The Commitment Function of Angry Facial Expressions

Lawrence Ian Reed1, Peter DeScioli2, and Steven A. Pinker1
1Department of Psychology, Harvard University, and 2Department of Political Science, Stony Brook University

Abstract
What function do facial expressions have? We tested the hypothesis that some expressions serve as honest signals of subjective commitments—in particular, that angry faces increase the effectiveness of threats. In an ultimatum game, proposers decided how much money to offer a responder while seeing a film clip depicting an angry or a neutral facial expression, together with a written threat that was either inherently credible (a 50-50 split) or less credible (a demand for 70% of the money). Proposers offered greater amounts in response to the less credible threat when it was accompanied by an angry expression than when it was accompanied by a neutral expression, but were unaffected by the expression when dealing with the credible threat. This finding supports the hypothesis that angry expressions are honest signals that enhance the credibility of threats.

Keywords
facial expressions, emotions, threat, punishment

Received 12/19/13; Revision accepted 3/16/14

Angry expressions are among the most common signals in human social life. An unruly child, a colleague late for a meeting, or a business partner taking more than his or her fair share is likely to elicit an unmistakable glower composed of slanted brows, glaring eyes, and tight lips. An angry face can escalate hostilities and lead to destructive behavior, including violence. Intuitively, people sense that angry expressions may also keep the peace by conveying social expectations and personal boundaries: The threat of aggression can avoid actual aggression. This intuition, however, is theoretically incomplete and empirically unsupported. In the experiments reported here, we tested a particular version of it: that angry expressions function as credible threats in bargaining.

The intuitively obvious theory of facial expressions is that they signal to conspecifics current emotional states (Izard, 1971) or action tendencies (Fridlund, 1994; Kraut & Jonston, 1979). For example, expressions of anger communicate a signaler's intention to approach the signals' receiver aggressively (Adams, Ambady, Macrae, & Kleck, 2006; van Honk & Schutter, 2007; Yik, 1999), whereas expressions of happiness signal the absence of threat (Ramachandran, 1998) and cooperative intent (Reed, Zeglen, & Schmidt, 2012).

The problem with a simple signaling theory is that it is not obvious why it is adaptive for an organism to telegraph its internal state or impending action, or why a perceiver should take such signals at face value. Darwin (1872/1998) argued that expressions may originally have had a noncommunicative adaptive function, namely, preparing the organism to respond to environmentally recurrent stimuli by regulating sensory intake. Widened eyes in expressions of fear, for example, increase peripheral vision to facilitate identification of potential threats (Susskind, Lee, Cusi, Grabski, & Anderson, 2008). Similarly, the constriction of the nose and mouth in expressions of disgust functions to expel and prevent internal exposure to noxious stimuli (Chapman, Kim, Susskind, & Anderson, 2009). Darwin argued that even these functions may no longer apply in modern humans: Expressions today are evolutionary vestiges, automatized through Lamarckian use and disuse. Other features of expressions merely represent an overflow of energy

Corresponding Author:
Lawrence Ian Reed, Department of Psychology, Harvard University, Cambridge, MA 02138
E-mail: lawrenceianreed@fas.harvard.edu
from an aroused nervous system (Darwin, 1872/1998; Fridlund, 1992).

A breakthrough in specifying a communicative function for facial expressions has come from game theory: the hypothesis that expressions communicate an individual's commitments to carry out threats or promises in bargaining situations (Hirshleifer, 1987; Schelling, 1960). A commitment ties one's hands, making it difficult to renege on the threat or promise. The goal of commitment is to influence the partner's choices by leaving him or her the last opportunity to decide the outcome (Schelling, 1960). This lack of freedom of choice can, paradoxically, improve one's bargaining position.

