantly more sadness in the sad condition ($M = 4.02$, $SD = .80$) than in the other three conditions ($M = 1.08$, $SD = .15$), $F(1, 66) = 999.18$, $p < .001$, $\eta^2_p = .87$. Moreover, participants in the sadness condition reported being significantly more sad after hearing sad ingroup members ($M = 4.46$, $SD = .77$) than sad outgroup members ($M = 3.58$, $SD = 1.26$), $F(1, 66) = 28.80$, $p < .001$, $\eta^2_p = .304$.

Facial mimicry

It is important to note that FaceReader only analyses faces when they are frontal, when they are completely “in frame,” and when they are not obscured (e.g. by manipulating the hands in front of the face). The output from FaceReader consists of continuous intensity ratings of different AUs for each frame (25 ms). The facial displays of the models lasted 3–4 s, and we measured the onset of an expression in the participant within the whole time frame of stimulus duration. All outputs were intensity scores times frequency of occurrence. We defined mimicry index as FaceReader’s output for the specific AUs per Emotion condition. Higher scores imply a higher degree of emotional mimicry. Following van der Schalk and colleagues (2011a), we examined AU4 for anger, AU5 for fear, AU6 and AU12 for happiness, and in addition, AU15 for sadness. Only the blocks of which 70% of the frames could be analyzed by FaceReader were used for analyses. This implies that there are complete data for 58 of the participants in the facial mimicry study and for 54 participants in the vocalization study. With respect to the current videos, we also noticed that some AUs (e.g. AU7, AU20) appeared to be problematic and incorrectly recognized by Face Reader (FR), due to light flashes of the camera, or participants’ face touching. Fortunately, this was not the case for the core AUs that were tested here.

We analyzed the effects of Emotion and Group membership on each overall emotion AU score separately, as our hypotheses differed per emotion. We could not use a multilevel approach in our analyses, because the data were too skewed due to the

large number of zeros (no occurrence of an AU in a given condition). Since we did not aim to test differences between models, but rather to test overall mimicry across all models displaying a specific emotion, this did not pose a problem in our view. Following the procedure by van der Schalk and colleagues (2011a), we used AU4 for anger, AU5 for fear, both AU6 and AU12 for happiness, and AU15 for sadness. Per AU score, we first conducted a 4 (WS: Emotion: anger, fear, happiness, sadness) by 2 (WS: Group membership: ingroup vs. outgroup) repeated measures ANOVA, and then used planned contrasts to examine the effects of Emotion and Group membership on mimicry of each emotion separately. When necessary, Greenhouse Geiser correction was applied. Bonferroni adjusted alpha levels of .01 per test (.05/5) were used. The first planned contrasts compared the relevant Emotion condition with the other Emotion conditions. For example, for the fear AU score, we first contrasted the fear condition to the anger, sadness, and happiness conditions. The second contrast compared the ingroup and outgroup conditions on the fear AU score. M and SDs can be found in Table 3.

Facial study

Anger. The 4 by 2 repeated measures ANOVA on AU4 score yielded no significant main effect of Emotion, $F(1.88, 107.18) = 2.14, p = .04, \eta^2_p = .056$. We also found no significant main effect of Group membership, $F(1, 57) = 1.66, p = .20, \eta^2_p = .028$, nor an interaction effect, $F(2.03, 115.52) = 1.07, p = .21, \eta^2_p = .027$.

Fear. The 4 by 2 mixed ANOVA on the AU5 score showed no significant effects of Emotion, $F(2.62, 149.46) = .86, p = .39, \eta^2_p = .017$, Group membership, $F(1, 57) = .11, p = .74, \eta^2_p = .002$, or the interaction, $F(3, 171) = .32, p = .81, \eta^2_p = .006$.

Table 3. Means (*M*) and standard deviations (SDs) of AU-activity per emotion by group membership in facial study.

