The Function of Assertion and Social Norms

Peter J. Graham

The Oxford Handbook of Assertion
Edited by Sanford Goldberg

Subject: Philosophy, Philosophy of Language Online Publication Date: Mar 2019
DOI: 10.1093/oxfordhb/9780190675233.013.30

Abstract and Keywords

A proper function of an entity is a beneficial effect that helps explain the persistence of the entity. Proper functions thereby arise through feedback mechanisms with beneficial effects as inputs and persistence as outputs. We continue to make assertions because they benefit speakers by benefiting speakers. Hearers benefit from true information. Speakers benefit by influencing hearer belief. If hearers do not benefit, they will not form beliefs in response to assertions. Speakers can then only maintain influence by providing true information, often enough. The function of assertion is then inducing true hearer belief. When interests conflict, however, some mechanism must ensure that speakers provide true information often enough, instead of deceiving, or providing information regardless of quality. In humans, a core mechanism stabilizing true assertion involves social norms for truth telling. We tell the truth partly because we prescribe and enforce telling the truth.

Keywords: assertion, proper function, social norms, strong reciprocity, animal signals, epistemic vigilance

Tools, biological organs, and innate and learned behavioral traits, among other entities and activities, have proper functions, functions *proper* to them. The proper function of a screwdriver—its *own* function, the function *proper* to it—is to drive in screws. The proper function of a heart is to pump blood. The proper function of a handshake is to greet another. You could use any of these things for some other reason or purpose, but then that would be just *your* purpose, not its *own* proper function.

Proper functions are not limited to ordinary tools, biological traits, and learned behaviors. This allows for the very real possibility that linguistic devices—words and moods, for example, and even speech act types—have proper functions. That's the main idea, after all, in Ruth Millikan's 1984 masterwork, *Language, Thought and Other Biological Categories* and many other works, including her 2005 book, *Language: A Biological Model*.

Taking a page out of Millikan's playbook, but without endorsing every step along the way, I shall argue that assertion—the speech act we usually, though not always, use the indicative mood to perform—has the proper function of inducing true hearer belief.[1]

Why this holds for assertion, I shall furthermore argue, involves, at least in part, the role of social norms for truth telling. Assertion has the function of inducing true belief partly because of social norms for truth telling.[2]

1. Proper Functions

Let's begin with functions in general. What are functions, especially proper functions?

Philosophers tend to think of functions either as what a designer intends (intentional design functions), or as deriving from a subset of the item's actual effects (causal role functions), or as deriving from a subset of its past effects that help explain its current existence or configuration (etiological functions). I shall assume, with Millikan and many others, that *proper* functions are *etiological* functions.

1.1. Etiological Functions and Functional Explanations

And so, roughly speaking, a proper function of a device (of a heart, for example) is the effect of past activity (pumping blood) of past tokens (or stages) of the device that explains later tokens (or stages) of the device. Proper functions thereby arise from history: past effects of the device enter into a feedback mechanism that explain the persistence of the device. A proper function of a device explains its persistence, its continued reproduction or replication, its *survival*.

Here's a working definition. X in a system S has doing F as a proper function iff:

1. X does F in S.
2. F benefits S.
3. X exists (persists) in S because F benefits S.

Proper functions are those effects that benefit the system and that explain why, via a feedback mechanism, the persistence of the item (Graham 2014; McGlaughlin 2001).

A proper function attribution is then shorthand for a functional, historical explanation for the persistence of the item. Screwdrivers exist now *because* they were reproduced on the model of past screwdrivers that drove in screws. Mating displays exist now *because* past mating displays led the displaying individuals to reproduce, and thereby to produce offspring that would perform the same mating display in turn. When there are functions, past beneficial effects explain persistence.

A complete functional explanation won't leave us in the dark as to why the function helps stabilize persistence. We reproduce screwdrivers to drive in screws because we want them to. If we found they didn't, or if we didn't care to drive in screws anymore, we'd quit

making screwdrivers. Or maybe we'd make them for some other reason, and then their function (their *stabilizing* function) would eventually change. The male behaves a certain way to attract the female. If she responds, that stabilizes the behavior. But if she doesn't, the male gives up, tries something different, or goes out of business altogether.

1.2. Feedback Mechanisms

Biological reproduction is one of the most discussed feedback mechanisms generating functions. Having babies is one way to make copies of your current traits. If your traits kept you alive and helped explain how you managed to make babies, then the trait tokens in your offspring are explained, in part, by past effects of previous tokens of those trait types.