An inherent problem in making a commitment is how to signal it credibly. Many commitments are enforced by third parties, such as legal systems or constraints in the situation (e.g., burning one's bridges). Alternatively, a subjective commitment is enforced by the individual's internal emotional state (R. H. Frank, 1988; Nesse, 2001; Pinker, 1997): The actor is in the throes of an emotion that makes it difficult to renege, even if carrying out the threat or promise is costly. Because this emotional enforcement is internal and unobservable, subjective commitments must be credibly displayed to be effective. This need for a credible display supplies the elusive communicative function for facial expressions.

Two sets of findings support the hypothesis that expressions credibly signal subjective commitments. First, expressions are related to emotional experiences (Rosenberg & Ekman, 2005) and convey information about intentions and action tendencies (Fridlund, 1994; Kraut & Jonston, 1979; Yik, 1999). Second, expressions are associated with involuntary neurological mechanisms (Rinn, 1984) and are difficult to fake (Ekman, Roper, & Hager, 1980). This association between motivational and expressive systems could allow expressions to be predictive of future behavior and hence credible signals.

A major problem for this credible-signaling hypothesis, however, is that individuals can gain advantages through deceptive signaling (Dawkins & Krebs, 1978), fooling perceivers into believing a threat or promise but then reneging if the costs are too high. If actors can dissociate facial displays from underlying psychological states, then they could gain a strategic advantage by deploying these signals deceptively (“bluffing”). In this case, facial expressions would be “cheap talk,” and perceivers would evolve or learn to ignore the signals. Only empirical tests can establish whether humans in fact treat emotional expressions as credible signals of bargaining resolve. We report two such tests, focused on expressions of anger.

Anger is an aversive emotion experienced in response to an undervaluing of one's welfare by another individual (Ortony, Clore, & Collins, 1988; Sell, Tooby, & Cosmides, 2009). Its function is to motivate angry parties to approach the angering parties and incentivize them to recalibrate their valuation (Carver & Harmon-Jones, 2009; Fischer & Roseman, 2007; van Honk & Schutter, 2007). These incentives—withstanding cooperation, dissolving the relationship, and inflicting harm—must be honestly communicated to be effective. Thus, expressions of anger communicate confrontational and unyielding states (Hinde, 1985; Keltner & Haidt, 2001; Morris & Keltner, 2000) and the intention to approach with aggressive intent (Adams et al., 2006; Yik, 1999).

Anger and threats often occur in bargaining situations when two individuals have conflicting preferences about the division of resources, such as the sale price for a trade, but also have a common interest in coming to an agreement rather than walking away from a deal. In such situations, individuals can use the threat of rejection to demand a particular share of the pie.

### The Current Investigation

We used a modified ultimatum game to investigate the effects of angry expressions on the credibility of threats in bargaining. In the standard ultimatum game, a proposer decides how much money to offer a responder, the responder decides whether to accept or reject the offer, and if the offer is rejected, both players receive nothing (Guth, Schmittberger, & Schwarzer, 1982). Though a rational, self-interested responder should accept any offer greater than zero, responders are commonly spiteful, rejecting offers they perceive as unfairly low. This behavior has been interpreted as a manifestation of anger (Straub & Murnighan, 1995), and this interpretation is supported by the finding that rejections are accompanied by activity in brain areas associated with negative emotion (Sanfey, 2003). It is commonly thought that proposers are more generous than rationality predicts because they anticipate and avoid provoking the responder's anger. This idea implies that when communication is possible, responders can gain by credibly communicating their anger, yet the effectiveness of such signals has not been demonstrated. The mere existence of expressions associated with anger is not sufficient, as these expressions may be evolutionary vestiges, spillovers of arousal, or cheap talk. To test the credibility of angry expressions, we altered the ultimatum game such that before making an offer, the proposer viewed a demand purportedly issued by the specific responder (Experiment 1) or a typical responder (Experiment 2) who displayed either an angry or a neutral facial expression.