Facial behavior	Emotion Condition							
	Anger		Fear		Happiness		Sadness	
	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD
AU4								
Ingroup	5.26	9.34	4.53	8.56	3.43	7.41	3.91	6.65
Outgroup	6.96	11.43	4.02	8.56	3.36	6.91	4.70	6.03
Total	6.11	10.43	4.28	8.53	3.39	7.13	4.31	6.33
AU5								
Ingroup	0.04	0.18	0.09	0.30	0.05	0.20	0.08	0.29
Outgroup	0.08	0.34	0.12	0.41	0.06	0.16	0.05	0.23
Total	0.06	0.27	0.11	0.35	0.05	0.18	0.07	0.26
AU6								
Ingroup	3.70	7.06	1.76	3.18	4.09	6.82	2.15	5.25
Outgroup	3.81	6.78	3.61	8.10	5.60	10.44	3.83	6.69
Total	3.75	6.89	2.69	6.20	4.84	8.81	2.99	6.04
AU12								
Ingroup	7.73	13.62	6.53	11.32	10.62	14.30	5.09	9.67
Outgroup	5.24	7.48	6.26	11.47	14.15	20.75	7.88	15.04
Total	6.48	11.01	6.39	11.35	12.39	17.83	6.49	12.67
AU15								
Ingroup	3.09	5.19	1.99	4.15	2.53	6.60	2.93	5.54
Outgroup	1.60	3.76	2.52	4.74	2.04	6.78	2.05	4.19
Total	2.35	4.57	2.25	4.44	2.29	6.67	4.49	4.91

Note. Mean values in bold represent convergent emotion responses. Results are based on a sample of $N = 58$.

Happiness. The 4 by 2 mixed ANOVA on the AU6 scores showed a significant main effect of Emotion, $F(2.23, 127.33) = 3.70, p = .006, \eta^2_p = .08$. The planned contrast comparison showed that AU6 was expressed more frequently in the happiness condition than in the anger, fear, or sadness condition, $F(1, 57) = 7.22, p = .01, \eta^2_p = .02$. No significant main effect of Group membership was found, $F(1, 57) = 4.92, p = .03, \eta^2_p = .08$. The interaction effect was not significant either, $F(2.45, 139.43) = .76, p = .41, \eta^2_p = .016$.

The 4 by 2 mixed ANOVA conducted on AU12 scores showed a significant main effect of Emotion, $F(2.28, 130.21) = 7.64, p < .001, \eta^2_p = .15$. The planned contrast showed that AU12 was expressed more frequently in the happiness condition than in the other three conditions ($M = 6.45, SD = 8.45$), $F(1, 57) = 18.27, p < .001, \eta^2_p = .05$. The main effect of Group membership was not significant, $F(1, 57) = 1.15, p = .29, \eta^2_p = .001$, nor was the interaction effect, $F(3, 171) = 2.52, p = .07, \eta^2_p = .04$.

Sadness. The 4 by 2 mixed ANOVA conducted on the AU15 scores showed no significant effects for Emotion, $F(2.50, 142.50) = .06, p = .96, \eta^2_p = .001$, Group membership, $F(1, 57) = 1.69, p = .20, \eta^2_p = .029$, or the interaction, $F(2.44, 139.26) = .68, p = .45, \eta^2_p = .015$.

Vocalization study

For the analysis of FaceReader output in response to the vocalizations, we followed exactly the same procedure as for the facial stimuli. Mimicry was operationalized in the same way as in the facial study (AU4 for anger, AU5 for fear, AU6 and AU12 for happiness, and AU15 for sadness). Table 4 reports all M and SDs.

Table 4. Means (*M*) and standard deviations (SDs) of AU-activity per emotion by group membership in vocal study.

