Emotional expressions as social signals of rejection and acceptance: evidence from the Affect Misattribution Paradigm

Heerdink, M.W.; van Kleef, G.A.; Homan, A.C.; Fischer, A.H.

DOI
10.1016/j.jesp.2014.09.004

Publication date
2015

Document Version
Final published version

Published in
Journal of Experimental Social Psychology

Link to publication

Citation for published version (APA):

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.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Emotional expressions as social signals of rejection and acceptance: Evidence from the Affect Misattribution Paradigm

Marc W. Heerdink *, Gerben A. van Kleef, Astrid C. Homan, Agneta H. Fischer

University of Amsterdam, The Netherlands

HIGHLIGHTS

• Happy expressions signal more acceptance than other emotional expressions.
• Angry expressions signal more rejection than other negative emotional expressions.
• The predicted associations are reliable at presentation times of only 50 ms.
• Effects are consistent across three conceptualizations of acceptance and rejection.
• These findings help explain consequences of expressed emotions in dyads and groups.

ABSTRACT

Inclusion in social groups is vital to human survival and wellbeing. We propose that emotional expressions signal acceptance versus rejection to observers. Based on this idea, we hypothesized that happy facial expressions prime acceptance, whereas angry expressions prime rejection. In six experiments using the Affect Misattribution Paradigm (Payne, Cheng, Govorun, & Stewart, 2005), we tested to what extent observers associate facial expressions (angry, happy, sad, fearful, and neutral) with three different operationalizations of acceptance and rejection (accept/reject, warm/cold, close/distant). A meta-analysis on these experiments revealed that angry expressions were more strongly associated with rejection than other (negative) expressions, and that happy expressions were more strongly associated with acceptance than other facial expressions. Effects were stable and robust at presentation times of 50 ms and higher and were similar across conceptualizations of acceptance/rejection. We discuss implications for theorizing on the social functions of emotions and the processing of emotional expressions.

Given the evolutionary significance of group life (Cosmides & Tooby, 1992; Dunbar, 1992), it would be adaptive for human beings to be sensitive to moment-to-moment variations in the extent to which fellow group members accept them (Baumeister & Leary, 1995). But how do people gauge their level of acceptance? Adopting a social–functional approach to emotion (Fischer & Manstead, 2008; Frijda & Mesquita, 1994; Keltner & Haidt, 1999; Van Kleef, 2009), we propose that individuals use the emotional expressions of others as implicit signals of acceptance versus rejection. Emotional expressions inform observers about a person's specific evaluation of a situation, and they communicate social motives and behavioral intentions (Fridlund, 1994; Hess & Fischer, 2013; Knutson, 1996). Thus, different emotional expressions – even two different negative emotional expressions – may have different implications for an observer's relation to the group. Here, we aim to show that different emotional expressions signal different degrees of acceptance and rejection. We develop and test the hypotheses that (i) angry facial expressions are more strongly associated with rejection than other (negative) facial expressions, and that (ii) happy facial expressions are more strongly associated with acceptance than other facial expressions. Furthermore, we test whether these associations may be found even when facial expressions are presented for very short durations.

The social–functional approach to emotions posits that emotional expressions play a vital role in regulating social life (Keltner & Haidt, 1999). These social functions are typically investigated by studying the consequences of emotional expressions for (social) behavior within a particular context. Research into such social consequences indicates that the effects of any given emotional expression may differ considerably depending on the social context, the individual's resources, and which type of consequences are investigated. This can be illustrated by the case of anger: Some research has documented destructive consequences of anger expressions, such as lowered relationship satisfaction.
and increased conflict in romantic relationships (Sanford & Rowatt, 2004), an increased likelihood of divorce (Gottman & Levenson, 2002), and retaliation and impasses in conflict resolution (Friedman et al., 2004; Kopelman, Rossete, & Thompson, 2006; Van Kleeft & Côté, 2007). However, other research has documented favorable outcomes of anger expressions, such as greater concessions from counterparts in negotiations (Van Kleeft, De Dreu, & Manstead, 2004), increased effort and task performance of subordinates (Sy, Côté, & Saavedra, 2005; Van Kleeft, Homan, Beersma, & Van Knippenberg, 2010), increased conformity of deviant group members (Heerdink, Van Kleeft, Homan, & Fischer, 2013), enhanced learning performance of students (Van Doorn, Van Kleeft, & Van der Pligt, 2014), and long-term improvement of intimate relationships (Fischer & Roseman, 2007).