Further, to examine whether the expression specifically changed the threat's credibility (rather than changing mood or some other nonspecific factor), we also manipulated the inherent credibility of the threat by varying the amount demanded by the responder. In many
bargaining situations, negotiators converge on a 50-50 split, either because it is perceived as inherently fair or because it is a salient focal point (Schelling, 1960). Accordingly, research with ultimatum games shows that responders commonly accept offers of 50% and higher, whereas lower offers are increasingly likely to be rejected the further they deviate downward from 50% (Camerer, 2003). For example, List and Cherry (2000) reported (for $20 stakes) approximately 20% rejections for offers of 40% to 49%, 50% rejections for offers of 30% to 39%, and 70% rejections for offers less than 25%.

A threat by a responder to reject a lowball offer and accept an even split is thus inherently credible and need not be reinforced by an honest signal. In contrast, a responder who demands 70% of the pot on pain of spitefully rejecting lower offers is not making an inherently credible threat, because the responder can expect the proposer to see the offer as unfair, and because carrying out the threat is against the responder’s own interests. To prove that this threat is not a bluff, the responder needs an additional guarantor. According to the credible-signal hypothesis, that guarantor is anger, signaled by an angry expression, and it will be effective in inducing the proposer to make a more generous offer. Specifically, proposers will make a higher offer when an inherently noncredible threat (a demand for 70% of the pot) is accompanied by an angry expression than when it is accompanied by a neutral expression, but will be unaffected by these expressions when they accompany an inherently credible threat (50%). According to the alternative hypotheses (anger expressions as vestiges, spillover, and cheap talk), anger expressions will not affect proposers’ offers at all.

Experiment 1

Method

This experiment examined proposers’ offers in an ultimatum game with threats from responders. Before making an offer, each proposer was told that he or she would be paired with a specific videotaped responder. Threats consisted of a brief video clip of either a neutral or an angry facial expression paired with a written message demanding either 50% or 70% of the pot. This resulted in a 2 (expression) × 2 (demand) design.

Participants. Eight hundred seventy participants (571 male, 299 female) were recruited on Amazon’s Mechanical Turk (MTurk), an online crowd-sourcing Web site at which individuals sign up to complete tasks. It has been used in previous research in psychology and experimental economics and has a large and diverse subject pool (Buhrmester, Kwang, & Gosling, 2011; Horton, Rand, & Zeckhauser, 2011). Participants’ mean age was 30.21 (SD = 9.34); their racial distribution was as follows: 79% Caucasian, 7% African American, 11% Asian American, and 3% other. Participants were randomly assigned to be proposers or responders.

Procedure. Participants were given a consent form and a description of the procedure. Those in the proposer role read that they had a $1.00 endowment and could decide how much to offer to a responder. We chose to use a $1.00 endowment because it is similar in size to the endowments used in other online studies involving economic games and yields results comparable to those obtained in laboratory settings with greater stakes (Amir, Rand, & Gal, 2012). Proposers were told that the responder would decide whether to accept or reject the offer, and that if it was rejected, both players would receive $0. In addition to earning 40¢ for completing the study, participants could earn up to an additional $1.00 depending on the decisions that they and the assigned responders made. These additional earnings were paid using the “bonus” feature on MTurk. All participants answered three comprehension questions (correct responses were required for participation) and completed a single trial of the task (~5 min).

Prior to making their offers, proposers viewed a threat purported to come from the assigned responder, who was said to have “provided messages and been videotaped.” The threat consisted of a 6-s video clip and a written demand. We used video clips because they provide more information to perceivers than static images do (Ambadar, Schooker, & Cohn, 2005). Each clip depicted the same female actor, who had been instructed to create specific facial actions and expressions. She displayed either a neutral expression (no facial actions) or an angry expression (frowning eyebrows, intensely staring eyes, and a closed mouth with lowered corners; Ekman & Friesen, 1975). At the conclusion of the clip, the video screen turned black. Each clip was recorded at 30 frames per second. This produced a set of 180 sequential 640- × 480-pixel full-color images. Thus, the clip duration of 6 s was similar in length to the average 4 to 6 s reported for spontaneous expressions (M. G. Frank, Ekman, & Friesen, 1993; Schmidt, Ambadar, Cohn, & Reed, 2006).

