The Emotional Moves of a Rational Actor: Smiles, Scowls, and Other Credible Messages

Lawrence Ian Reed 1,* and Peter DeScioli 2

1 Department of Psychiatry, McLean Hospital, Belmont, MA 02478, USA
2 Department of Political Science, Stony Brook University, Stony Brook, NY 11794, USA; pdescioli@gmail.com
* Correspondence: lreed@mclean.harvard.edu; Tel.: +1-412-401-7658

Academic Editors: Roman Sheremeta and Eric Schniter
Received: 21 January 2017; Accepted: 28 March 2017; Published: 31 March 2017

Abstract: Many scholars turn to emotions to understand irrational behavior. We do the opposite: we turn to rationality and game theory to understand people’s emotions. We discuss a striking theory of emotions that began with the game theory of credible threats and promises, then was enriched by evolutionary biology and psychology, and now is being tested in psychological experiments. We review some of these experiments which use economic games to set up strategic situations with real payoffs. The experiments test whether a player’s emotional expressions lend credibility to promises, threats, and claims of danger or hardship. The results offer insights into the hidden strategies behind a warm smile, an angry scowl, a look of terror, and eyes of despair.

Keywords: emotion; game theory; facial expression; commitment; signaling

1. Introduction

A buyer makes a lowball offer on a house. The seller’s face swells with anger, their voice trembles, and they demand no less than the list price. When the buyer quibbles again, the seller throws a fit, storms out of the room, and slams the door behind them. Could this unruly display of outrage actually help the seller get their price?

A classic, brilliant analysis by the economist Thomas Schelling explains why a tantrum-throwing hotheaded seller might actually reap more profits than a calm, collected, rational one [1]. The secret is the idea of a strategic commitment, when a person purposely limits their own freedom in order to influence someone else’s decisions. For instance, a clever military general could burn the bridges behind them to convince the enemy they will never retreat, which could persuade the enemy to retreat instead. Or, an enterprising merchant could sign a legal contract with a lender stipulating that failure to repay the loan will be punished by law; by committing to their own punishment in this eventuality, the merchant could win the lender’s trust. Analogously, a seller whose outrage is out of control has committed to mutual destruction if their demands are not met, which in certain circumstances could command a better price. In each case, an actor ties their own hands to a specific course of action, limits their own freedom, and thus, paradoxically, improves their bargaining position.

Schelling is a game theorist but he also, accidentally, invented a startling and profound theory about the psychology of human emotions and emotional expressions. Only decades later did scholars begin to imagine that emotions might not only give a strategic edge, but that some emotions were sculpted by natural selection for precisely this function. Economist Robert Frank’s Passions Within Reason [2] combined ideas from game theory, evolutionary biology, and psychology to make a compelling case that passions like love, sympathy, jealousy, and rage were designed by natural selection to function as internal commitment devices. The theory of strategic commitment became the basis for a psychological theory of emotions [2–5].
Another few decades later, the commitment theory has seeped only slowly into actual psychological research on emotions. A small but growing set of studies have begun to test predictions of commitment theories for different emotions. For example, Gonzaga et al. [6] studied how nonverbal displays of love including smiles, gestures, and leaning forward are associated with the signaler’s feelings of love and their partner’s estimates of the signaler’s love. In another line of work, researchers examined how people use smiles to assure partners that they will cooperate [7–12]. A third set of studies looks at how participants’ feelings of guilt [13,14], blushing faces [15,16], and written apologies help to repair trust after a transgression [17].

Moreover, this research fits into a broader literature about how our emotions also perform many other strategic functions (besides commitments), such as evolutionary functions surrounding cooperation [18–20], reciprocity [21–23], dominance [24–26], and other social challenges (for review, see [27]).

Here we review a few of our own psychological experiments that test commitment theories for different emotions, focusing particularly on facial expressions of emotion [28–31]. We also discuss experiments about how emotions can lend credibility to other types of claims besides commitments, including bolstering claims of danger, hardship, misfortune, and other messages that might be doubted. These experiments use economic games to set up strategic situations with real payoffs on the line. The experimenter measures or manipulates (using a confederate) a player’s facial expression, which is observed by another player who then makes a choice in response. This basic setup provides a means to observe the strategic consequences of different emotional expressions under varying circumstances. The results can offer a few tips for clever bargaining and might even illuminate the very nature of human emotions.