Facial behavior	Emotion Condition							
	Anger		Fear		Happiness		Sadness	
	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD
AU4								
Ingroup	3.20	4.18	3.74	5.82	4.52	7.36	3.65	4.97
Outgroup	3.57	4.72	3.85	5.34	4.37	6.62	3.89	4.65
Total	3.38	4.44	3.79	5.56	4.44	6.96	3.77	4.79
AU5								
Ingroup	0.03	0.14	0.02	0.09	0.15	0.65	0.06	0.31
Outgroup	0.79	4.50	0.11	0.60	0.20	0.63	0.05	0.18
Total	0.41	3.19	0.07	0.43	0.17	0.64	0.05	0.25
AU6								
Ingroup	2.69	5.16	2.36	3.65	1.91	3.14	1.60	3.25
Outgroup	1.92	4.25	1.68	3.85	1.47	2.89	1.10	1.84
Total	2.30	4.72	2.02	3.75	1.69	3.02	1.35	2.64
AU12								
Ingroup	10.45	16.55	9.78	11.56	8.92	8.98	8.55	10.63
Outgroup	7.00	12.36	8.77	12.02	9.00	11.36	9.01	11.03
Total	8.73	14.64	9.28	11.75	8.96	10.18	8.78	10.78
AU15								
Ingroup	1.15	3.29	1.38	3.29	1.88	4.00	2.31	4.98
Outgroup	2.25	4.36	3.09	6.24	2.09	4.21	2.94	6.69
Total	1.70	3.88	2.24	5.04	1.99	4.08	2.63	5.88

Note. Mean values in bold represent convergent emotion responses. Results are based on a sample of $N = 54$.

Anger. We conducted a 4 by 2 repeated measures ANOVA on AU4 and found no significant effects for Emotion, $F(2.34, 123.99) = 1.13, p = .24, \eta^2_p = .027$, Group membership, $F(1, 53) = .24, p = .62, \eta^2_p = .005$, or the interaction, $F(2.67, 141.73) = .10, p = .94, \eta^2_p = .002$.

Fear. The 4 by 2 mixed ANOVA on AU5 yielded no significant effects for Emotion, $F(1.09, 57.89) = .40, p = .31, \eta^2_p = .02$, Group membership, $F(1, 53) = 2.04, p = .16, \eta^2_p = .034$, or the interaction, $F(1.11, 58.77) = .48, p = .26, \eta^2_p = .024$.

Happiness. The 4 by 2 mixed ANOVA on AU6 scores showed no significant main effect of Emotion, $F(2.01, 106.21) = 1.31, p = .15, \eta^2_p = .036$, but we found a significant main effect of Group membership, $F(1, 53) = 7.82, p = .007, \eta^2_p = .129$, with ingroup-vocalizations eliciting more AU6 responses ($M = 2.14, SD = 3.88$) than outgroup-vocalizations ($M = 1.54, SD = 3.33$). The interaction was not significant, $F(3, 159) = .10, p = .96, \eta^2_p = .002$.

The 4 by 2 mixed ANOVA conducted on AU12 scores showed no significant effects of Emotion, $F(2.48, 131.22) = .05, p = .97, \eta^2_p = .001$, Group membership, $F(1, 53) = .06, p = .98, \eta^2_p = .055$, or the interaction, $F(3, 159) = .93, p = .43, \eta^2_p = .015$.

Sadness. The 4 by 2 mixed ANOVA on the AU15 scores showed no main effect of Emotion, $F(3, 159) = 1.33, p = .27, \eta^2_p = .025$, no significant effect for Group membership, $F(1, 53) = 3.07, p = .09, \eta^2_p = .055$, nor a significant interaction, $F(2.64, 140.16) = .70, p = .49, \eta^2_p = .015$.

Interpersonal Reactivity Index

We also examined whether the IRI was correlated with the Facial mimicry scores of the different AUs in both the facial and vocal study. The correlations and p -values can be found in Table 5. We only found a significant difference between the mimicry of AU15 within the sadness condition and the overall IRI score ($r = .287, p = .03$) and the Emotional Contagion subscale ($r = .268, p = .04$). All other correlations were nonsignificant.

Table 5. Pearson correlations of mimicry scores and IRI scores (facial study).

		Pearson correlations					
		Overall mimicry	AU4 mimicry	AU5 mimicry	AU6 mimicry	AU12 mimicry	AU15 mimicry
IRI	Pearson's r	−0.067	0.017	−0.022	−0.089	−0.132	0.287
	p -Value	0.623	0.899	0.870	0.511	0.329	0.030
EC	Pearson's r	−0.207	−0.038	−0.061	−0.049	−0.300	0.268
	p -Value	0.123	0.779	0.654	0.715	0.023	0.043
PT	Pearson's r	0.123	0.083	0.110	−0.068	0.097	0.189
	p -Value	0.364	0.539	0.413	0.618	0.471	0.160
PD	Pearson's r	−0.020	0.007	−0.071	−0.075	−0.035	0.154
	p -Value	0.883	0.956	0.600	0.580	0.795	0.253

The mean frequency of AU scores per participant in the corresponding emotion (e.g. AU4 in the anger Emotion condition) was used as the mimicry score per AU.

