Experimental work on the norms of assertion

John Turri

Communication is essential to human society, and assertion is central to communication. This article reviews evidence from life science, cognitive science, and philosophy relevant to understanding how our social practice of assertion is structured and sustained. The principal conclusion supported by this body of evidence is that knowledge is a central norm of assertion—that is, according to the rules of the practice, assertions should express knowledge.

KEYWORDS
assertion, communication, knowledge, norms

Communication is endemic to life. Even the simplest organisms, including bacteria, communicate (Keller & Surette, 2006). In the thriving interdisciplinary study of animal communication, communication is understood as sending or receiving a signal whose function is to provide information to another organism for use in making decisions (Bradbury & Vehrencamp, 2011). Decades of research in this field has made it clear that communication is an adaptive behavioral trait shaped by natural selection, confirming speculation dating to the origin of modern evolutionary biology (Darwin, 1872). Communication systems evolved because they benefit sender and receiver (Maynard Smith & Harper, 2004). Adaptive and stable communication systems make it worth the sender’s effort to send signals and worth the receiver’s effort to monitor and parse signals.

Sometimes, the signaler’s interests align with the receiver’s, but other times they do not. When the sender and receiver benefit similarly from how receivers respond to signals, receivers can count on honest signals. For example, prey and predator have a mutual interest in sharing certain information. Prey will often stare intently at nearby predators and follow their movements. This signals to the predator that it has lost the element of surprise, which is often enough to call off the hunt (Hasson, 1991). Prey obviously benefit from not being hunted, and predators benefit from avoiding hunts with a very low probability of success. By contrast, when the preferences of sender and receiver diverge, senders can also benefit from sending dishonest signals. For example, in some species, over two-thirds of predator alarm calls are false and are often merely intended to scare conspecifics away from a preferred food source or to gain mating opportunities (Haftorn, 2000; Wheeler & Hammerschmidt, 2013).

What keeps senders from routinely sending false or misleading signals when it could benefit them? The simple answer is that, in the long run, receivers adapt: They evolve to better detect dishonesty in a signal, ignore certain signals, and attend to more honest signals. Stable and enduring communication systems thus include features that promote honest signaling. Researchers have identified several mechanisms that promote honest signaling. Two are most relevant to the present discussion.
The first mechanism is attending preferentially to **performance signals**, which only some signalers can produce (Hurd & Enquist, 2005). Some performance signals are indexed to physical characteristics such as body size. For instance, smaller toads cannot croak as deeply as larger toads, so lower-frequency croaks are restricted to larger specimens (Davies & Halliday, 1978). Signalers cannot send dishonest signals that their body size prevents them from sending. Other performance signals are "information constrained" (Hurd & Enquist, 2005). For example, consider pursuit-deterrent signaling by potential prey. An antelope stares at a lioness in the brush and follows her movements, thereby signaling to the lioness "both its alerted state and the futility of continuing the hunt. This signal can be performed only by a signaler who knows the location of the hidden predator" (Hurd & Enquist, 2005, p. 1160). To take another example, neighboring sparrows usually share at least two songs in their repertoire. Sparrows view established neighbors as less threatening than strangers (Temeles, 1994). A neighbor’s song will be heard frequently, even when the neighbor is respecting the bird’s territory. A sparrow typically responds to a neighbor’s song by singing a different song it shares with the neighbor ("repertoire matching"), which expresses tolerance. By contrast, a sparrow typically responds to a stranger’s song by imitating it ("song matching"), which expresses aggressive intent (Vehrencamp, 2001). Because repertoire matching "requires knowledge of the singer’s repertoire," or having "committed that bird’s repertoire to memory," it is an informationally constrained signal of neighborhood (Beecher, Campbell, Burt, Hill, & Nordby, 2000, pp. 22 and 25).

The second mechanism is **social policing**, which involves testing for honesty and retaliating for dishonesty. This is important for conventional signals, whose form is arbitrarily associated with their significance and, thus, are not constrained as performance signals are. For example, in some sparrow species, the amount of dark plumage on the head and throat signals social dominance (Rohwer, 1977). This visible marker of dominance is a "badge of status" that allows conspecifics to settle disputes over resources without resorting to potentially harmful fighting (Dawkins & Krebs, 1978). The benefit of a large badge is deference from individuals with small badges. But if a larger badge confers such advantages, then why do subordinates not simply molt into plumage that resembles a higher rank? Because a system of "social control" actively prevents it (Moller, 1987). Individuals with large badges frequently challenge one another, ensuring that pretenders will be exposed. Aggressive punishment of dishonest signalers has also been observed in lizards and wasps (Thompson & Moore, 1991; Tibbetts & Dale, 2007; Tibbetts & Izzo, 2010). But punishment does not always take the form of outright aggression. Sometimes the cost is diminished reputation and distrust from other group members, known as "skeptical responding" (Cheney & Seyfarth, 1988; Gouzoules, Gouzoules, & Miller, 1996). The sophistication of skeptical responding in some monkey species is truly remarkable. For example, a vervet monkey who gives false leopard alarm calls will eventually be ignored on subsequent leopard alarm calls, but this skepticism is not transferred to that monkey’s eagle alarm calls.