Learning is another well-known feedback mechanism. You try one behavior, get feedback, try it again, modify what you've done, or try something else instead. Eventually you keep doing the same thing because it works. You've learned how to do something. I tie my shoes a certain way—a behavioral trait of mine—because of valuable feedback I received as a child and continuing feedback I receive every day. That's why the particular way I tie my shoes has the function of tying my shoes. In this case I'm aware of what I'm doing and why I'm doing it. But in other cases, as Millikan emphasizes, we learn to do something even when we're not aware that we are doing it or why.

Social learning—learning from others—is another way behavior gets reproduced for its effects. It worked for them, and now that you've copied their behavior, it works for you too. Another mechanism is manufacturing. We keep making screwdrivers on the model of past screwdrivers because of their effectiveness at driving in screws. Maybe manufacturing is a form of learning. We've learned what works; that's why we keep making them.

Self-maintenance is another still (McGlauglin 2001). Complex systems—living organisms, beehives, and social groups—persist through an interesting form of part-whole causation. The parts help cause the whole, and the whole helps cause the parts. These complex systems then maintain themselves, as it were. (Echoes of Kant from the *Critique of Judgement*.) When your heart pumps blood, it keeps you alive. By keeping you alive, it contributes to its own replication: the individual cells of your heart will be continually replaced through the normal functioning of your metabolism. By having pumped blood, past stages of your heart contribute to its current stage. By pumping blood your heart thereby contributes to your self-maintenance, and your self-maintenance thereby contributes to the replication and repair of your heart.

These mechanisms, as I have described them, may or may not involve selection over variants. Though many theorists require selection over variants (either conscious, cultural, or natural) for etiological functions, I do not. Functions even so require history, however brief.[3]

In the next section I shall argue that the function of assertion is to induce true hearer belief. I do not know whether *asserting* is an innate repertoire triggered by normal development, but I know *communicating informatively* is. Learning surely helps explain why we assert. You assert something because you want someone to believe it, and she does. As a result, you try the same strategy again on another occasion. Self-maintenance plays a role as well. As social, linguistic creatures, using language with others helps us get around. If we think of speech acts and other linguistic devices as tools, then manufacturing doesn't seem too far-fetched either. Maybe "cultural evolution" would be a better label for this process.

Which mechanism plays the major role? I do not know. It is a difficult empirical question. But I do think we have good reason to believe, as we'll learn by the end of the chapter, that culture plays an important role, especially the role played by social norms for truth telling. Social norms for truth telling help explain why we keep reproducing *true* assertions.

2. The Function of Assertion

Why believe assertion has the proper function of inducing true belief in the audience?

2.1. Coady's Argument

In "Testimony and Observation" from 1973, C. A. J. Coady offered a thought experiment designed to show that any practice of reporting—of making assertions about verifiable states of affairs not immediately verifiable to the audience—must be, for the most part, reliable. He elaborated further in his landmark 1992 book, *Testimony: A Philosophical Study*.

His thought experiment involved a community of Martians. "Let us suppose for the moment," he writes,

> … that [these Martians] have a language we can translate (there are difficulties in this supposition …) with names for distinguishable things in their environment and suitable predicative equipment. We find however, to our astonishment, that whenever they construct sentences addressed to each other in the absence (from their vicinity) of the things designated by the names, but when they are, as we should think, in a position to report, they seem to say what we (more synoptically placed) can observe to be false.
>
> (Coady 1992, 85)

Coady then asked whether any hearer would have any reason to rely on what a speaker says in such a situation. He replies that

> … any Martian has four powerful reasons for not relying on what others appear to be telling him: (i) he finds their "reports" false whenever he checks personally on them, (ii) he finds reliance upon them consistently leads him astray in practice, (iii) he finds himself utterly unreliable in what he tells others and it is, at least in part, possible that he is not atypical, (iv) others often give chaotically different reports on those matters beyond his checking.
>
> (Coady 1992, 87)

Coady inferred there would then be no practice of making and accepting reports. Given what each Martian can expect of the other,

> … it is … very hard to imagine the activity of reporting in anything like its usual setting with the Martians, for there would surely be no reliance upon the "reportive" utterances of others. [T]he Martian community cannot reasonably be held to have the practice of reporting …
>
> (Coady 1992, 87)

Coady then concluded that any practice of reporting must be reliable.