Each clip was paired with a written demand (placed directly below the clip) for either an equal split of 50% each (the inherently credible threat) or an unequal split of 70% for the responder and 30% for the proposer: “If you offer me less than 50 cents [70 cents], I will reject your offer.” We chose a demand of 70% to represent a threat of borderline credibility. Our choice was based on an initial pilot study in which threats were composed of a neutral expression paired with a demand of 60%, 70%, 80%, or 90%; in that study, 70% demands elicited the most compliance.
Immediately following the presentation of the threat, proposers were asked, “How much money do you choose to transfer to the Receiver (0–100 cents)?” After responding to this question, they were asked to rate how happy and how angry the purported responder (i.e., the actor in the video) was, using 7-point scales (1 = not at all, 7 = extremely).

Proposers had been told that their partners had specified a minimum amount they would accept without forcing forfeiture of the entire pot. In actuality, participants in the responder role made their decisions for each possible offer using the strategy method (Selten, 1967), stating how they would react to each amount that a proposer might offer. After data collection (i.e., not in real time), proposers and responders were randomly matched, responders’ choices were implemented for the offers they actually received, and the payouts were calculated accordingly.

Results and discussion

As a manipulation check, we first examined proposers’ ratings of the purported responder’s (i.e., the actor’s) emotions. Proposers rated the responder in the anger video clip as more angry, $F(1, 431) = 17.87, p < .001$, and less happy, $F(1, 431) = 21.01, p < .001$, than the responder in the neutral video clip. The ratings of anger were not significantly affected by the size of the demand, $F(1, 431) = 4.43, p = .036$, and demand size did not interact with expression, $F(1, 431) = 0.28, p = .601$. The ratings of happiness decreased significantly with the size of the demand, $F(1, 431) = 7.95, p = .005$, though demand size did not interact with expression, $F(1, 431) = 0.86, p = .769$. For responders, the average minimum acceptable amount specified was $31.15¢ (SD = 19.32)$. Proposers’ offers did not differ by gender or ethnicity (all $p$s > .05), so we aggregated across these categories in subsequent analyses.

The data of interest are shown in Figure 1. Following demands of 50%, proposers did not give higher offers if they had viewed an angry expression ($M = 49.98\%, SD = 10.95$) than if they had viewed a neutral expression ($M = 50.55\%, SD = 11.17$). In contrast, following demands of 70%, they gave significantly higher offers if they had viewed an angry expression ($M = 60.92\%, SD = 11.23$) than if they had viewed a neutral expression ($M = 54.77\%, SD = 20.73$), $t(215) = -2.71, p = .007$. A 2 × 2 analysis of variance (ANOVA) revealed significant main effects for facial expression, $F(1, 431) = 4.19, p = .041$, and demand size, $F(1, 431) = 30.96, p < .001$, and, crucially, a significant interaction between facial expression and demand size, $F(1, 431) = 6.08, p = .014$.

In sum, the data support the credible-signal hypothesis. When faced with a dubious demand of 70% of the pot, proposers who viewed an angry expression offered more money than did those who viewed a neutral expression. When faced with a more credible demand of 50%, an extraneous guarantor was otiose, and proposers were unaffected by the facial expression in the threat.

Experiment 2

Method

In Experiment 1, we used deception, telling participants that the videos showed their actual partners. It is possible that the angry face caused proposers to make inferences about their specific partner (e.g., that she had a malevolent disposition) and did not operate solely as a signal of credibility. In Experiment 2, we tested whether we would observe the same effect if participants were not deceived about their partners: The person in the video was described as a typical responder rather than the specific assigned partner. We tested whether an angry expression would increase the perceived credibility of a threat even when it did not provide information about a specific partner.