2. Emotions Are Algorithms

The idea that emotions can function as strategies runs counter to the folk wisdom that emotions are mindless, irrational urges that usually cause worse decisions [32]. The notion that emotions are dumb, pointless, primitive, and reckless is a staple of Western philosophy running from the Stoics to Descartes and Spinoza, who viewed the emotions as unruly passions that should be reined in by reason and morality [33]. The same notion is found in the id and superego of Freudian psychology and it continues today in some strands of behavioral economics and political psychology, which tend to view psychology in general, and especially emotions, mainly as a source of bias and error [34]—the opposite of smart, rational, strategic moves.

However, this view of emotions as a source of bias is at odds with mainstream psychology [35,36]. The mind is the source of all experience and behavior, not only people’s errors and quirks. While some social scientists might feel satisfied to conclude that a person behaved rationally, to a psychologist such an observation is only the beginning of inquiry, not the end. To illustrate, when a child learns a language or a person visually perceives the depth of objects, they have acted in close accord with rational choice since the expected benefits of perceiving depth or learning language outweigh the costs. Yet, merely appreciating their rationality does not begin to uncover how these mental processes actually work. Rational choice, even when accurate, is an insufficient explanation of behavior, and the same applies to irrationality, which only negates an insufficient explanation.

Instead, mainstream psychology and cognitive science view the mind as a computational system composed of complex, specialized algorithms that learn language, perceive depth, weigh opportunities, assess dangers, reason about minds, recognize anger, evoke compassion, and innumerable other mental operations underlying our experience and behavior [5]. Irrational urges are no longer part of the proper lexicon of psychological science, any more than spirits and demons are.

Modern psychology is based on the computational theory of mind [5]. From this perspective, the mind is a control system that receives input from the senses, processes this information with evolved cognitive algorithms, and ultimately outputs the patterns of muscle contractions that comprise behavior. All mental phenomena, including experience, learning, emotions, culture, relationships, etc.,
Games 2017, 8, 18

come from information-processing algorithms. Especially now, at a time when computer algorithms are increasingly able to understand speech, recognize faces, drive cars, defeat experts in games, learn our taste in products, and so on, it is all the more clear that algorithms provide a powerful lens for understanding complex behavior.

Algorithms in both minds and machines show the hallmarks of rationality, intelligence, and creativity. Algorithms can steer around obstacles, weigh costs and benefits, make guesses about unknown values, statistically analyze data, simulate possible outcomes, randomize choices to confuse an opponent, or innovate novel solutions that have never been seen before, like a new melody, chess maneuver, or mathematical proof.

These abilities set algorithms in minds and machines apart from inanimate objects: rocks, mountains, clouds, planets, and stars do not avoid obstacles, calculate costs, make guesses, imagine possibilities, outwit opponents, or create new solutions. Rationality and purpose are the defining property of living organisms, sharply distinguishing them against the backdrop of a vast universe of lifeless, purposeless things. For this reason, evolutionary biology has always concentrated on the purposes and functions of biological traits [37–39]. The theory of natural selection explains how nonliving matter can under certain circumstances become enmeshed in a feedback loop of self-replication, which causes replicators to evolve adaptive traits like protective armor, deadly weapons, and a variety of algorithms for making decisions and orchestrating behaviors. Hence, natural selection explains the striking enigma of purpose and design in living things without a designer.

Emotions are purposeful algorithms like all of the contents of an evolved, living mind [27]. Emotions are essential survival tools for a human’s challenging expeditures through rugged terrain with scarce resources, stalked by predators and strangers, plagued by microscopic pathogens, haggled by merchants and partners, and tangled in political intrigues. For instance, fear protects us against predators and hazards, disgust steers us away from pathogens in rotten corpses and putrid water, and anger prepares us to fight for vital resources.