Findings regarding the social consequences of other emotions are similarly mixed. For instance, some studies indicate that expressions of happiness increase affiliative and cooperative tendencies among observers, especially in communal relationships (Clark, Pataki, & Carver, 1996). Other studies, however, suggest that expressions of happiness evoke exploitation, especially in competitive settings (for a review of this literature, see Van Kleeft, De Dreu, & Manstead, 2010). Furthermore, sad and fearful expressions have been found to increase affiliation and helping (Barnett, Howard, Melton, & Dino, 1982; Clark, Oullette, Powell, & Milberg, 1987; Clark & Taraban, 1991; Yee & Greenberg, 1998), especially in communal relationships (Clark et al., 1996). However, other research has shown that people tend to avoid interactions with (chronically) sad individuals when possible, because such interactions tend to be draining and not socially rewarding (e.g., Coyne, 1976).

To better understand how the same emotional expression may have both positive and negative social consequences, we believe that it is useful to draw a distinction between short-term, immediate signals conveyed by emotional expressions, and the longer-term consequences of expressing an emotion. In this view, emotional expressions convey elemental social signals to an observer that are relatively stable across situations (Fridlund, 1994). How these social signals affect an observer (i.e., their consequences) does not only depend on the social signal itself, but also on contextual factors that determine the relevance of a particular social signal to one’s current goals. Focusing on the social signals conveyed by emotional expressions thus also helps us gain an understanding of the type of contextual factors that may be relevant in determining the consequences of a particular emotional expression. Our focus here is on social signals of acceptance versus rejection, as this constitutes a key dimension of social life.

### Emotional expressions as signals of acceptance and rejection

Acceptance and rejection may be seen as the extremes of a bipolar dimension that represents one’s level of acceptance. To be accepted is a fundamental human need (Baumeister & Leary, 1995), and accordingly experiences that negatively affect one’s level of acceptance have a great impact on people. For instance, research has shown that feeling rejected is a highly aversive experience (Williams, 2007), with a neural activation pattern similar to physical pain (Eisenberger & Lieberman, 2004). Rejection can be a strong motivator of both antisocial behavior (Wesselmann, Butler, Williams, & Pickett, 2010; Williams, 2007) and behavior aimed at regaining acceptance such as ingratiation (Romero-Canyas et al., 2010) and conformity (Heerdink et al., 2013). At the other end of the dimension, acceptance can be conceptualized as a state of increased (social) safety that facilitates development and self-expression (Heerdink et al., 2013; Ryan & Deci, 2000).

Our hypotheses focus on two prevalent emotion displays that we believe to be highly consequential for an observer’s position in the group: anger and happiness. The expression of happiness is typically associated with affiliative social motives (Fischer & Manstead, 2008; Fridlund, 1994). People with an intention to affiliate smile more (Clark et al., 1996; Kraut & Johnston, 1979), and those who smile are also perceived as having affiliative intentions (Hess, Blairy, & Kleck, 2000; Knutson, 1996). In terms of consequences, happiness has been theorized to improve social bonds (Fischer & Manstead, 2008). We therefore predict that happy expressions are interpreted as signals of acceptance.

Anger, on the other hand, is an emotion that is often linked to antisocial behavior and aggression (Averill, 1982). People express anger when they intend to change another person’s behavior (Fischer & Roseman, 2007), thereby signaling that certain behavior is unacceptable. The expression of anger in close relationships is predictive of short-term divorce (Gottman & Levenson, 2002), and it is related to both decreased relationship satisfaction and increased conflict (Sanford & Rowatt, 2004). Yet, Fischer and Roseman (2007) found that anger can also be effective in eliciting behavioral change (see also Heerdink et al., 2013). These social consequences of angry expressions align with the effects of social rejection described above. We therefore propose that angry expressions may be seen as signaling a (temporary) problem in the relationship between people, and we predict that angry expressions are interpreted as signals of rejection.

Like happy expressions, sad and fearful expressions have also been found in some studies to increase affiliation, particularly in communal relationships (Clark et al., 1996). However, these social consequences are typically attributed to these emotional expressions signaling a need for help (Clark et al., 1996), rather than signaling acceptance. Thus, although the social consequences of sadness and fear may partially overlap with those of happiness in some cases, we argue that these expressions convey different social signals (i.e., a need for help vs. acceptance, respectively). We therefore expect happy and angry expressions to be stronger signals of acceptance and rejection than fearful and sad expressions.

If expressions of happiness and anger are indeed robust signals of acceptance versus rejection, these associations may be expected to generalize to other conceptualizations of the acceptance/rejection dimension. Williams and Bargh (2008) suggest that acceptance and rejection are grounded on the experiences of warmth and coldness, respectively (see also Zhong & Leonardelli, 2008). Warmth is the first dimension on which people judge others (Fiske, Cuddy, & Glick, 2007), which is consistent with the possibility that signs of acceptance versus rejection can be quickly gleaned from others’ nonverbal behavior. Similarly, it has been argued that social distance shares a conceptual basis with other kinds of distances (e.g., spatial; temporal; Trope & Liberman, 2010; see also IJzerman & Semin, 2009). It follows that acceptance and rejection are linked to closeness and distance as well.