I think Coady missed a premise (Graham 2000). With the premise he needs, his argument goes like this:

> **(1)** If everyone in a community, when making reports, always makes false reports, then no one will have any reason for relying upon and accepting those reports. They have "four powerful reasons not to do so."
> **(2)** A central reason (motive) for making reports is to get others to believe them.
> **(3)** But if no one accepts those reports, then speakers will have no reason to think that making those reports will lead hearers to believe them. (This is the missing premise.)
> **(4)** So if no one accepts reports, speakers will stop making them.

In short: if all falsehoods uttered, then no acceptances; if no acceptances, then no reason to report; if no reason to report, no reporting. Contraposed: if reports, then not all reports are false. Hence: if reports, then truths. (Echoes of Kant on the impossibility of a world where no one keeps their promises.)

This argument does not show that any practice of reporting is necessarily reliable. At best it shows not every report can be false. And "not all are false" surely does not entail "most are true." Many reporting practices may be reliable, but not necessarily so. It depends on the details of the case.

Though this argument does not go as far as Coady intended, it does bring out a very important fact about linguistic practices in general: speakers and hearers *both* need some incentive to participate. Speakers, presumably, benefit in some way by influencing hearers. But if hearers receive no benefit from being influenced, they will probably stop

responding in the desired way. Unless hearers get something out of accepting reports, they will not accept them. And if they will not accept them, speakers will not benefit by making them; speakers will lose their ability to influence hearers by making reports. Speakers then won't bother to make those reports anymore. Hearer benefits then partly explain speaker production.

2.2. Millikan's Argument

Ruth Millikan argues for a related conclusion about the indicative (the declarative) mood, the linguistic device we paradigmatically use to make reports (in particular) and assertions (in general).

She argues in general that (i) language devices produce effects that interest speakers often enough to encourage continued replication by those speakers only if hearers produce hoped-for responses often enough, and (ii) that hearers continue to produce those hoped-for responses often enough only if the results are of interest to hearers (Millikan 2004, 25). Language devices then persist only if their use benefits both senders and receivers.

She then asks why the indicative mood persists. Though inducing false beliefs sometimes benefits speakers, and it may also sometimes benefit hearers, she answers that hearers benefit by receiving true information. Dan Sperber agrees: "From the point of view of receivers, communication, and testimony in particular is beneficial only to the extent that it is a source of genuine (and of course relevant) information" (Sperber 2001, 404). Millikan then notes that if hearers never got true information from assertive utterances of declaratives in the indicative mood, they would quit forming beliefs in response. And then speakers could not even get across false information when they wanted to. And if hearers quit forming beliefs in response, then speakers would quit making those utterances. There would be no point. True hearer belief thus explains why the indicative mood persists. Millikan's argument (applied to the indicative mood) is the same as Coady's (as applied to reporting).

2.3. The Function of Assertion

Millikan then puts two and two together and draws a further conclusion about the *function* of the indicative mood: inducing true hearer belief is a proper, stabilizing function of the indicative mood:

> If no true beliefs ever resulted from hearer interpretations of indicative sentences, it is clear that indicative syntactic patterns would cease to be used first by hearers and, as a result, by speakers in the ways they now are. … The … stabilizing function of the indicative mood is thus the production of a true hearer belief. "The function" of the indicative mood is to convey information.
>
> (Millikan 1984, 58–59)[4]

Though Millikan argues that it's the indicative mood that has the proper function of inducing true hearer belief—that it gets reproduced because of this effect—the reasoning applies equally well to the speech acts we frequently perform when using the indicative mood, typically assertions (and not just reports). Millikan even goes on to apply this reasoning to assertive speech acts herself (Millikan 2005, ch. 8).

Plugging all of this into our definition, the function of assertion is inducing true hearer belief among speaker-hearer pairs, for:

1. Speaker assertions induce true hearer belief.
2. Inducing true hearer belief benefits both speakers and hearers.
3. Inducing true hearer belief explains why assertive communication persists.

I'm persuaded by these arguments. Assertion has the proper function of inducing true hearer belief.