Exploratory analyses

FACS data

In the original study (Study 2; van der Schalk, Fischer, et al., 2011), data were manually coded with the Facial Action Coding System (FACS). To compare the performance of FaceReader with manually FACS-coded result, we decided to code a subset of the sample (the facial study of 39 participants) using FACS. We coded the AUs that were included in the original study, that is, AU4 (anger), AU5 (fear), AU6 and AU12 (both for happiness). We then ran the same analyses as was done in our analyses of the FaceReader data of the facial study, using Bonferroni adjusted alpha levels of .0125 per test (.05/4).

Anger

The 4 by 2 repeated measures ANOVA on the AU4 score yielded a significant main effect of Emotion, $F(2.27, 86.28) = 11.25, p < .001, \eta^2_p = .281$. AU4 was expressed more frequently in the anger condition ($M = .25, SD = .29$) than in the fear ($M = .09, SD = .20$), happiness ($M = .03, SD = .08$), and sadness condition ($M = .16, SD = .07$). We did not find a significant main effect of Group membership, $F(1, 38) = 2.17, p = .15, \eta^2_p = .054$, or an interaction effect, $F(2.02, 76.94) = .26, p = .67, \eta^2_p = .010$.

Fear

The 4 by 2 mixed ANOVA on the AU5 score showed a significant effect of Emotion condition, $F(1.10, 41.77) = 3.58, p = .003, \eta^2_p = .204$. AU5 was expressed more frequently in the fear condition ($M = .07, SD = .14$), than in the anger ($M = .01, SD = .03$), happiness ($M = .00, SD = .00$), and sadness condition ($M = .00, SD = .00$). Neither the main effect of Group membership, $F(1, 38) = 1.13, p = .29, \eta^2_p = .029$, nor the interaction effect, $F(1.09, 41.40) = .26, p = .27, \eta^2_p = .032$, was significant.

Happiness

The 4 by 2 mixed ANOVA on the AU6 scores showed no significant effects; neither of Emotion, $F(1.43, 54.43) = .64, p = .27, \eta^2_p = .034$; nor of Group membership, $F(1, 38) = .55, p = .46, \eta^2_p = .014$; or the interaction effect, $F(1.43, 54.21) = .10, p = .64, \eta^2_p = .005$.

The 4 by 2 mixed ANOVA conducted on AU12 scores showed a significant main effect of Emotion, $F(1.22, 46.17) = 11.17, p < .001, \eta^2_p = .421$. AU12 was expressed more in the happiness condition ($M = .29, SD = .27$) than in the anger ($M = .04, SD = .11$), fear ($M = .04, SD = .11$), and sadness condition ($M = .02, SD = .07$). The main effect of Group membership was not significant, $F(1, 38) = .10, p = .75, \eta^2_p = .003$, nor was the interaction effect, $F(1.91, 72.62) = 1.01, p = .21, \eta^2_p = .04$.

In short, in contrast with most findings reported by FaceReader, the FACS coding results show that the emotional mimicry effect was not only present for happiness but also for anger and fear, replicating the mimicry results of van der Schalk, Fischer, et al. (2011) and other studies (see Hess & Fischer, 2013). The effect of group membership was not replicated, however, similar to the results of the FaceReader data.

Discussion

This study aimed to replicate the effect of group membership on the mimicry of negative emotions (van der Schalk, Fischer, et al., 2011, Study 2), and to test the Contextual Model of Emotional Mimicry by extending these findings to facial reactions to emotional vocalizations. We will first discuss the findings with respect to the facial stimuli, as this was the main aim of this replication study.