Based on these theoretical considerations, we formulated two hypotheses: Happy facial expressions are associated with acceptance to a greater extent than other facial expressions (Hypothesis 1); and angry facial expressions are associated with rejection to a greater extent than other negative emotional facial expressions (Hypothesis 2). We further examined to what extent these associations generalize across various conceptualizations of acceptance versus rejection (i.e., accept/reject, warm/cold, or close/distant). If the predicted effects are robust, we should find that the associations emerge irrespective of the particular conceptualization of acceptance versus rejection.

### The Affect Misattribution Paradigm

To test whether facial expressions are indeed associated with rejection and acceptance, we conducted a series of six experiments using the Affect Misattribution Paradigm (AMP; Payne, Cheng, Govorun, & Stewart, 2005). The AMP measures implicit associations by assessing the extent to which primes in a task, in which the primes activate conceptual knowledge that is subsequently misattributed to the ideograms (e.g., Blaison, Imhoff, Hühnel, Hess, & Banse, 2012; Gawronski & Ye, 2014).
The advantage of using the AMP is that it yields reliable results that are relatively unaffected by metacognitive strategies (e.g., hypothesis guessing; Payne et al., 2005). Because emotional expressions are rich and complex stimuli that may trigger a number of different processes in observers (e.g., Hareli & Hess, 2010; Van Kleef, 2009), this feature helped us exclude the influence of explicit strategies and processes that might interfere with the participants’ primary associations when perceiving facial expressions in real-world situations. Using the AMP thus allowed us to assess the associations between facial expressions and acceptance/rejection in a relatively unbiased way.

In our experiments, we manipulated the facial expressions and the conceptualization of the acceptance/rejection dimension. In addition, we exploratorily varied the presentation time of the facial expressions (i.e., Prime Duration) to see which presentation durations produce reliable effects. We tested the four shortest presentation times allowed by our equipment (ranging from 17 ms to 67 ms). The initial experiments were intended to explore the influence of various combinations of Response Dimension and Prime Duration. As our insights progressed, our experiments increasingly focused on confirming these explorations and testing our specific hypotheses.

The results of the experiments are analyzed and reported as a meta-analysis. This strategy was chosen because the procedures of the six experiments were largely identical, and writing them up as separate studies would entail considerable repetition. Furthermore, the meta-analytic report as a whole provides a more reliable, more complete, and more general picture than each of the studies in isolation, and the greater quantity of data included in the meta-analysis provides the statistical power needed to obtain reliable estimates of the effect sizes.

Method

The six experiments that were conducted to test our hypotheses shared a mixed Response Dimension × Prime Duration × Facial Expression design, in which Response Dimension and Prime Duration were manipulated as between-subjects factors and Facial Expression was manipulated as a within-subjects factor. The exact levels of Response Dimension, Prime Duration, and the Facial Expressions that were included in each of the six experiments are displayed in Table 1.

Procedure

Participants were seated individually in cubicles. The set-up of the cubicles ensured that the top of the computer screen was approximately aligned with the participant’s eyes. Participants received either monetary compensation or course credit.

Affect Misattribution Paradigm (AMP)

The task used in the experiments was procedurally similar to the AMP as originally introduced by Payne et al. (2005). In a number of trials, the participant’s task was to ‘intuitively judge’ the meaning of Chinese ideograms, and to press one of two keys to indicate which of two alternatives best represented the meaning of the ideogram. These Chinese ideograms were preceded by primes of different facial expressions. To test our hypotheses, we examined how the facial expression primes influenced participants’ responses. Depending on the Response Dimension condition, the response alternatives were acceptance/rejection, warm/cold, or close/distant. Using these conceptual response dimensions made our task more similar to semantic variants of the task (e.g., Blaison et al., 2012) than to the original, affective AMP as introduced by Payne et al. (2005). Besides tapping into acceptance versus rejection, these word pairs vary on a valence dimension. Thus, depending on the pattern of findings, effects could potentially be interpreted in terms of valence. In order to allow us to judge our findings in relation to a valence explanation, we collected valence ratings for the three word pairs in a separate study (described in Appendix A). The results indicated that the word pairs differ considerably in the extent to which they represent positive versus negative valence (accept/reject is more strongly valenced than are warm/cold and close/distant). We return to this issue in the Discussion.