Emotions are rational in ways that should even impress an economist. In particular, emotions are designed to make good tradeoffs between costs and benefits. For instance, the emotion of fear weights the cost posed by a danger in proportion to the benefits of ignoring it, so that we feel less fear of a danger when there are greater benefits of confronting it. For example, a person who is starving feels less fear when hunting dangerous prey such as a poisonous snake, a massive elk, or an electric eel. Similarly, the emotion of disgust weighs the chance of exposure to pathogens against the benefits of eating questionable food, showing affection, or caring for a sick loved one [40]. Hence, we feel less disgust about putrid water when we are dehydrated, we feel less disgust about the saliva of an attractive mate, and we feel less disgust about a festering wound when helping a loved one—because in these cases the benefits partially offset the costs of pathogen exposure. (For illustrating examples of tradeoffs in fear and disgust, see the popular television series Naked and Afraid about survival in the wilderness.)

One colorful example comes from decades of research on fear in lizards [41,42]. This research shows that a lizard’s decision to flee an approaching predator depends on costs and benefits, such as their current distance from a safe hideout, which determines the cost (chance of death) if the predator gives chase. This research surrounding ‘optimal escape theory’ finds that lizards and many other animals flee when the expected cost, which increases as the threat gets closer, exceeds the benefits of staying put. The lizard’s fear is a paragon of rational choice.

There is an important respect in which emotions differ from prototypical rationality. Namely, a person often does not know the underlying purpose of their emotion [43]. For instance, before science discovered microorganisms, no one could have known that the feeling of disgust aims to prevent infection by pathogens. Still in modern times, people feel disgust toward things like a dead cockroach even if they consciously know it has been sterilized [44]. Moreover, our emotions can go against our conscious purposes. For instance, someone might wish to get on a plane but they are prevented by a fear of heights. This could appear irrational but if you think about it, the fear algorithm
has made a very reasonable estimation that a height of thirty thousand feet poses a great danger. It is not a purposeless and arbitrary urge, but shows a clear goal of protection, albeit at odds with the person’s goals. We also cannot fault our fear mechanisms too much if they ignore our conscious beliefs. Though we might be correct about planes, our conscious beliefs are mistaken in many other matters of safety (climbing ladders, riding horses, challenging a powerful rival), and fear better protects us by overriding these foolhardy beliefs.

Herein lies the likely origin of the myth that emotions are irrational. It is not that our emotions are actually pointless, but that we are often in the dark about their goals, and yet still under their sway. This creates the conscious experience that our mind has been hijacked by a powerful external force. In these cases, a person or poet could say they were overcome by irrational feelings and this would be fairly accurate, since they do not understand them; they could convey the same message of incomprehension by saying they were possessed by demons. However, such an account is not suitable for scientific theories of behavior, which are tasked with understanding both conscious and unconscious mental processes. From a scientist’s perspective, a person’s emotions are quite orderly, purposeful, and intelligent. Their magnitudes and proportions fit closely to the exigencies of the situation. To survive the rigors of evolutionary competition, emotions had to be smart, precise, and efficient in pursuit of goals, even if their aims can differ from our own.

3. Facial Expressions as Guarantors of Promises and Threats

If emotions are algorithms that routinely calculate costs and benefits, then perhaps it will seem less impossible that they could also perform strategies, as Frank [2], Hirshleifer [3], Nesse [4], and others have suggested. We focus particularly on how facial expressions of emotion could help credibly communicate messages.

Some of our emotion algorithms include characteristic expressions that accompany them. In Darwin’s original treatise on emotions, he theorized that emotional expressions have mainly physiological functions, preparing the face to better respond to challenges in the environment [45]. Modern research expands on this theory with the idea that emotional expressions also function to communicate with others [46]. Importantly, theories about communication also raise additional adaptive problems surrounding deception and credibility, which brings a strategic element to the study of expressions.

We look at how different emotional expressions can help in strategic games by lending credibility to two types of commitments, promises and threats, and then next, claims of danger and hardship.

3.1. Smiles and Promises

In game theory, a person makes a promise to commit to a cooperative action even if a selfish action would yield greater payoffs for themselves [1,47]. A person who merely states a promise is not too credible because they could simply break their word if it suits them when the time comes. However, if a person had an emotional commitment mechanism which binds them to their promise, like fear binds us from dangers, and, critically, if they can display this commitment to the promise with a facial expression such as a smile, then they could enhance promise’s credibility.