We found hardly any facial mimicry effects of emotion on the basis of the FaceReader data: participants only smiled more (cheek raising as well as lip corner puller) when watching smiling faces. However, they did not frown more when watching angry faces, raise their eyelids more when watching fearful faces, or show more lip corner depressor movements when watching sad faces. This pattern of results does not replicate the original findings by van der Schalk, Fischer, et al. (2011), who found that participants mimicked anger, fear, and happiness (sadness was not included in that study). In addition, we did not find any effect of group membership, whereas van der Schalk et al. found that facial mimicry was more pronounced for ingroup than outgroup faces, in the case of anger and very marginally in the case of fear ($p = .09$). It should be mentioned that the effect for anger in van der Schalk, Fischer, et al. (2011) was also quite small ($t(147) = 2.29, p = .01$).

Our results with emotional self-reports do replicate part of the original findings: similar to van der Schalk, Fischer, et al. (2011), we found emotional contagion effects, such that when watching angry, fearful, happy, or sad faces, participants reported more of the feelings they watched than any of the other feelings (e.g. participants reported feeling more fear when they had seen a fearful face than when they had seen an angry, sad, or happy face). In line with van der Schalk, Fischer, et al. (2011), we also did not find an effect of group membership on emotional contagion, with the exception of the response to an angry face, where we found that participants frowned more to ingroup than outgroup faces.

With respect to the vocalization study, we found a similar pattern of results. There were no convergent facial responses upon hearing emotional vocalizations of anger, fear, or sadness, even not for happiness. In addition, we did not find any interactions with group membership of the individual who produced the vocalizations. With respect to emotion self-reports, like van der Schalk, Fischer, et al. (2011), we found emotional contagion effects for all emotions. Thus, participants reported to feel more angry, happy, afraid, and sad, after hearing matching sounds. Interestingly, however, we also found effects of group for angry, fearful, and sad sounds, in line with the hypothesis. In other words, participants reported to feel more anger in response to ingroup anger vocalizations, fear in response to ingroup fearful vocalizations, and sadness in response to sad ingroup vocalizations in contrast with similar sounds of the outgroup.

In conclusion, the most important results of van der Schalk, Fischer, et al. (2011) were not replicated, that is, no support was found for more emotional mimicry of ingroup compared to outgroup faces. This means that the Contextual Model of Emotional Mimicry is not supported. Our results do not yield support for the Matched-Motor Hypothesis either, however, because in order for this model to be supported, we would have had to find facial mimicry of emotions in the facial study.

There are several explanations for this lack of replication. A number of differences in methodology and design may have contributed to differences in results. First of all, although we used the same task, cover story, and stimuli, and a similar sample of students, there were some differences with van der Schalk, Fischer, et al (2011). Specifically, van der Schalk and colleagues used a complete between-subjects design, whereas we used a complete within-subjects design. This means that our participants were presented with all stimuli, and they thus had the possibility to compare faces from ingroup and outgroup. Given that the study was presented as a study on the recognition of emotional expressions of native Dutch versus immigrant models, we would expect that the ethnicity of the models would become more salient if the participants would indeed see the different faces. Thus, if anything a within-subjects design may have enhanced rather than diminished differences in mimicry of ingroup and outgroup faces. We therefore do not think that this can explain our lack of group membership effects.

Another difference is that van der Schalk and colleagues FACS coded participants' faces, whereas we used automated coding software, FaceReader. In order to rule out this possibility as an explanation for the failure to replicate van der Schalk's results, we therefore coded a subset of the data using FACS. Exploratory analyses revealed stronger effects of emotional mimicry for anger (AU4) and fear (AU5) in the facial mimicry resulting from FACS coding. This could suggest that FaceReader was not sensitive enough to capture all facial actions that were visible on the face, especially the frown (AU4) and the upper eyelid raiser (AU5).

Another explanation may be the involvement of the participants. Facial mimicry is an automatic process, but it does require engagement in the task. In order to emotionally mimic, participants have to take the task seriously, attend to the stimuli, and they should be motivated to try to detect emotions in the faces they watch. It could be that participants in the current study were less motivated than in the previous studies. Because we used a complete within-subjects design, the task was much longer than in the original study, which could have reduced engagement in the task and made participants more bored and less attentive. This is supported by the fact that whereas we found a mimicry effect of laughter in the facial study, we did not find this effect anymore in the vocalization study. This complete absence of mimicry in response to vocalizations also contradicts previous findings from our lab (Hawk et al., 2012). Because the vocalization study was always presented after the facial study, this may suggest that participants were even more bored or unmotivated by the end of the study, although they did report emotions in both the facial and vocalization study that were in line with the hypothesis and suggest emotional contagion.