Behavioral ecologists describe "receiver retaliation" as a "behavioral rule" that disincentivizes dishonest conventional signaling. Indeed, they consider this rule to be “the key factor” making conventional communication systems evolutionarily stable (Bradbury & Vehrencamp, 2011, p. 411). Absent retaliatory costs, dishonesty would proliferate and eventually conventional signals would be ignored. Senders would not benefit from producing them and so would eventually stop doing so. Conventional communication would be abridged severely, if not abrogated entirely.

In humans, assertion is a principal means of communicating information. Just like animal communication more generally, human communication is an adaptive trait that benefits sender and receiver. Human communication is subject to the same general evolutionary pressures as any other animal communication system. Accordingly, we should expect human communication to exhibit important similarities to other animal communication systems. Spoken human language is a paradigm example of conventional communication. An assertion’s content is arbitrarily associated with its form and human speech is very cheap to produce. Competent speakers can generally assert what they want, when they want, while incurring negligible production costs. Thus the question arises: What prevents humans from dishonestly asserting to their own advantage? Of course, humans often lie and mislead (Vrij, 2008). So more specifically the question is this: What prevents humans from dishonestly asserting enough to destabilize the practice?

Recent work in cognitive science and philosophy supports a specific answer to that question: The human practice of assertion is at least partially sustained by a socially policed information constraint, namely, knowledge (Turri, 2016b).
Humans expect assertions to express knowledge and react negatively to assertions failing to meet that standard. In the literature, this sort of view is often described as proposing a “knowledge norm” or “knowledge rule” for assertion (Benton, 2014).

The conversational give-and-take surrounding assertion is the most obvious way humans police the practice (Turri, 2014). We prompt assertions with formulations that explicitly presuppose knowledge, just as we test for honesty and accuracy by challenging assertors with questions that explicitly refer to or imply knowledge. For instance, one way of prompting someone to make an assertion is to ask, “What time is it?” But an equally effective, and practically interchangeable, prompt is to ask, “Do you know what time it is?” (Turri, 2010: 458). Competent speakers respond to the two questions similarly. And once you have made an assertion, even if the content of your assertion has nothing to do with you or what you know, it is normally appropriate to ask you, “How do you know that?” (Slote, 1979; Unger, 1975, pp. 263–264; Williamson, 2000). When you are asked a question, even if the question has nothing to do with you or what you know, it is normally acceptable to abstain by saying, “Sorry, I don’t know” (Reynolds, 2002: 140). Relatedly, as abstentions, “I don’t know,” “I can’t tell,” and “I can’t say” are practically interchangeable (Turri, 2011,p. 38). The parable of Cain and Abel contains perhaps the most famous abstention in literary history. In one translation of the story, when asked, “Where is your brother Abel,” Cain answers, “I know not: Am I my brother’s keeper?” (Koenig & Harris, 2005, 1266 ff.). And when asked why they are good at answering questions, children will say, “Because I know.”

In addition to observational evidence, developmental findings, and precedents from animal communication studies, it is possible to experimentally test whether a knowledge rule is central to our social practice of assertion. We should expect skilled practitioners to reliably identify what should and should not be done according to the rules of the practice. In other words, the intuitive judgments of skilled practitioners are a source of evidence about what the practice’s rules are. That is why philosophers and cognitive scientists contributing to research on assertion agree that theories in this area “must face the linguistic data” (Douven, 2006). For instance, if knowledge is a norm of assertion, then judgments about what someone should assert (“assertability judgments”) will be sensitive to knowledge judgments and, furthermore, to judgments about the intuitive requirements of knowledge, such as truth.