The instructions emphasized that there was no objectively ‘right’ answer in the task, and made no mention of response speed. Thus, a speed/accuracy trade-off explanation is unlikely to account for the results described below (see also the null findings with regard to response latencies, described under Auxiliary analyses below).

Participants were presented with 20 trials in each facial expression condition (80–100 trials in total, depending on the facial expressions included in the study; see Table 1). Trials were presented in two blocks, each containing half of the trials. There was a short break between the blocks.

Each trial began with a facial expression prime, presented for 17, 33, 50, or 67 ms (depending on study and condition). These Prime Durations represent 1 to 4 screen refresh cycles on the 60 Hz computer monitor that was used (1/60th to 4/60th of a second). The facial expressions were immediately followed by a Chinese ideogram that was randomly selected from a set of 100 unique ideograms. After 150 ms, this ideogram was replaced by a mask consisting of noise that remained on screen until a response was given. Participants then pressed a key on the keyboard (‘C’ or ‘M’) to indicate whether they thought this ideogram represented a word related to rejection (i.e., rejection, cold, or distant, depending on study and condition), or whether it represented a word related to acceptance (i.e., acceptance, warm, or close). The meaning of these two keys was counterbalanced between participants.

Primes

The images used as primes were taken from the Amsterdam Dynamic Facial Expression Set (ADFES; Van der Schalk, Hawk, Fischer, & Doosje, 2011). Ten different individuals were selected, five male (numbers 1, 2, 3, 5 and 6), and five female (1, 2, 3, 4, and 5), each posing five different facial expressions (neutral, happy, angry, sad, or fearful). With regard to the valence of the expressions, the ADFES validation study indicates that the happy facial expressions ($M = 5.88$ on a 7-point scale) were rated as more positive than the angry ($M = 2.45$), fearful, ($M = 2.37$), and sad ($M = 2.28$) facial expressions (Van der Schalk et al., 2011, Table 2). Fearful facial expressions did not differ in valence from either the angry or sad facial expressions, but angry facial expressions were rated as significantly more positive (i.e., less negative) than sad facial expressions. (No valence ratings were reported for the neutral expressions.) We desaturated the images (to make them black-and-white), slightly blurred and cropped them to remove distracting features (e.g., hair, ears), and resized them to $160 \times 240$ pixels. Example images are displayed in Fig. 1.

Results

Analytic strategy

A two-stage meta-analysis (Simmonds et al., 2005) was conducted to assess the overall effects of the facial expression primes and the impact of the Response Dimension and Prime Duration manipulations on the responses given by participants. In the first stage, effect sizes were estimated for each unique between-subjects condition within each study. This means that, for instance, two sets of effect sizes were obtained from Study 2, which had a 1 (Response Dimension: warm–cold) × 2...
(Prime Duration: 17 ms vs 50 ms) between-subjects design: warm-cold × 17 ms; and warm-cold × 50 ms. Slicing up the data in this fashion resulted in a total of 19 datasets, which were treated as separate experiments. This approach is recommended when the effect sizes from one study can be assumed to be independent (Borenstein, Hedges, Higgins, & Rothstein, 2010). This assumption was met because between-subjects manipulation of the moderators of interest ensured that each participant contributed to only one of the effect sizes obtained from a study. Other potential sources of effect size dependency, such as the location and the experimenter, were constant across all experiments, making it unlikely that any two conditions from the same experiment would be more related than two conditions from different experiments. In the second stage of the meta-analysis, the effect sizes obtained from the first stage were aggregated, and the overall impact of the moderators was assessed.

To obtain the effect size estimates in the first phase, mixed-effect logistic regression models were fitted to the individual conditions using the lme4 package (version 0.999999.2; Bates, Maechler, & Bolker, 2013) for R (version 3.0.1; R Core Team, 2013). A random intercept for participant was specified. The model estimated the following planned contrasts: 1) between the happiness condition (−1) and the four other conditions (each coded as 1/4); 2) between the anger (1) and fear/sad conditions (each coded as −1/2); 3) between the neutral (−1) and the three negative emotional conditions (each coded as 1/3); and 4) between the fear (−1) and sad (1) conditions. In those studies that omitted either the fear, or the sad condition (see Table 1), only the first three contrasts were fitted, and the contrast weights were adjusted (e.g., in contrast 1, all conditions except happiness received weights of 1/3 instead of 1/4). Note that contrasts 3 and 4 were specified in this way to obtain an orthogonal set of contrasts. Although no hypotheses were formulated about these contrasts, the fact that they are part of the same model (and therefore adjusted for while estimating contrasts 1 and 2) necessitates us to report them in order to provide a complete picture of the results.

