


# Strategically Communicating Minds

Hugo Mercier<sup>1</sup>, Guillaume Dezeache<sup>2</sup>, and  
Thom Scott-Phillips<sup>3,4</sup>

<sup>1</sup>Cognitive Science Institute - Marc Jeannerod UMR 5304, CNRS; <sup>2</sup>Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University; <sup>3</sup>Department of Cognitive Science, Central European University; and <sup>4</sup>Department of Anthropology, Durham University

Current Directions in Psychological  
Science  
2017, Vol. 26(5) 411–416  
© The Author(s) 2017  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0963721417703155  
www.psychologicalscience.org/CDPS  


## Abstract

Several recent theories postulate communicative functions for cognitive mechanisms previously thought to have individualistic functions—in particular, reasoning and metacognition. These theories join older theories suggesting that many of our behaviors have communicative functions, for instance to communicate emotions or to influence how people perceive us. Using the framework of the evolution of communication, we offer a series of questions to test these hypotheses. The first question is whether the mechanism enables effective communication. The second question takes into account the different strategic incentives between agents who send signals and those who receive them, asking whether receivers can discriminate beneficial from harmful signals. However, serving a function well is not sufficient evidence that a mechanism evolved to this end in particular. Accordingly, the third question bears on whether the mechanism serves other purported functions well and the fourth on whether some of its features can be explained as specifically serving a communicative function. An overview of the literature suggests that these questions have been experimentally addressed for some cognitive mechanisms (reasoning in particular) but not others. This framework thus opens up avenues for further research that will enable researchers to better test hypotheses regarding the communicative functions of cognitive mechanisms.

## Keywords

communication, evolution, reasoning, metacognition, emotions, self-presentation

A recent trend in cognitive science is to attempt to better understand some cognitive mechanisms by postulating that they have a communicative function. Two recent examples are Mercier and Sperber's (2011) argumentative theory of reasoning and Shea et al.'s (2014) dual-process theory of metacognition. Mercier and Sperber (2011) suggest that the main function of human reasoning is communicative and, more specifically, argumentative: Reasoning would have evolved so that senders (i.e., individuals who send information) can provide arguments to convince receivers (i.e., individuals who receive information), whereas receivers can evaluate these arguments to decide whether they should change their minds. This exchange of arguments would help senders and receivers communicate when trust is not sufficient to guarantee that messages can be safely accepted. Shea et al. (2014; see also Yaniv & Foster, 1995) suggest that metacognition should be divided in two different sets of processes, an implicit one whose function is to regulate behavior and cognition and an explicit one whose function is to communicate to others

one's degree of confidence. Thanks to this ability to communicate confidence, senders and receivers should be better able to reach an agreement. These theories can be contrasted with more standard theories holding that reasoning and metacognition serve the individual function of regulating one's personal cognition, in particular by correcting mistaken intuitions (e.g., Kahneman, 2011; Stanovich, 2004).

The recent trend can be related to older theorizing that postulated communicative functions for a range of other behaviors (and thus, implicitly, for the cognitive mechanisms guiding these behaviors). Two salient examples are emotional expressions and self-presentation. Several researchers have postulated that many, if not all, emotional expressions—facial expressions in particular—serve communicative functions. For

---

## Corresponding Author:

Hugo Mercier, Cognitive Science Institute - Marc Jeannerod, CNRS UMR 5304, 67 Boulevard Pinel, 69675 Bron, France  
E-mail: hugo.mercier@gmail.com

instance, the function of the expression of anger would be to communicate that one is about to act aggressively toward the receiver (Fridlund, 1997). Regarding self-presentation, a long tradition in sociology and social psychology suggests that many of our behaviors aim at influencing how others perceive us—as a way of either forcing this perception onto them or communicating it to them (e.g., Leary, 1995).

Following Sperber's (2001) lead, the goal of this article is to use the framework of the evolution of communication to specify what these ascriptions of communicative functions mean and to highlight ways to test them. We focus here on these four examples, but the framework can easily be extended to related proposals (e.g., Baumeister & Masicampo, 2010; Johnson & Fowler, 2011).

## An Evolutionary Perspective on Communication

In an evolutionary perspective, an influential way to define communication is in terms of adaptive function. Communication occurs when an action (a signal) produced by an individual organism causes a change (a reaction) in another organism, where both the signal and the reaction have evolved to fulfill these purposes (Maynard Smith & Harper, 2003; Scott-Phillips, 2008). If the action has evolved to fulfill these purposes but the reaction has not, then the interaction is coercive; and if the reaction has evolved for these purposes but the action has not, then the interaction is a cue.

This adaptationist approach suggests a series of tests to evaluate whether a given cognitive mechanism serves a communicative function (see Maynard Smith & Harper, 2003). The first test is straightforward: Does the cognitive mechanism fulfill this communicative function well? This means that this mechanism should allow senders to send messages that are understood by receivers, thus allowing them to communicate effectively.