One series of studies investigated whether the norm of assertion is, at the very least, “factive” or truth-entailing (Turri, 2013; Turri & Blouw, 2015). A factive norm implies assertions should express truths. Participants considered a simple story about Maria. Maria is a watch collector who owns so many watches that she cannot keep track of them all by memory alone, so she maintains a detailed inventory of them. She knows that the inventory, although not perfect, is extremely accurate. One day someone asks Maria whether she has a 1990 Rolex Submariner in her collection. She consults the inventory and it says that she does have one. At the end of the story, one group of people
was told that the inventory was right, whereas another group was told that the inventory was wrong—that was the only difference between conditions. Everyone then answered the same question: Should Maria say that she has a 1990 Rolex Submariner in her collection? When the assertion would be true, virtually everyone said that Maria should make the assertion. But when the assertion would be false, the vast majority said that she should not make the assertion. When asked to explain their evaluation, most people said that the statement’s truth value was more important than Maria’s evidence.

This same basic pattern persisted when people were questioned in different ways and across other differences that often can influence evaluative judgments and social cognition (Turri, 2013). For example, the pattern was robust against participant age and gender. It persisted whether the stakes were low (a “neighbor asking out of idle curiosity”) or high (a “federal prosecutor asking in the course of an official investigation”). It persisted when the stimuli were systematically switched so that participants evaluated negative assertions (e.g., "I don’t have one" as opposed to "I do have one"). It persisted when testing different narrative contexts (Turri, 2017a). The same basic pattern also persisted when people performed a much more open-ended task of identifying what Maria should say, rather than merely rating their agreement with the statement that Maria should make the assertion (Turri, 2013). When the assertion would be true, the vast majority of people answered that Maria should assert that she owns the watch. But when the assertion would be false, very few people answered that way. Instead, the most common response was that Maria should assert that she “probably” owns one, which, on the most natural interpretation of the case, is true because of Maria’s evidence (see also Turri, 2015b).

Some researchers have suggested that assertability judgments are sensitive to truth value only because truth value affects whether an assertion would be immoral, irrational, impolite, or some other negative status (Kvanvig, 2011; Pagin, 2015). The underlying concern is that patterns favoring “true assertions actually track” considerations that are not “proper to assertion” (Pagin, 2015, p. 22). This concern has been directly tested. Participants considered simple scenarios and recorded judgments about assertability, truth value, morality, rationality, etiquette, and legality. Controlling for the influence of the other qualities, evaluations of truth value were by far the strongest predictor of assertability attributions (Turri, 2017a).

In another set of studies, participants evaluated agents in many different situations, with different types of evidence and with different amounts at stake (Turri & Buckwalter, 2017). For example, one group evaluated Christina, a barista in charge of updating the coffee shop menu each day. To some customers with severe nut allergies, it matters whether the coffee contains pine nuts. While working on today’s menu, Christina notices a persistent pattern in the supplier’s shipments, which strongly suggests that the latest shipment of coffee does not contain trace amounts of pine nuts. Should Christina write on today’s menu that the coffee does not contain traces of pine nuts, effectively making a written assertion to customers? In addition to judging whether Christina should make a written assertion, participants also judged whether the relevant proposition is true, whether she believes the proposition, whether she has good evidence for the proposition, and how important it is whether the proposition is true. Of all these judgments, knowledge judgments most strongly predicted assertability attributions. A related study found that knowledge judgments mediate the effect of truth on people’s evaluation of how others should behave (Turri, 2015b, experiment 4). This suggests that truth influences such evaluations because it influences knowledge judgments—that is, whether a proposition is true affects whether people think that an agent knows that it is true, which in turn guides behavioral evaluations.

A third set of studies tested the knowledge account directly and in the simplest way possible: by intervening on knowledge (Turri, 2015c). This study manipulated the presence or absence of knowledge by including it as an independent variable in the experimental design. This is important because if assertions should express knowledge, then manipulating the presence or absence of knowledge should significantly affect people’s assertability judgments. Across a variety of scenarios, varying whether the agent knows the relevant proposition, while holding all else equal, had an extremely large effect on judgments of assertability. For example, consider Mallory, who manages the local farmer’s market. One of her employees is interested in improving the health of his diet. The employee asks Mallory whether avocados have vitamin K. Should Mallory say that avocados have vitamin K? In one version of the story,
Mallory knows that avocados have vitamin K. In the other version, she does not know. Nearly everyone who read the first story judged that Mallory should make the assertion, but nearly no one who read the second story did.