In the second stage of the meta-analysis, a random-effects meta-analytic model was fit to the regression coefficients and standard deviations obtained from the first step using the metafor package (version 1.9.1; Viechtbauer, 2010) for R.² Using the random-effects model instead of a traditional fixed-effects model has the advantage that heterogeneity among effect sizes is allowed, thereby reducing the influence of any within-experiment dependency of effect sizes. The results of the meta-analysis are reported as odds ratios (ORs), followed by the associated 95% confidence intervals. Conventional p-values are also reported where applicable.

Note on figures

To ensure correct interpretation of the figures (forest plots), some explanation is in order. The figures are divided in two parts. The top part contains the results for each of the individual datasets for which an effect size was estimated. The label (left column, e.g., ‘accept/reject 17 ms (exp 1)’) indicates the Response Dimension (acceptance/rejection) and Prime Duration (17 ms) condition, and the experiment that this dataset originates from (Experiment 1). The right column presents the effect sizes (ORs) and their 95% confidence intervals numerically. In the center, these effect sizes and confidence intervals are graphically presented relative to a reference line set at an OR of 1 (which indicates no difference). Larger blocks denote that the dataset contained more participants, and therefore received more weight in the meta-analysis. Finally, the bottom part presents the meta-analytic summary. If there is no moderation, the overall ‘main effect’ is presented; if there is moderation, the summary is split out for each level of the moderator.

Participants and data cleaning

Table 2 summarizes the total number of participants in each experiment, and the results of the following checks that ensured data quality. First, participants were excluded from the sample if they could read Chinese, or if they used one of the two response keys in over 90% of the trials. Then, outliers were identified by first determining the distribution of log latencies in each condition within each experiment, and then removing responses with log latencies that deviated more than three standard deviations from the mean log latency in the condition. Finally, participants whose responses were outliers in more than 10% of the cases were dropped from the sample. Without dropping outliers, the analyses produce similar effect estimates as reported below, but with wider confidence intervals. Significant effects remain significant, with the exception of contrast 2, which becomes marginal.

Hypothesis 1. Happy facial expressions and acceptance

According to our first hypothesis, happy facial expressions should be more strongly associated with acceptance than other facial expressions. For this hypothesis to be supported, the first contrast should show an OR smaller than 1, which indicates that the likelihood of a rejection-related response is lower (i.e., the likelihood of the opposite acceptance-related response is greater) after a happy facial expression prime than after another facial expression prime. An initial random-effects model without any moderators showed the predicted effect, OR = 0.735 [0.648, 0.835], p < .001. Remaining heterogeneity (Qw [18] = 204.17, p < .001) suggested potential moderators, however, so we examined the influence of both Response Dimension and Prime Duration. The impact of Response Dimension on the contrast between happy and other facial expressions was found to be marginally significant,  qdimension(2) = 5.80, p = .055. Exploring the pattern of moderation indicated that the contrast between happy and other facial expressions was slightly more pronounced for the accept/reject Response Dimension (OR = 0.622 [0.516; 0.750]) than for the warm/cold (OR = 0.769 [0.635; 0.933]) or close/distant (OR = 0.887 [0.704; 1.118]) response dimensions. Because of the weak evidence for moderation, and because the simple effect was in the predicted direction for

---

² This approach was recommended by the author of the meta-analytic software package that was used; correspondence is available upon request.
each Response Dimension, we did not consider the influence of Response Dimension further.

For Prime Duration, on the other hand, we found much stronger evidence that it affected the relative likelihood of rejection responses after happy facial expression primes compared to other facial expressions, \( Q_{\text{prime}}(3) = 33.96, p < .001 \). The results of this analysis (summarized in the bottom part of Fig. 2) indicate that happy facial expressions did not affect the likelihood of a rejection response at Prime Durations of 17 ms (\( OR = 1.004 [0.847, 1.190] \)) and 33 ms (\( OR = 1.131 [0.868, 1.473] \)), but the hypothesized decrease in likelihood of a rejection-related response for happy, relative to other facial expressions was found at Prime Durations of 50 ms (\( OR = 0.620 [0.551, 0.698] \)) and 67 ms (\( OR = 0.647 [0.559, 0.748] \)). Accordingly, inclusion of Prime Duration as a moderator decreased the observed heterogeneity, although not to non-significance: \( Q_w(15) = 56.65, p < .001 \). In sum, our first hypothesis, that happy facial expressions would be more strongly associated with acceptance than other facial expressions, was supported for Prime Durations of 50 ms and up.