However, senders and receivers often have different strategic incentives. As a result, senders might have incentives to send signals that are dishonest or in some other way costly to receivers. So for communication to be evolutionarily stable, widespread dishonesty must be either prevented or defended against. Evolutionary biologists have identified several possible means by which this can occur (for a review, see Maynard Smith & Harper, 2003). Humans frequently interact with unrelated individuals who have different incentives from their own and possess cognitive mechanisms that evolved to evaluate communicated information (Sperber et al., 2010). Thus, a second test of whether a particular cognitive mechanism has a communicative function is whether receivers are able to evaluate communicated

information in a broadly accurate way, rejecting the most harmful signals while accepting the most beneficial ones.

Having said that, showing that a cognitive mechanism fulfills a particular function well is necessary but not sufficient. A mechanism can do something well without being adapted especially for it—color vision did not evolve for traffic signals or hands for holding pens. To test whether the actual function of a given cognitive mechanism is more general than the communicative function, one can see how this mechanism performs noncommunicative functions that have also been ascribed to it. If the mechanism does not fulfill these other functions well, this argues in favor of the communicative function. Relatedly, and as a final test, some traits of the cognitive mechanisms should serve better as a communicative function than other functions.

## Reasoning

**Does it enable effective communication?** When people can discuss things with each other, their performance increases on a wide variety of tasks—logical, mathematical, and inductive problems; schoolwork; medical diagnoses; economic predictions; and so forth (Mercier, 2016a; Mercier & Sperber, 2011). More specifically, group members with sound insights into the problem usually manage to convince others to accept these insights. That the exchange of arguments is critical to this process has been demonstrated by analyses of the transcripts of discussions (e.g., Moshman & Geil, 1998) and by eliminating other potential explanations (for instance, that the expression of confidence might be sufficient to account for the observed spread of the correct answers; Trouche, Sander, & Mercier, 2014).

**Can receivers discriminate beneficial from harmful signals?** Reasoning helps discriminate between beneficial and harmful signals by discriminating between good and poor arguments, which should correlate with overall signal quality in most cases. At least when they are motivated, people discriminate between good and poor arguments, being influenced by the former while they reject the latter (Hahn & Oaksford, 2007; Petty & Wegener, 1998). In particular, there is no experimental evidence that people are easily taken in by sophistry or fallacious arguments.

**Does it serve other purported functions less well?** Reasoning largely fails to fulfill the individual function usually ascribed to it: A majority of participants fail to correct their mistaken intuitions even when doing so requires only elementary logical or mathematical reasoning well within their abilities (Frederick, 2005; Wason, 1966).

**Can its features be explained as specifically serving a communicative function?** Reasoning has several features that seem particularly well suited for a communicative function. For instance, reasoning has been shown to have a strong confirmation bias (more accurately called a *myside bias*; see Mercier, 2016b) so that when people generate reasons, they overwhelmingly find reasons that support their prior beliefs (Mercier, 2016b). This is largely why reasoning fails to correct solitary reasoners' intuitions. By contrast, the *myside bias* makes sense if reasoning has an argumentative function, as it helps people find arguments to defend their opinions.

### **Metacognition**

**Does it enable effective communication?** When people face decisions together for which finding reasons is difficult (e.g., because the tasks are perceptual), they effectively communicate their degree of confidence, allowing the group to agree on the answer of the most confident individual, who, in this type of task, tends to be right (Bahrami et al., 2010). Analyses of transcripts have revealed the richness of the linguistic means through which people manage to communicate their degrees of confidence (Fusaroli et al., 2012).

**Can receivers discriminate beneficial from harmful signals?** Here, the honesty of a signal depends on whether the expressed degree of confidence (as interpreted by the receiver) correlates with the likelihood that the message is sound. Some researchers have claimed that receivers do not discriminate well between senders who consistently express confidence, whether they are right or wrong, and senders who are better calibrated (Anderson, Brion, Moore, & Kennedy, 2012; Kennedy, Anderson, & Moore, 2013). However, a literature review and further experiments have shown that, on the whole, misplaced expressions of confidence tend to degrade a sender's reputation, with the consequence that the expression of overconfidence is not a sound strategy in repeated interactions (Vullioud, Clément, Scott-Phillips, & Mercier, 2017).

**Does it serve other purported functions less well?** The main other function suggested for metacognition is to regulate individual thought and behavior. If implicit metacognition fulfills this function well (Shea et al., 2014), explicit metacognition suffers from several issues in this respect, being affected by some consistent biases. For instance, people consistently overestimate the degree of precision with which they can make estimates (Yaniv & Foster, 1995).