Participants. Five hundred eighty-eight participants on MTurk (345 male, 243 female) were randomly assigned to be proposers or responders. Their mean age was 30.80 (SD = 10.08); their racial distribution was as follows: 79% Caucasian, 7% African American, 12% Asian American, and 2% other.
Results and discussion

As in Experiment 1, a manipulation check confirmed that proposers rated the responder in the anger video clip as more angry, $F(1, 290) = 51.59, p < .001$, and less happy, $F(1, 290) = 18.41, p < .001$, than the responder in the neutral clip. The ratings of anger were not significantly affected by the size of the demand, $F(1, 290) = 3.32, p = .068$, and demand size did not interact with expression, $F(1, 290) < 0.001, p = .995$. The same pattern held for ratings of happiness: no significant main effect of demand, $F(1, 290) = .468, p = .995$. The same pattern held for rat-

Procedure. The procedure was the same as in Experiment 1 with one exception: Rather than being told that the threat came from their assigned partner, participants were told that the threat was “typical of a Responder in this scenario” and that they would be paired with a randomly selected responder. Following data collection, proposers and responders were randomly matched. Each responder’s choice was implemented for the specific offer of the assigned proposer, and the payouts were calculated accordingly.

General Discussion

Results of both experiments support the hypothesis that angry facial expressions function as credible threats in bargaining situations: An angry expression resulted in increased offers when paired with a less-than-reasonable and less-than-plausible demand, but not when paired with a fair, plausible, and hence credible demand. The effects were strong enough to induce proposers to make generous offers (ones that favored their partners over themselves, which is highly unusual in ultimatum games), and they occurred whether the perceivers viewed the facial expression of a person described as their specific assigned responder or a person described as a generic responder. These findings suggest that angry expressions are not cheap talk but are perceived as honest signals of a person’s willingness to carry out a mutually harmful threat. Thus, the expression and perception of anger can function to avoid spiteful behavior in bargaining situations, benefiting both senders and receivers.

The data thus support the adaptive function of emotions proposed by R. H. Frank (1988) and Hirshleifer
(1987), who argued that the expressive components of emotion serve as credible signals communicating subjective commitment. In this case, angry facial expressions serve as an honest signal that one is motivated to act against rational self-interest and reject low offers. But the logic may apply to positive emotions as well: Other experiments suggest that smiles serve as credible signals of cooperative intent (Reed et al., 2012) and may guarantee promises against the suspicion of rational defection and double-crossing in the same way that angry expressions guarantee threats against the suspicion of bluffing.

One surprising aspect of our findings is that we observed an effect of an angry expression even though it was created using a directed facial-action task rather than spontaneously emitted. In the context of the experiment, this expression was in fact not honest but literally faked, yet the participants treated it as if it were honest. Future research can examine whether the effects would generalize to spontaneous expressions, as well as the sending of commitment signals, by observing responders’ threats in face-to-face interactions allowing spontaneous expressions of emotion.

The fact that bargaining offers are mediated by internal emotional states speaks not only to theories of the adaptive function of the emotions and emotional expressions but also to the study of bargaining. Experienced negotiators know that although negotiations are affected by rational self-interest, they are also affected by the parties’ emotional reactions, including perceptions about how much each side values the other side’s welfare. This is why, for instance, prospective home buyers are advised not to start with a lowball offer because it will “insult” the home owner, and salespeople feign (and sometimes honestly develop) warm and friendly relationships with their customers to increase the likelihood of profitable sales. These responses appear to be irrational quirks, and are exploited by seasoned negotiators, but they may have a rationale in the less structured negotiations that make up informal social cooperation. The study of the role of emotions in negotiation and bargaining thus offers the prospect of illuminating both irrational and rational responses.

Author Contributions
L. I. Reed, P. DeScioli, and S. A. Pinker developed the study concept and contributed to the study design. Data collection was performed by L. I. Reed and P. DeScioli. L. I. Reed and P. DeScioli performed data analysis and interpretation. L. I. Reed, P. DeScioli, and S. A. Pinker wrote the manuscript. All the authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests
The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References
on evolution and optimality (pp. 301–326). Cambridge, MA: MIT Press.