In previous experiments [28], we tested this idea by examining the verbal promises and facial expressions of participants who played a game of cooperation, the prisoner’s dilemma. In the game, each player decides whether to cooperate, benefiting their partner, or to defect, earning more payoffs for themselves. Before making their decisions, participants could speak freely with the other player about what they both would choose. During these pregame interactions, we observed and measured participants’ verbal promises and facial expressions, especially their smiles.

We recorded participants’ faces with a video camera. After the study, we coded their expressions using the Facial Action Coding System (FACS) [48]. We particularly looked for indicators of smiles and FACS distinguishes different features for genuine, felt, or Duchenne smiles versus more contrived, non-felt, non-Duchenne smiles through the presence of the orbicularis oculi muscle (the so-called
Duchenne marker) [49,50]. Overall, we coded occurrences of felt smiles, non-felt smiles, brow furrowers, lip pressers, and dimples. We examined whether participants’ smiles, both felt and non-felt, really did predict whether they chose to cooperate, and also whether their partner who received the smile was more likely to cooperate as well.

We found that players who showed more smiles, both felt and non-felt, were more likely to cooperate (see Figure 1). This is consistent with the idea that smiles provide a predictive signal of cooperation. For players who received felt or non-felt smiles, they were not necessarily more likely to cooperate (a nonsignificant trend for felt smiles and only marginally greater for non-felt smiles). This observation could mean that recipients remained skeptical of these smiles, or alternatively, that the recipient believed the smile but saw it as an opportunity to take advantage of a cooperator.

![Graph](image_url)  
**Figure 1.** Mean duration (seconds) of Duchenne smiles during the discussion period. Error bars represent ± 1 SE. Adapted from [28].

The observation that smiles are predictive of cooperation has also been found in other research [10]. Moreover, in contrast to our findings, some experiments have also found evidence that receivers of more smiles are more likely to cooperate in return [9–12,51].

### 3.2. Anger and Threats

A person makes a threat to commit to a harmful action against the recipient if their conditions are not met, even if following through with it damages the threatener themselves [1,47]. Just like promises can easily be broken, a person who only states a threat is not too credible, especially if they would suffer too if the threat was realized, like if they had to quit their job or go to prison for assault. The emotion of anger might provide an extra guarantor that the threatener will actually do what they said, even if it leads to mutual destruction. If anger sets an internal commitment device, then people also need to display this emotional commitment if it is to influence the recipient.

In previous research [31], we investigated whether angry facial expressions serve this function by lending credibility to threats of mutual harm. To test this hypothesis, participants played a bargaining game, the ultimatum game, in which a proposer offers a split of a pot of money and a responder decides whether to take the offer or to destroy the whole pot, leaving both players empty-handed [52]. We conducted the game online with one dollar stakes, and our version had a few important twists. The responder could first send to the proposer a written message along with a brief video of their
expression, before the proposer made their offer. This allowed the responder to make a demand. Moreover, the responder was actually a confederate actress who sent a preset message and was trained to show either a genuine angry expression or a neutral expression.

We also varied the inherent credibility of the demand, by making use of a common finding in bargaining research that responders, even without a threat, frequently reject offers less than 50%, even though it means they will earn nothing instead. So, in one condition, the responder made an inherently credible demand for 50%. In another condition, the responder made an outrageous demand, saying that they would accept no less than 70%. We designed this threat to be less credible: would the responder really choose to receive no money rather than accept an offer of 50% or 70% of the pot?

For the responder’s outrageous demand of 70%, the proposer offered more money, on average, when the responder showed an angry facial expression than when their face was neutral. That is, the responder’s questionable demand was seen as just cheap talk unless it was accompanied by an additional guarantor of an angry face (see Figure 2).

![Figure 2](image-url)

**Figure 2.** Mean proposer offer by facial expression (neutral or angry) and the size of the demand (50% or 70%) in the threat. Error bars represent ± 1 SE. Reprinted from [31].

4. Facial Expressions as Credible Claims of Hardship and Danger

Emotional expressions might lend credibility to other claims too, besides commitments. Any time a person, or another organism, communicates with someone else, there is a chance the speaker is lying, especially if they might have conflicting interests with their audience. Hence, the credibility theory for emotional expressions could extend to other types of messages such as claims of hardship or danger.
4.1. Sadness and Loss

People experience all sorts of hardships including injury, famine, disease, unemployment, and other mishaps and misfortunes that lead to feelings of sadness [53]. Those who suffer hardship often seek help from others by pleading their case. However, these messages could be deceptive. For instance, a student who says they lost their paper might not have written it at all, or someone who claims unemployment benefits might actually be working under the table.