**Can its features be explained as specifically serving a communicative function?** Some features of explicit metacognition might be well suited for communicative

functions. For instance, the tendency to offer estimates that are more precise than warranted has been cast as a way of increasing the relevance of one's statements rather than as exerting an undue influence on others (Yaniv & Foster, 1995). Accordingly, overprecision does not seem to impair communication, as receivers tend to adjust for senders' overprecision (see Vullioud et al., 2017).

### **Emotional signals**

**Do they enable effective communication?** Subtle changes in facial musculature can be detected quickly and their emotional significance accurately evaluated by recipients over time (Jack, Garrod, & Schyns, 2014). If the extent to which those emotional signals are universal is debated, locally specific facial configuration and dynamics enable effective communication of emotions (Jack, Garrod, Yu, Caldara, & Schyns, 2012).

**Can receivers discriminate beneficial from harmful signals?** It has long been claimed that receivers react automatically to emotional signals (Hatfield, Cacioppo, & Rapson, 1994). If this were so, senders could abuse emotional signals to their advantage by repeatedly sending signals that benefit them but not receivers. In fact, reactions to emotional signals are heavily modulated by characteristics of the source and of the situation (for a review, see Dezechache, Mercier, & Scott-Phillips, 2013). For instance, signals of pain are likely to elicit an empathetic response if the receiver expects to collaborate with the sender, but no response or a counterempathetic response if the receiver expects to compete with the sender (Lanzetta & Englis, 1989).

**Do they serve other purported functions less well?** According to the two-stage models of the evolution of facial displays (Shariff & Tracy, 2011), facial expressions of emotions first evolved to serve intrapersonal sensory regulatory functions before being selected for their communicative function. Some effects of emotional expressions are compatible with an individual function. For instance, the widening of the eyes in fear increases sensory acquisition (Susskind et al., 2008). However, the modifications of sensory exposure in producers of facial display are limited, and they cannot account for the evolution of other prototypical facial movements (such as the activity of the zygomaticus major in the smile).

**Can their features be explained as specifically serving a communicative function?** Following Ekman (2007), a tradition in psychology has viewed emotional expressions as automatic readouts of inner states. However, and as argued by Fridlund, it is unlikely that humans would have evolved traits that can be detrimental to them, for instance by automatically revealing fear when under

stress. In a number of studies, Fridlund (1994) also has shown that the intensity of emotional facial displays (notably smiles) is modulated by the presence of an audience and the sociality of the context.

## **Self-presentation**

**Does it enable effective communication?** A tremendous amount of research has been dedicated to showing how self-presentation affects senders' behaviors, but relatively little work has been dedicated to receivers' reactions to these behaviors (see Leary, 1995). For instance, social psychologists have extensively studied the influence that dissonance reduction has on individuals' behaviors but have paid little attention to how individuals who reduce dissonance (or who fail to do so) are perceived by others (Cooper, 2007). As a result, it is difficult to tell if self-presentation enables either effective communication or if it effectively serves coercive goals for senders (i.e., getting receivers to form a favorable impression of them, whether it is warranted or not).

**Can receivers discriminate beneficial from harmful signals?** As mentioned above, little work has focused directly on the question of how people react to self-presentation. However, some work in a related area bears on this question. One of the goals that self-presentation might serve would be to make one's lies harder to detect by restraining behaviors that are perceived as cues to lying, such as fidgeting, or by encouraging behaviors that are perceived as cues to truth telling, such as looking the receiver in the eyes (The Global Deception Research Team, 2006). If this were the case, then receivers might be made to accept lies, raising the question of why they would keep attending to these signals. Recent work in lie detection suggests, however, that people in fact pay little attention to such unreliable cues, focusing instead on the content of the message and on whether the situation makes a lie more or less likely (Hartwig & Bond, 2011). Thus, although we cannot exclude that self-presentation can consistently deceive receivers, there is, to the best of our knowledge, no evidence that this is the case either.

**Does it serve other purported functions less well?** Some mechanisms of self-presentation, such as dissonance reduction, have been postulated to serve nonsocial, regulatory functions (see Cooper, 2007). However, self-presentation can lead senders to the adoption of misguided beliefs and costly behaviors. For instance, dissonance reduction can lead participants to believe that an excruciatingly boring task was not so bad after all (Festinger & Carlsmith, 1959), and it can lead them to engage in personally costly but socially desirable behaviors, such as having an instrument installed in their car to measure its speed (for a review, see Stone & Fernandez, 2008).

**Can its features be explained as specifically serving a communicative function?** The type of work cited above suggests that, at least in the case of dissonance reduction, features of self-presentation are best explained by a social than by an individual function. However, the relative lack of work on receivers' reactions to self-presentation makes it difficult to tell whether the features of self-presentation fit best with a coercive or a communicative function.