3. The Stability Problem

This form of reasoning is an instance of a more general fact about communication in the animal kingdom. Here John Maynard Smith and David Harper from their 2003 book, *Animal Signals*:

> [If] a signal alters the behavior of others it must, on average, pay the receiver of the signal to behave in a way favourable to the signaler; otherwise receivers would cease to respond… . It follows that the signal must carry information—about the state or future actions of the signaler, or about the external world—that is of interest to the receiver. This information need not always be correct, but it must be correct often enough for the receiver to be selected to respond to it… . [Both] parties must on average benefit from the exchange; otherwise, the signaling system would cease to exist.
>
> (Maynard Smith and Harper 2003, 3)

"Honest" (reliable) signals explain the persistence of animal signals; that's why the function of animal signals is to convey accurate information. Though human assertions are different from (nonhuman) animal signals in many ways, they both share the function of conveying information from the speaker to the hearer. Human assertion is not the only communicative device that has the function of inducing true (accurate) belief (representation) in the receiver.

3.1. Influence over Information

But any signaling system faces the stability problem. The problem is straightforward. For sender-receiver systems to persist, both senders and receivers must benefit, as we've seen. To explain stability, one must then explain why. Since we know the hearer only benefits (on balance) if the sender sends accurate information, then the problem of

explaining stability becomes the problem of explaining why the sender keeps on sending "honest" (accurate, reliable) signals. If the signals are not reliable enough, the receiver will stop responding, and then the sender will quit sending. The system will collapse. Why then does the sender continually send (accurate) information, when presumably the sender sends a signal to influence the recipient? Why should the sender continue to send reliable ("honest," accurate) signals? For without reliability, to repeat, there's no stability. The *stability* problem is then the *reliability* problem: what explains the reliability (the "honesty") of signal systems?[5]

3.2. When Interests Coincide

When interests coincide, this problem is not really a problem. Kin selection (we take care of our own) or coordination (we want to work together) often straightforwardly explains the persistence of reliable signals.[6]

For example, one of the best-known forms of animal communication is the bee dance. Genetically related honey bees perform complicated dances—most famously the "waggle dance"—that reliably informs its hive mates where to find food. This form of communication is stable: both senders and receivers benefit from senders sending, and receivers receiving, signals that accurately convey the location of food. The sender benefits by sending a reliable signal. The proper function of the waggle dance is then to "induce true belief" about where to find food.

Both sides can even benefit when *overall* interests conflict. The trick is just to find a case where both sides benefit from a specific piece of information. For example, for a predator to succeed, often she needs to have the element of surprise. She does not want to attack if she knows she has been detected. That might be a huge waste of time and energy. And the prey, obviously, does not want to be attacked. If the prey detects the predator, he thereby gains by signaling to his predator that he sees her. The predator then receives the beneficial information that she's been seen, and so won't waste her time. Both sides win. Both sides benefit from an accurate signal, even though interests overall clearly conflict.[7]

3.3. When Interests Conflict

But what if a sender can gain from an unreliable—a "dishonest"—signal? What if influencing the receiver conflicts with sending accurate information? Then we should expect pressure for senders to produce unreliable signals. A predator that sends the signal "I'm a conspecific, come mate with me!" to unsuspecting prey will clearly benefit by sending that signal to a gullible receiver. When interests conflict in the particular case, we are one step away from the stability problem.

For then we should equally expect a defensive response in receiver: ignore the signal. Because "accepting" the signal is not in the receiver's interest, we should expect pressure for receivers to ignore unreliable signals. The signal will then either lose its function, change its function, or go out of existence altogether. Unreliable signals might

work for a time, but they won't last. Lie too many times and people just won't listen anymore. Aesop's fable about the boy who cried 'Wolf!' too many times clarly illustrates this point. So if senders benefit through influence, then when interests conflict, won't senders eventually misinform often enough and thereby inadvertently destroy the stability of the system?

Applying this line of reasoning to human animals, we can then ask what keeps human assertive communication reliable enough, and so up and running. Information is often very valuable. Keeping it to yourself might seem like a good thing to do. We may lie or mislead, and often do. Feeding others misinformation might serve our interests in any number of ways (Dessalles 2007; Sperber 2001; Trivers 2011). Too much lying and deception or overall incompetence thereby threatens to destabilize assertive communication (Green 2009, 157). Shouldn't the interests of speakers eventually ruin their ability to influence hearers, and thereby undermine their interest in influencing hearers? That's the stability problem.

If that's the problem, what's the solution for humans? Given conflicting interests, and given that assertive communication persists, what explains the reliability, and so the stability, of assertive communication? In other words, given that we have conflicting interests, but given that assertive communication remains up and running for all that, how have we solved the stability problem?