In previous research [30], we investigated whether sad expressions might enhance credibility for claims of hardship and loss. Everyday observations and previous research shows that people feel sadness in response to loss [53]. Some researchers have hypothesized that sad expressions signal these losses in order to get help from others [4,54–57].

We tested whether sadness enhances credibility by using a game about sharing, called the dictator game [58], in which a player called the dictator decides how much of a pot of money to give to recipient. Our version of the game was embellished with a story to set the stage for a questionable claim of hardship. Participants played different roles in the story. They read that although the story was fictional, the interaction and the resulting payoffs were real, and they would earn actual money based on their decisions.

In the game, two players go fishing for the weekend. Each player learns that they caught 90 fish (worth 1 cent each), but a storm struck and each player has a 50% chance that they will lose their catch in the rough seas on the way home. If one player suffers a loss, then the other one decides how many of their fish to share with the unlucky one.

Then, the first player finds out that they made it home successfully with all of their fish. Next, they receive a message from the second player, stating that all of their fish were lost in the storm, along with a brief video of their facial expression. The first player could not find out if the second player’s claim of loss was actually true, which was explained in the game instructions, so it was up to the first player whether they believed the message. In reality, the second player was a confederate actress who sent a preset message claiming a loss, and she was trained to show a genuine sad expression or a neutral expression in the video. Finally, in response to the message, the first player decided how many fish (0–30) to send to the other player.

We found that participants shared significantly more fish (money) with their partner when they showed a sad expression in the video compared to a neutral expression. Specifically, participants sent on average about 17 fish (standard deviation = 13) to the sad fisher, whereas they sent about 11 fish (standard deviation = 19) to the fisher with a neutral expression. This result supports the hypothesis that sad expressions add credibility to statements of loss, and more generally, support credibility theories for emotional expressions [2,3,5].

An alternative interpretation is that a recipient’s sad expression automatically activates an observer’s empathy and increases their generosity, apart from their judgments of credibility. We examined this possibility in an additional experiment. We ran the same design with one key difference: participants were told by the experimenter that their partner really lost their fish, removing any doubts about the credibility of the message. We found that under certainty, participants no longer shared more money with the recipient who showed a sad expression compared to a neutral expression. This finding indicates that sad expressions did not increase generosity in general, but specifically operated to enhance the credibility of loss when it was uncertain.

4.2. Fear and Danger

Danger lurks around every corner: road hazards, rabid beasts, hidden poisons, stalking strangers, greedy corporations, divisive demagogues, marauding invaders. However, those who warn us of danger could be lying to use our fear against us. They could sell us unnecessary repairs, bogus remedies, useless insurance, or even costly wars. Hence, as illustrated by stories like The Boy Who Cried Wolf, warnings of danger pose a problem of credibility for both sides of the message [59,60].
In previous work [29], we conducted experiments to see whether facial expressions of fear might help bolster a warning’s credibility. To look at this, we examined how participants respond to warnings of danger paired with a fearful or neutral face.

We designed an incentivized lie detection game in which participants guessed whether a speaker was lying or telling the truth about a danger. We used a simple, colorful story to set the scene for the task. Participants read that they would play roles in a fictional scenario with a real partner and real payoffs for their judgments. Participants played the role of an explorer traveling with a scout, played by another participant who was actually a confederate actress.

A band of warriors was coming after the explorer to take their gold. Only the scout knew, after climbing a tree, whether the warriors were coming from the east or west. Moreover, the scout could either be loyal or a traitor, randomly determined by drawing a card. The scout reports the direction to the explorer, and the explorer must decide if the scout is lying or truthful. If the explorer is correct, they escape the warriors and keep their gold, worth real money in the study (50 cents). If the explorer guesses incorrectly, they are captured and lose their money.

Before deciding whether the scout is lying, participants saw a video clip showing the scout’s message (“The warriors are coming from the east”) paired with a fearful or neutral expression. Moreover, to make participants initially skeptical, they read that there was a 25% chance that the scout was assigned to be loyal and truthful, and a 75% chance the scout was a traitor instructed to lie.