A fourth study investigated whether knowledge was a stronger indication of assertability than either believing or being certain of a true proposition (Turri, Friedman, & Keefner, 2017). Each group read the same basic story, with one small difference. The first group was told that the agent believes a true proposition; the second group was told that the agent is certain of that same true proposition; the third group was told that the agent knows the true proposition. People then rated whether the agent should perform a variety of actions, including asserting a proposition. To illustrate, consider the following example. The water at Metro Beach was recently tested and declared unsafe for swimming. However, the health department botched the test, and, as a matter of fact, the water is perfectly safe for swimming. It is a hot summer day and Alicia decides to go to Metro Beach. She examines the water and now she thinks (is certain/knows) that the water is safe for swimming. Should Alicia tell other people at the beach that the water is safe for swimming? When people were told that Alicia knows, the central tendency was to agree that she should make the assertion. By contrast, when people were told that Alicia thinks or is certain, the central tendency was to disagree. It is worth emphasizing that this experiment held truth constant across the three conditions. That is, the comparison was not simply between knowledge, belief, and certainty. Rather, it was between knowledge, true belief, and true certainty (i.e., being certain of a proposition that is true). Whatever difference remains is attributable to knowledge specifically. And it was knowledge specifically that led people to judge that the assertion should be made. In a related study, people were much more willing to allow attributions of assertability and certainty to come apart than to allow attributions of assertability and knowledge to come apart (Turri, 2016f).

The link between knowledge and assertability has also been tested using peculiar scenarios that critics claim intuitively break that link. For example, critics claim that knowledge and assertability intuitively come apart in “fake barn” cases (Coffman, 2014; Lackey, 2007; Smithies, 2012). Suppose Sarah looks out her car window and sees a roadside barn as she drives along. Everything about Sarah and the barn is normal. But Sarah does not realize that the area she is driving through is being used as a movie set and the set designers have constructed many fake barn façades that look just like real barns. Sarah is looking at the one real barn among all the nearby fakes. Critics claim that Sarah clearly does not know that it is a barn, even though she should, if asked, say that it is a barn. Researchers recently tested whether this captures how competent speakers actually judge such cases (Turri, 2016a; for related research on such cases, see Colaco, Buckwalter, Stich, & Machery, 2014; Horvath & Wiegmann 2016; Turri, Buckwalter, & Blouw, 2014; Turri, in press c). In one study, people read one of two versions of a story about a woman who sees a barn while driving with her son. In one version, she is looking at the only “expensive barn” amidst many “cheap barns.” In the other version, she is looking at the only “real barn” amidst many “fake barn façades.” At the end of the story, her son points and asks if it is a barn. Whether the contrast was cheap barns or fakes, participants overwhelmingly judged that Sarah knows and should tell her son that it is a barn.

Critics have also claimed that “selfless assertions” (Lackey, 2007; Pritchard, 2014) provide intuitive examples where agents should assert propositions that they do not believe or, hence, know. Researchers recently tested how competent speakers actually judge such cases. Although people strongly judge that “selfless assertors” should assert the relevant propositions, people equally strongly judge that “selfless assertors” believe and know the propositions they should assert (Turri, 2015d). To illustrate, a widely discussed example of “selfless assertion” features Stella, a devoutly religious creationist who teaches science to fourth graders (Lackey, 2007, p. 599). Stella’s “deep faith” includes “a belief in the truth of creationism and, accordingly, the falsity of evolutionary theory.” Nevertheless, Stella “fully recognizes” the “overwhelming scientific evidence against creationism and in favor of evolutionary theory.” This leads Stella to tell her students, “Modern humans evolved from more ape-like ancestors called hominids.” Should Stella make this assertion? Does she believe that humans evolved? Does she know that humans evolved? Competent speakers judged that she should make the assertion. But they also judged that she believes and knows that humans evolved.

It is worth noting that although knowledge sets a standard for what should be asserted, people reliably distinguish between assertions that should not be made and blameworthy assertions (Turri, 2013; Turri & Blouw, 2015; see also...
Williamson, 2000: 256). From the fact that someone makes an assertion that he should not make, people do not automatically infer that he should be blamed for the assertion. This fits a pattern in commonsense morality and epistemology: People generally acknowledge the possibility of blameless transgressions. For instance, when judging moral obligations, people reliably distinguish between a person not keeping a promise he is obligated to keep, on the one hand, and being blameworthy for not keeping it, on the other (Buckwalter & Turri, 2015; Turri, 2016e; Turri, 2017b). People also distinguish between propositions that a person should believe, on the one hand, and propositions that a person cannot be properly blamed for believing, on the other (Turri, in press a).