4. What Stabilizes Assertion?

Coordination, vigilance, and reciprocity are three frequently discussed stabilizing mechanisms.

4.1. The Scope of Coordination

When Millikan treats this issue, she largely focuses on cases of common or coordinate interests (2005, ch. 3). We frequently coordinate when using language. For example, we want to talk to each other about birds or baloney, so we coordinate on the words "bird" and "baloney" to induce thoughts in each other about birds and baloney. Since we both benefit, and those benefits sustain our use of those words, the function of those words is to induce those thoughts, even when we use those words on particular occasions for other purposes, as I just did here. That stabilizes our use of words.

But what about the indicative mood, and so reports and assertions? Millikan again focuses on coordinate interests. Speakers often want to coordinate with hearers; that's why they tell the truth. Suppose you want to help me out and pick me up at the airport. If so, I'll be sure to tell you the correct time of my arrival. It doesn't serve me to mislead you (McCready 2015, 33). What would be the point?

Millikan thus seems to think that coordination is sufficient to stabilize the indicative mood (and so, by our lights, assertion), just as coordination is clearly sufficient to stabilize

driving on the right (or on the left), even if we lie more often than we engage in reckless driving. (Note echoes of Lewis's [1969] convention of truthfulness-in-L.) According to Millikan, we've solved the stability problem by not really running into it in the first place, as our interests align well enough, often enough, to begin with. The coordinating effect is the stabilizing effect; the coordinating function is the stabilizing function.

Now Millikan *may* be right that in the human case coordination is enough to ensure the reliability—truth *often enough*—of assertion. I can't say for certain that she isn't. But what if she isn't? Sperber et al. worry that "while the interests of others often overlap with our own, they rarely coincide exactly" (2010, 360). They may conflict more often than we think. Maybe coordination alone then isn't enough. Using the indicative mood is clearly sometimes like driving on the right, but not always. When it comes to assertion, maybe we need some other reason to stay in our lane.

We should then ask whether there are other mechanisms at work that also play a role in stabilizing the reliability of assertion. Other than coordination, what else might explain the stability of our practice of assertion? It is a question worth pursuing.

4.2. Epistemic Vigilance

Sperber et al.'s answer involves "epistemic vigilance," the variety of forms of "filtering" out of unreliable assertive speech acts that we possess, the various ways we assess whether to believe what we have been told. One form involves making use of background knowledge about the reliability of the speaker, either in general, in circumstances like these, or on the subject matter at hand. Another form involves making use of background information about the subject matter itself. There are others still, including argumentation. If I challenge you to give reasons and you can't, then I've just saved myself the trouble of believing something that might just be false. They conclude that these protective measures help explain the stability of assertive communication (Dessalles 2007, 330–332; Sperber 2001, 406, 410; Sperber et al. 2010, 360, 368). I think they are certainly right that these protective measures *help* explain the reliability, and so the stability, of assertion (Graham 2010, 171–174). They are certainly a part of the overall picture. Are there other contributing factors? I believe there are.

4.3. Reputation: Direct and Indirect Reciprocity

Reputation is one. If I observe your behaviour, I can make predictions about how you will behave in the future, since I can reidentify you and remember how you have behaved in the past.

When it comes to assertions, I can keep track of your record for accuracy, judging either your competence or your sincerity. If I find you are unreliable for either reason, maybe I won't believe you next time around. I certainly won't rationally believe you the third time around, as the old saying goes.

But since you, the sender, know this too, then, provided you will want to maintain your ability to influence me—to get me to believe your assertions—you'll have a reason to assert truly even in cases where you might otherwise have an incentive to lie or mislead, assuming you know you'll be interacting with me again. You have a reason to tell me the truth now—even though you would otherwise rather not—so that you can continue to influence me in the future.

This is more than just the hearer protecting himself through vigilance. This is vigilance playing a part in a feedback mechanism *to the sender*, giving her an incentive to tell the truth (McCready 2015, 22–24, 35; Nowak 2006). She *learns* to tell the truth.

This is a case of so-called direct reputation (or direct reciprocity) where Annie develops a reputation with Barney. If Annie lies to Barney today, and Barney figures that out, then Barney does not believe Annie the next day. Annie then learns that she's better off sticking to the truth, so as not to forgo future opportunities to influence Barney.