We found that 25% of participants thought the scout with a neutral expression was telling the truth. In contrast, when the scout showed a fearful expression, 39% of participants judged that they were telling the truth, a significant increase.

These findings support the hypothesis that fearful expressions increase the credibility of claims of danger. However, another possibility is that when participants see a fearful face they become more credulous for any message rather than a message of danger specifically. To test this possibility, we examined a case where a speaker claims that a danger is absent rather than present. If fear boosts the credibility of danger specifically, then it will not do so for a claim that danger is absent.

We designed a new lie detection task with a different backstory to set the stage for a questionable message that danger is absent. Participants played the role of a sheriff while their (confederate) partner played the role of a merchant who might be hiding stolen bags of gold. Participants viewed a video clip of the merchant who claimed that a danger was absent (“There is no gold in this cart”) while showing either a fearful or neutral expression. As in the previous game, participants read that there was a 25% chance the merchant really had no gold and was honest and a 75% chance they really had the gold and were lying. If the sheriff guessed correctly, they recovered the gold and earned a real reward (50 cents). If they guessed incorrectly, then they lost the gold.

In this case, participants were no more likely to believe the merchant with the fearful expression than the merchant with the neutral expression (27% versus 22% believed the message, respectively, a nonsignificant difference). Hence, a fearful expression bolsters the credibility of claims of danger, specifically, rather than other messages such as oppositely claiming no danger.

This observation could point to a hidden cost of credible emotional expressions. Fear can bolster a claim of danger but it might simultaneously make it more difficult to hide a danger. For instance, smugglers of contraband could find it difficult to suppress emotions of fear when interrogated by the border patrol. In this way, fearful expressions that are designed to automatically and credibly signal a danger would also inadvertently give away the smuggler’s hidden cargo. Similarly, a political leader who tries to calm a panicked citizenry might accidentally betray themselves by expressing the real danger ahead.

5. Discussion

The idea that emotions can function as commitments is a striking insight that integrates psychology, economics, evolution, and games. We reviewed some previous research (see Introduction)
and a few of our own experiments that test these ideas for different emotional expressions, and we hope they will receive increasing attention in future research on the psychology of emotions.

Strategic theories of emotions also provide an opportunity to reflect on how rationality is related to psychology. Many social scientists have come to view psychology as a source of aimless forces that warp our rational judgment. In reality, rationality and purpose are the defining feature of all of life, distinguishing living organisms from lifeless matter. This is why economics and evolutionary biology are so compatible, both deal in costs, benefits, goals, and strategies.

Our minds and emotions are rational and goal-seeking. They weigh costs and benefits, and they make strategic moves. The field of psychology asks us to look deeper beneath rationality to study the information-processing algorithms that allow us to perceive depth, comprehend sentences, find our way home, haggle with a merchant, support a leader, cooperate with a partner, and other common challenges. Good theories for these behaviors will take the form of proposing efficiently designed algorithms that could explain them.

In contrast, theories that propose only aimless urges, clumsy errors, and various irrationalities are inconsistent with the basic nature of a living mind. Errors are inevitable in life, but it is the good decisions and impressive feats that pose the most challenging puzzles for psychology.

For a closing example, consider the algorithms in a bird’s brain that control flight [61]. Birds fly with more skill and precision than any manmade machine due to the exquisite engineering of their flight controllers, compliments of natural selection. Although we could catalogue the bird’s errors, this would hardly tell us how it flies. A good theory of bird flight should propose algorithms that are at least as capable as the flight controllers designed by professional software engineers, since birds outperform these machines. If bird flight and countless other animal behaviors call for theories about high-performance algorithms, then perhaps we should expect no less from theories about human behavior. After all, the same process of natural selection behind the hummingbird’s acrobatics is also the artful programmer behind an angry scowl, an amorous gaze, and tears of sorrow.

Author Contributions: L.I.R. and P.D. contributed equally to all parts of the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References


60. McCullough, M.E.; Reed, L.I. What the face communicates: Clearing the conceptual ground. *Curr. Opin. Psychol.* 2016, 7, 110–114. [CrossRef]


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).