In sum, we do not think that the lack of replication can be explained by differences in tasks, stimuli, or sample. Because ingroup effects of facial mimicry have been found in other studies as well (e.g. Bourgeois & Hess, 2008), we suggest that the most likely explanation is the difference in design, and the consequences for participants' attention and engagement. In addition, we should also note that the effect of group membership reported by van der Schalk, Fischer, et al. (2011) was a small effect, and thus may have been further diminished by this difference in design. More generally, such an explanation would be in line with the general assumption of the Contextual Model of Emotional Mimicry (Hess & Fischer, 2013, 2014), which posits that it is less likely that people mimic emotions if they do not know and are not interested in why the other person is expressing an emotion, which is the case when they merely watch a stranger on a

video. Although this was also the case for the original study of van der Schalk, Fischer, et al. (2011), the alleged lesser engagement of the participants in the present study may have added to this effect. New research in which participants' motivation to decode others' emotions is better operationalized, would shed new light on the social context effect of emotional mimicry and contagion.

Note

1. Based on our G*power analysis, which indicated that we needed 53 participants (see section on "Power analysis"), we initially recruited and tested 60 participants. Because we could only use the data of 40 individuals in the first sample for the facial mimicry study and the data of 37 individuals for the vocalization study, we decided to run the study a second time, testing 22 additional participants.

Acknowledgments

The authors wish to thank Delano de Klerck for his help in collecting the data and Xia Fang for FACS coding the videos of the first batch of participants. Moreover, the authors wish to thank the European Association of Social Psychology for awarding this study with the preregistration grant.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors wish to thank the European Association of Social Psychology for awarding this study with the preregistration grant.

References

- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self-scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596–612. doi:10.1037/0022-3514.63.4.596
- Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, *77*, 343–352. doi:10.1016/j.biopsycho.2007.11.008
- Banissy, M. J., Sauter, D. A., Ward, J., Warren, J. E., Walsh, V., & Scott, S. K. (2010). Suppressing sensorimotor activity modulates the discrimination of auditory emotions but not speaker identity. *The Journal of Neuroscience*, *30*(41), 13552–13557. doi:10.1523/JNEUROSCI.0786-10.2010
- Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I show how you feel": Motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, *50*, 322–329. doi:10.1037/0022-3514.50.2.322
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910. doi:10.1037/0022-3514.76.6.893
- Cootes, T. F., & Taylor, C. J. (2004). *Statistical models of appearance for computer vision. Imaging science and biomedical engineering*. Manchester, UK: University of Manchester.

- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology*, 10, 85.
- Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, 39, 39–45. doi:10.1111/1467-9450.00054
- Duclos, S. E., Laird, J. D., Schneider, E., Sexter, M., Stern, L., & Van Lighten, O. (1989). Emotion-specific effects of facial expressions and postures on emotional experience. *Journal of Personality and Social Psychology*, 57, 100–108. doi:10.1037/0022-3514.57.1.100
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. doi:10.3758/BF03193146
- Friesen, W. V., & Ekman, P. (1983). *EMFACS-7: Emotional facial action coding system*. San Francisco, CA: University of California, San Francisco.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York, NY: Cambridge University Press.
- Hawk, S. T., Fischer, A. H., & van Kleef, G. A. (2012). Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions. *Journal of Personality and Social Psychology*, 102(4), 796–814. doi:10.1037/a0026234
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, 40, 129–141. doi:10.1016/S0167-8760(00)00161-6
- Hess, U., Blairy, S., & Kleck, R. E. (2000). The influence of expression intensity, gender, and ethnicity on judgments of dominance and affiliation. *Journal of Nonverbal Behavior*, 24, 265–283. doi:10.1023/A:1006623213355
- Hess, U., Philippot, P., & Blairy, S. (1999). Mimicry: Facts and fiction. In P. Philippot & R. S. Feldman (Eds.), *The social context of nonverbal behavior. Studies in emotion and social interaction* (pp. 213–241). Cambridge, UK: Cambridge University Press.
- Hess, U., & Fischer, A. (2013). Emotional mimicry as social regulation. *Personality and Social Psychology Review*, 17, 142–157. doi:10.1177/1088868312472607
- Hess, U., & Fischer, A. (2014). Emotional mimicry: Why and when we mimic emotions. *Social and Personality Psychology Compass*, 8, 45–57. doi:10.1111/spc3.12083
- Izard, C. E., Dougherty, F. E., Bloxom, B. M., & Kotsch, N. E. (1974). *The differential emotion scale: A method of measuring the meaning of subjective experience of discrete emotions*. Nashville, TN: Vanderbilt University, Department of Psychology.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103, 54–69. doi:10.1037/a0028347
- Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, 20, 165–182. doi:10.1007/BF02281954
- Laukka, P., Elfenbein, H. A., Söder, N., Nordström, H., Althoff, J., Chui, W., ... Thingujam, N. S. (2013). Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations. *Frontiers in Psychology*, 4, 353. doi:10.3389/fpsyg.2013.00353
- Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology and Economics*, 7(4), 227–236. doi:10.1037/npe0000028
- Magnée, M. J. C. M., Stekelenburg, J. J., Kemner, C., & de Gelder, B. (2007). Similar facial EMG responses to faces, voices, and body expressions. *Neuroreport*, 18(4), 369–372. doi:10.1097/WNR.0b013e32801776e6
- Noldus. (2014). *FaceReader: Tool for automatic analysis of facial expression: Version 6.0* [Software]. Wageningen, The Netherlands: Noldus Information Technology B. V.
- Sauter, D. (2013). The role of motivation and cultural dialects in the in-group advantage for emotional vocalizations. *Frontiers in Psychology*, 4, 814. doi:10.3389/fpsyg.2013.00814
- Sauter, D., Crasborn, O., & Haun, D. B. M. (2010). The role of perceptual learning in emotional vocalizations. *Journal of Acoustical Society of America*, 128, 2476. doi:10.1121/1.3508875

- Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408–2412. doi:[10.1073/pnas.0908239106](https://doi.org/10.1073/pnas.0908239106)
- Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity*, *1*, 353–376. doi:[10.1080/152988602760328012](https://doi.org/10.1080/152988602760328012)
- Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, *43*, 433–443. doi:[10.1111/1467-9450.00312](https://doi.org/10.1111/1467-9450.00312)
- Stel, M., & Vonk, R. (2009). Empathizing via mimicry depends on whether emotional expressions are seen as real. *European Psychologist*, *14*(4), 342–350. doi:[10.1027/1016-9040.14.4.342](https://doi.org/10.1027/1016-9040.14.4.342)
- Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of personality and social psychology*, *84*(3), 558.
- van der Schalk, J., Fischer, A., Doosje, B., Wigboldus, D., Hawk, S., Rotteveel, M., & Hess, U. (2011a). Convergent and divergent responses to emotional displays of ingroup and outgroup. *Emotion*, *11*, 286–298. doi:[10.1037/a0022582](https://doi.org/10.1037/a0022582)
- van der Schalk, J., Hawk, S. T., Fischer, A., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, *11*, 907–920. doi:[10.1037/a0023853](https://doi.org/10.1037/a0023853)
- Van Kuilenburg, H., Wiering, M., & den Uyl, M. (2005). A model based method for facial expression recognition. In J. Gama, R. Camacho, P. Brazdil, A. Jorge, & L. Torgo (Eds.), *Lectures notes in computer science: Vol. 3720. Machine learning: ECML 2005* (pp. 194–205). Berlin, Germany: Springer-Verlag. doi:[10.1007/11564096_22](https://doi.org/10.1007/11564096_22)
- Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., ... Scott, S. K. (2006). Positive emotions preferentially engage an auditory-motor “mirror” system. *The Journal of Neuroscience*, *26*, 13067–13075. doi:[10.1523/JNEUROSCI.3907-06.2006](https://doi.org/10.1523/JNEUROSCI.3907-06.2006)