**Hypothesis 2.** Angry facial expressions and rejection.

Our second hypothesis was that anger would be more strongly associated with rejection than other negative facial expressions. For this hypothesis to be corroborated, the \( OR \) of the second contrast should be higher than 1, which indicates that the likelihood of a rejection-related response is greater after angry facial expression primes than after sad and fearful facial expression primes. The initial random-effects without moderators showed the predicted effect (see Fig. 3 for a summary) to be small but reliable, \( OR = 1.033 [1.001, 1.066] \), \( p = .043 \). There was little evidence of heterogeneity among the effect sizes (\( Q_w(18) = 8.96, p = .961 \)), which suggests that the effects were comparable across conditions. Thus, no further moderation analyses were conducted. In sum, our second hypothesis, that angry facial expressions would be more strongly associated with rejection than other negative facial expressions, was supported.

### Remaining contrasts

The final two contrasts that were specified in the models were also subjected to the meta-analytic procedure. Note that no hypotheses were specified for these contrasts. They are reported to provide a complete picture of the results. The relevant forest plots are available as Supplementary materials online.

The third contrast tested all negative emotional facial expressions (i.e., angry, fearful, and sad facial expressions) against neutral facial expressions. A model without moderators indicated that negative emotional facial expressions increased the likelihood of a rejection response relative to neutral facial expressions, \( OR = 1.073 [1.019, 1.129] \), \( p = .007 \). There was substantial heterogeneity among effect sizes, \( Q_w(18) = 33.74, p = .014 \). Only Prime Duration significantly moderated the contrast, \( Q_{\text{prime}}(3) = 18.46, p < .001 \). No effects were found at Prime Durations of 17 ms (\( OR = 0.968 [0.894, 1.048] \)) and 33 ms (\( OR = 0.875 [0.740, 1.035] \)). At Prime Durations of 50 ms (\( OR = 1.125 [1.067, 1.187] \)) and 67 ms (\( OR = 1.140 [1.071, 1.214] \)), however, negative facial expressions increased the likelihood of rejection-related responses compared to neutral facial expressions. The remaining heterogeneity was non-significant, \( Q_w(15) = 15.28, p = .431 \).

The results for the fourth contrast did not show a difference between the likelihood of a rejection-related response after sad facial expressions, compared to fearful facial expressions (\( OR = 1.013 [0.953, 1.077] \), \( p = .669 \)), with little evidence of heterogeneity among effect sizes (\( Q_w(8) = 9.91, p = .272 \)).

### Auxiliary analyses

In addition to these hypotheses tests, we explored whether the effects described above were moderated by participant sex. We found that it did not moderate any of the contrasts. Then, we explored whether response latencies were affected by the facial expression primes, by
running the same analyses as described above, but substituting response for response latency as the dependent variable. Happy facial expressions were responded to slightly faster than all other facial expressions ($\beta = -0.019 \ [ -0.037, -0.002]$), but no other differences were found. We are reluctant to interpret this effect, because it is small and does not relate in a meaningful way to the hypotheses.

**Discussion**

Drawing on social-functional accounts of emotion (Fischer & Manstead, 2008; Frijda & Mesquita, 1994; Keltner & Haidt, 1999; Van Kleef, 2009) and the evolutionary relevance of group life (Cosmides & Tooby, 1992; Dunbar, 1992), we developed the idea that facial
emotional expressions signal different degrees of acceptance versus rejection. A meta-analysis of six studies based on the AMP (Payne et al., 2005) supported our two hypotheses: Happy facial expressions were more strongly associated with acceptance than other facial expressions, and angry facial expressions were more strongly associated with rejection than other negative facial expressions. These associations emerged reliably at Prime Durations of 50 ms and higher, and they occurred regardless of whether the acceptance/rejection dimension was operationalized in a direct (acceptance/rejection), or indirect (warm/cold, close/distant) way. Together, these findings support the idea that observers use emotional expressions as immediate social signals regarding their own standing in the group.

The current findings have implications for theorizing on the social functions of emotions across different contexts. There is an increasing awareness that emotional expressions play a role in regulating social behavior (Keltner & Haidt, 1999) and engendering social influence (Van Kleef, Van Doorn, Heerdink, & Koning, 2011), but the ways in which such influence comes about is imperfectly understood. Our findings suggest that expressions of anger and happiness bring about social consequences in part because they emit signals of impending rejection versus acceptance, which may motivate behaviors aimed at securing acceptance versus self-expression and development. Viewing anger as a rejection signal thus helps to understand how the social consequences of anger may differ depending on contextual factors that determine the motivation to be reaccepted. For instance, it helps explain why angry expressions may both drive people apart by evoking aggressive tendencies (e.g., Van Kleef & Côté, 2007) and have the potential to draw people together by eliciting conformity from a deviant group member when this individual is highly dependent on the group (Heerdink et al., 2013). Furthermore, it helps explain why people are more likely to be categorized as outgroup members versus ingroup members when they express anger compared to happiness (Dunham, 2011).