## **Conclusion**

We have offered a series of questions, informed by the perspective of evolutionary adaptationism, that can help researchers test their hypotheses about the communicative functions of cognitive mechanisms. In the case of reasoning, the questions have all been experimentally addressed. However, for the other mechanisms, some questions have been the topic of little inquiry. In particular, work on self-presentation could be usefully extended so that these basic questions can be answered. We thus hope that the proposed framework will prompt research in these directions. Another avenue for further research is the influence of strategic incentives on senders. Senders likely evolved to modulate their signals as a function of the receivers' vigilance—in particular, to avoid the potential reputational costs of sending signals perceived as misleading. Thinking of senders' behavior in this light might help explain otherwise puzzling behaviors, such as why people take sunk costs into account (Arkes & Ayton, 1999) or why they exert caution in sending requests (Bohns, 2016).

## **Recommended Reading**

- Dezecache, G., Mercier, H., & Scott-Phillips, T. C. (2013). (See References). A review of work showing that people evaluate emotional communication.
- Leary, M. R. (1995). (See References). A classic introduction to the social psychology of self-presentation.
- Mercier, H., & Sperber, D. (2011). (See References). Argues that the main functions of human reason are communicative.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). (See References). Argues that some metacognitive processes serve a communicative function.

## **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.

## **Funding**

H. Mercier is supported by Agence Nationale de la Recherche (ANR) Grant ANR-16-TERC-0001-01. G. Dezecache is supported by the European Research Council (ERC) FRONTSEM project and ANR Grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL. T. Scott-Phillips received funding from the

ERC under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement No. 609819 (Somics project).

## References

- Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology, 103*, 718–735.
- Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin, 125*, 591–600.
- Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science, 329*, 1081–1085.
- Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review, 117*, 945–971.
- Bohns, V. K. (2016). (Mis)Understanding our influence over others: A review of the underestimation-of-compliance effect. *Current Directions in Psychological Science, 25*, 119–123.
- Cooper, J. (2007). *Cognitive dissonance: Fifty years of a classic theory*. London, UK: Sage.
- Dezecache, G., Mercier, H., & Scott-Phillips, T. C. (2013). An evolutionary approach to emotional communication. *Journal of Pragmatics, 59*, 221–233.
- Ekman, P. (2007). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. New York, NY: Macmillan.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*, 203–210.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42.
- Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. New York, NY: Academic Press.
- Fridlund, A. J. (1997). The new ethology of human facial expressions. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 103–129). Cambridge, UK: Cambridge University Press.
- Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science, 23*, 931–939.
- The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*, 60–74.
- Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114*, 704–732.
- Hartwig, M., & Bond, C. H. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*, 643–659.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York, NY: Cambridge University Press.
- Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology, 24*, 187–192.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences, USA, 109*, 7241–7244.
- Johnson, D. D., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature, 477*, 317–320.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Kennedy, J. A., Anderson, C., & Moore, D. A. (2013). When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior and Human Decision Processes, 122*, 266–279.
- Lanzetta, J. T., & Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *Journal of Personality and Social Psychology, 56*, 543–554.
- Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior*. Boulder, CO: Westview Press.
- Maynard Smith, J., & Harper, D. (2003). *Animal signals*. Oxford, UK: Oxford University Press.
- Mercier, H. (2016a). The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences, 20*, 689–700.
- Mercier, H. (2016b). Confirmation (or myside) bias. In R. Pohl (Ed.), *Cognitive illusions* (2nd ed., pp. 99–114). London, UK: Psychology Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral & Brain Sciences, 34*, 57–74.
- Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking & Reasoning, 4*, 231–248.
- Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 323–390). Boston, MA: McGraw-Hill.
- Scott-Phillips, T. C. (2008). Defining biological communication. *Journal of Evolutionary Biology, 21*, 387–395.
- Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science, 20*, 395–399.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Sciences, 18*, 186–193.
- Sperber, D. (2001). An evolutionary perspective on testimony and argumentation. *Philosophical Topics, 29*, 401–413.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*, 359–393.
- Stanovich, K. E. (2004). *The robot's rebellion*. Chicago, IL: Chicago University Press.

- Stone, J., & Fernandez, N. C. (2008). To practice what we preach: The use of hypocrisy and cognitive dissonance to motivate behavior change. *Social & Personality Psychology Compass*, *2*, 1024–1051.
- Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience*, *11*, 843–850.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, *143*, 1958–1971.
- Vullioud, C., Clément, F., Scott-Phillips, T., & Mercier, H. (2017). Confidence as an expression of commitment: Why misplaced expressions of confidence backfire. *Evolution & Human Behavior*, *38*, 9–17.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology: I* (pp. 106–137). Harmondsworth, UK: Penguin.
- Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General*, *124*, 424–432.