Due to space constraints, the present article cannot provide detailed coverage of all relevant recent research on norms of assertion. Some of this research provides evidence that the point of assertion is to transmit knowledge (Turri, 2016b), investigates relationships among attributions of assertability, knowledge, and perception (Turri, 2016c), investigates relationships among judgments about knowledge, assertability, and instructional demonstration (Buckwalter & Turri, 2014), and studies the relationship between evaluations of assertions and other speech acts, such as assurance (Turri, 2015a) and explanation (Turri, 2015e).

In summary, convergent evidence from animal communication studies, developmental findings on human children, observations of patterns in everyday discourse, and experiments with human adults all strongly support the conclusion that knowledge is a central norm of the social practice of assertion. It is worth emphasizing that this conclusion is phrased in terms of a central norm, not the one and only norm. The evidence reviewed here demonstrates an important normative relationship between assertion and knowledge, but it does not necessitate a unique interpretation of that relationship. For instance, one interpretation of the norm is that it imposes an exceptionless perfect requirement, one that applies strictly to each and every assertion. But another interpretation is that the norm imposes a stringent imperfect requirement, or very strong preference, which is satisfied if most of one's assertions express knowledge. Existing findings do not strongly favor one of these interpretations over the other. Consistent with all the findings to date, the practice might also be sustained by other norms pertaining to assertion, such as standards of criticism, punishment, or reward (for a general theoretical framework integrating such standards into an overall model of normative practice, see Turri, in press b). Also consistent with the available evidence is that the practice is also partly sustained by more general norms or evaluative presuppositions enabling social life, such as trustworthiness or cooperativeness (e.g., Tomasello & Vaish, 2013), or even by nonnormative factors enabling communication, such as a properly functioning mirror neuron system (e.g., Iacoboni & Wilson, 2006; Corballis, 2010). Philosophers have speculated that the norm of assertion is belief, certainty, or evidence (for reviews, see Weiner, 2007; Goldberg, 2015), and this might suggest that norms pertaining to these other categories also have a place in the practice of assertion. With the potential exception of a norm regarding certainty (see Turri, 2016c), no published findings support these hypotheses, so they remain speculative. At the same time, no published work rules them out, so they also remain worth investigating.

In light of these possibilities, as empirically informed work on this topic expands, it is important for researchers to avoid certain tendencies. On one hand, it is important to avoid an “either/or” approach, a principal mark of which is to assume that results suggesting the existence of one norm are automatically evidence against the existence of another, or that results suggesting the importance of a nonnormative factor are automatically evidence against the importance of norms. On another hand, and relatedly, it is important to avoid a narrow focus on alleged “counterexamples” to proposed norms of assertion. In order to have a fighting chance of understanding something as complicated as the human practice of assertion, we must not allow isolated observations or, especially, alleged intuitions about highly contrived cases to blind us to central tendencies that are abundantly evident elsewhere.

In closing, one important limitation of the research reviewed here is that the developmental, observational, and experimental evidence is mostly limited to North American anglophone populations. Informal consultation with bilingual adults who speak English and either Spanish, German, Arabic, Korean, or Assyrian suggests that many of the same patterns persist across cultures, but further research is clearly needed before drawing any firm conclusions. If knowledge is a central norm of the human practice of assertion, as it is in other animal communication systems, then we should find a strong connection between knowledge and assertability across cultures. Of course, this is consistent
with finding some cultural differences, such as those related to policing norms in general (e.g., Hamilton, Blumenfeld, Akoh, & Miura, 1990). Alternatively, further research could reveal vast differences suggesting that there is no such thing as the basic human practice of assertion but, rather, a constellation of human information-sharing practices sustained by different implicit rules.

ACKNOWLEDGEMENTS
For helpful feedback and discussion, I thank Vanessa Lam, Edouard Machery, Angelo Turri, and Sarah Turri. Thanks also to an anonymous reviewer for this journal. This research was supported by the Social Sciences and Humanities Research Council of Canada, the Ontario Ministry of Economic Development and Innovation, and the Canada Research Chairs program.

WORKS CITED
Buckwalter, W., & Turri, J. (2015). Inability and obligation in moral judgment. PloS One, 10(8), e0136589 https://doi.org/10.1371/journal.pone.0136589.g004.


---

**John Turri** is Canada Research Chair in Philosophy and Cognitive Science at the University of Waterloo (Ontario). He directs the Philosophical Science Lab and recently authored *Knowledge and the Norm of Assertion: An Essay in Philosophical Science* (Open Book Publishers).

---

**How to cite this article:** Turri J. Experimental work on the norms of assertion. *Philosophy Compass*. 2017;12:e12425. https://doi.org/10.1111/phc3.12425