Generally speaking, effects were found at presentation times of at least 50 ms. Although this is a shorter Prime Duration than is typically used in AMP studies (typical presentation times range from 75 ms up to 1500 ms), this finding is interesting in the context of the ongoing discussion about mechanisms that underlie AMP effects (e.g., Blaison et al., 2012; Gawronska & Ye, 2014; Payne et al., 2005). AMP effects are thought to be produced by a combination of ‘hot’ (or affective) mechanisms, in which the primes induce affect that is subsequently misattributed to the targets (Payne et al., 2005), and ‘cold’ (or semantic) mechanisms, in which the primes induce conceptual activation that is misattributed to the target (Blaison et al., 2012). Both mechanisms may simultaneously contribute to AMP effects (Gawronska & Ye, 2014). Consequently, it is difficult to pinpoint exactly which mechanism is responsible for our effects, because affective and semantic mechanisms would produce effects in the same direction (i.e., happy expressions could lead to either feeling accepted or thinking about acceptance, and both would elicit the same response). However, one set of AMP studies pitted these mechanisms against each other (Blaison et al., 2012). In these studies, angry and fearful expressions were used as primes, and the response alternatives were ‘angry-evoking’ and ‘fear-evoking’. Angry expressions were found to increase the proportion of ‘angry-evoking’ responses relative to fearful expressions. This is more in line with a semantic explanation than with an affective explanation, because the latter would predict an increased proportion of ‘fear-evoking’ responses instead. Thus, cold mechanisms may be relatively more potent than hot mechanisms. This suggests that our effects may have been primarily driven by cold mechanisms.

If it is indeed primarily cold, conceptual activation that underlies our findings, the findings with regard to Prime Duration may represent a lower bound for the presentation time to elicit sufficient conceptual activation to find cold AMP effects. As the encoding of the stimulus takes place during the presentation, a minimum Prime Duration of 50 ms may be required to allow cognitive flexibility in the encoding of facial expressions; that is, to categorize facial expressions using a contextually determined dimension (in this case, acceptance and rejection; analogous to how emotion words serve as context while categorizing facial expressions, Barrett, Lindquist, & Gendron, 2007). However, as noted by an anonymous reviewer, the conclusions drawn from this finding should be qualified by the fact that the presentation time only represents a fraction of the total time that participants took to produce a response (mean latencies in the current studies ranged from 665 ms to 962 ms). Future research may test these speculations more directly.

The finding that the warm/cold Response Dimension produced similar results as the other response dimensions is relevant to understanding how people process other people’s emotional expressions. As explained in the introduction, abstract concepts are grounded on more concrete representations, and these conceptual groundings are often revealed by linguistic metaphors. Such metaphors link rejection and acceptance to coldness and warmth (e.g., giving the cold shoulder, or being a warm person). However, metaphors about anger often link anger to higher temperatures instead (e.g., boiling with anger). Importantly, these metaphors appear to reflect the bodily experience of anger, as this experience is associated with higher skin temperature than neutrality and other emotions such as sadness (e.g., Nummenmaa, Glerean, Hari, & Hietanen, 2013). The finding that angry facial expressions increased the likelihood of ‘cold’ rather than ‘warm’ responses suggests that people focus on the social implications signaled by others’ facial expressions before attempting to understand the emotional experience of the other person (e.g., by simulation, Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005).

It is important to consider to what extent our findings may be accounted for by alternative explanations. As indicated before, the word pairs that we used to tap into acceptance versus rejection also vary on a valence dimension, and the same is true for the emotions we examined. One might therefore argue that our findings reflect a valence effect. Indeed, the negative emotional expressions (anger, fear, sadness) were generally more strongly associated with the negative outcome (rejection) than the neutral facial expressions, which were in turn more strongly associated with the positive outcome (acceptance). A more fine-grained analysis, however, reveals that various aspects of our findings do not fit a simple valence account.

First, angry facial expressions produced more rejection-related responses than fearful and sad facial expressions. A valence-based explanation may only account for these findings if angry facial expressions are construed as more negative than fearful and sad facial expressions. Yet, these emotions are not typically differentiated in terms of negativity (Russell, 2003). Moreover, the validation data for the ADFES (as discussed in the Method section) indicate that angry facial expressions were construed as less negative than sad facial expressions (Van der Schalk et al., 2011, p. 914) — the opposite of the pattern that would be required to produce our findings under a valence account.

Second, we found the predicted effects across three different conceptualizations of the acceptance/rejection dimension, which are not all equally clearly valenced. For instance, closeness may be either negative (when it concerns a crocodile or a foe) or positive (when it concerns a kitten or a friend). Indeed, the valence ratings that we collected in a separate study (see Appendix A) clearly indicate that acceptance/rejection is more explicitly valenced than warm/cold and close/distant. If valence were driving our effects, one would expect smaller effect sizes for more ambiguously valenced response options. Yet, we found that Response Dimension did not reliably moderate our findings,3 which indicates that the effect sizes were comparable between the more ambiguously valenced

---

3 Although we found marginally significant evidence for moderation of the happy vs. other facial expressions contrast by response dimension, the pattern of moderation does not meaningfully correspond either to the main hypotheses or to an alternative account in terms of valence. Because the evidence was only marginally significant, and because we found it in only one of the four contrasts, we do not attach much weight to this moderation.
response dimensions (warm/cold, and especially close/distant) and the most clearly valenced Response Dimension (acceptance/rejection). Thus, considering that the pattern of our findings matches neither the differences in valence of the stimuli (facial expressions), nor the differences in valence of the response dimensions, we believe that our effects are better understood in terms of the specific social signals of discrete emotional expressions, consistent with social-functional accounts of emotion.

Although Contrast 1 provided strong support for Hypothesis 1, the heterogeneity among effect sizes was not completely accounted for. This may indicate that there is a systematic source of variance (i.e., a potential moderator) that is omitted from the model. Since all experiments were identical, except for the manipulations of Response Dimension and Prime Duration, there were no more study-level moderators we could test. Thus it remains unclear what may have caused these variations, and why heterogeneity was only observed for Contrast 1. Future replication may elucidate whether this heterogeneity is a spurious finding or requires further inquiry. For present purposes, heterogeneity is unlikely to have influenced our conclusions, as the random-effects meta-analytic model takes heterogeneity into account (Borenstein et al., 2010).

In this article, we have provided a minimal demonstration of the proposed social-signaling function of angry and happy expressions: angry expressions signal rejection, and happy expressions signal acceptance. Furthermore, we have shown that these social implications of facial expressions are accessible nearly instantly to an observer; 50 ms and 67 ms are only 1/20th, or 1/15th of a second — less than the blink of an eye (Kaneko & Sakamoto, 1999). The immediacy of these effects points to the fundamental importance of the signals conveyed by emotional expressions as regulators of social life.

Acknowledgments

We would like to thank Juliane Degner for sharing the materials on which these experiments were based. This research was facilitated by a grant from the Netherlands Organisation for Scientific Research (NWO 452-09-010) awarded to the second author.

Appendix A

To compare the extent to which the three different Response Dimensions are linked to differences in valence, 53 participants (no demographic information collected) rated one (randomly selected) alternative from each of the acceptance/rejection, warm/cold, and close/distant word pairs. Thus, they completed three trials in total; the order of the word pairs was randomized. Ratings were made on a 7-point scale (1 = negative, 4 = neutral, 7 = positive). Using mixed-effect regression analysis, we tested whether the difference in valence of two words within a word pair differs between word pairs (e.g., the difference in valence of ‘accept’ and ‘reject’ is smaller or greater than the difference in valence between ‘warm’ and ‘cold’). Thus, we tested the Word pair (dummy-coded, using accept/reject as the reference condition) × (acceptance-rejection-related) Alternative interaction. Significance of the results was determined by bootstrapping (10000 resamples) using the bootME function in the lme4 package (Bates et al., 2013) for R (R Core Team, 2013).

A simple effect of Alternative indicated that within the accept/reject word pair, ‘accept’ was rated as more positive than ‘reject’, β = 2.16, 95% CI [1.90, 2.43], p < .001. More importantly, a significant interaction indicated that both the warm/cold (β = −0.48, 95% CI [−0.80, −0.18], p = .003) and close/distant word pairs (β = −0.94, 95% CI [−1.25, −0.63], p < .001) differed significantly less in valence than accept/reject. Follow-up analyses indicated that the differences were mainly prominent for the rejection-related words: ‘reject’ was rated as more negative (M = 1.48, SD = 0.64) than ‘cold’ (M = 2.77, SD = 0.96) and ‘distant’ (M = 2.83, SD = 1.36), both ps < .001. For the acceptance-related words, ‘accept’ (M = 5.92, SD = 1.02) was found to be more positive than ‘close’ (M = 5.33, SD = 1.05, p = .041), but not different from ‘warm’ (M = 6.23, SD = 0.92, p = .297). Thus, the three Response Dimensions differ in how strongly they are linked to valence.