Why "reciprocity"? Because it's a case of trading. You give me the truth today, and I'll believe you again tomorrow. You give me something, I'll give you something back. But if you mislead me today, all bets are off. Truth is the price you'll pay to keep your ability to influence me in play.

"Indirect reputation" (or "indirect reciprocity") has the same effect. Annie develops a reputation with an onlooker, Christine. Christine observes Annie lie to Barney today, then Christine does not believe Annie tomorrow. Annie then learns she's better off sticking to the truth. And then there is the case of "reported reputation" (also often grouped as a case of "indirect reciprocity") where Annie lies to Barney today; Barney tells Christine about Annie's lie; Christine does not believe Annie tomorrow. Or where Christine observes Annie lie to Barney today; Christine tells Dora about Annie's lie; Dora does not believe Annie tomorrow. You get the picture.

Damage to your reputation involves a loss of credibility, and so your ability to influence. It's a penalty you'll pay if caught. Not willing to take a risk, you are thereby deterred (when it works) from misleading or deceiving. Reputation is then a mechanism that helps stabilize assertion. [8]

Is there some other penalty you might pay if caught, some other penalty that might incentive honest communication and thereby deter dishonest communication? There is. If I catch you asserting something false, besides thinking twice about believing you in the future, I might also *punish* you for failing to live up to a *norm* for making assertions. I might thereby provide another incentivize for you to stay in line in the first place. I turn to that stabilizing mechanism next.

5. A Social Norm for Truth Telling?

This mechanism—the enforcement of social norms through punishing—is sometimes called “strong reciprocity” reciprocity” (Bowles and Gintis 2011). Since I think that this mechanism is a powerful mechanism in human affairs, with, as we’ll see, a clear application to problem of the stability of assertion, I’ll spend a good deal of time conveying how it works (though I’m sure social norms aficionados will think I haven’t spent enough). I’ll first describe the general phenomenon before applying it to the case of assertive communication.

5.1. Hold the Door for Strangers

What then are social norms? There are a number of different definitions of social norms within the philosophical and social scientific literature.[9] Our purposes won’t require an exact definition. The main idea is that humans have a strong tendency to “norm” one another. We influence one another’s thought and behaviour through expressing and acting on judgments about how we think we ought to think or behave. These attitudes create “normative” social pressure for others to conform. And because these normative attitudes are often very effective at influencing one another’s thought and behaviour, a great deal of human life becomes regular and predictable as result.

Here is a simple example. In many parts of the United States, when you walk through a door to enter a building, such as a store, and there are strangers walking in behind you, you are supposed to help with the door. Once you do, they are supposed to acknowledge your help. If you do not hold the door, they might think you are a jerk, or experience some other negative reaction. If you hold the door and they fail to express any gratitude, you too will react negatively. If you see them later in the store and they clearly need help bringing down a large object from a high shelf that you could easily reach, you might decide to simply walk past. You might even grunt disapprovingly as you do. You might, in effect, punish them for failing to express gratitude.

Because Americans typically have these normative attitudes about holding the door, you can bet that when you are going shopping that people in front of you will hold the door and the people behind you will express some degree of gratitude when you do. In South Korea, on the other hand, there is no such norm for holding the door. Koreans—insofar as they are in Korea—don’t have these normative attitudes. You just walk through and let the door close right behind you without a second thought. Koreans aren’t bothered by this at all. Indeed, if you, as an American, stop and hold the door, the Koreans behind you might just wonder what the heck you are doing.

5.2. Principles, Attitudes, Punishments

Social norms involve at least three components. First, when there is a social norm, we have a normative principle (Brennan et al. 2013) or behavioural rule (Bicchieri 2006, 2014, 2017) that one *ought* to A in C (prescriptive rule) or that one *ought not* A in C

(proscriptive rule). In logical space there are as many normative principles as you please. Some are objectively valid. Many are not. Some are embraced by groups; some are not. Normative principles are "out there" waiting to be embraced.

This takes us to our second component, normative *attitudes*. Where there is a social norm, a normative principle has been embraced by a group. How do we embrace a normative principle? Through *normative attitudes* that involve—that in various ways *embrace*—the normative principle (Brennan et al. 2013). Normative principle-involving normative attitudes include *judging* that one must A in C ("One must hold the door"), *demanding* that others A in C ("You really should hold the door"), *disapproving* of those who do not A in C ("Really pal, you could have easily helped with the door"), judging that it is *appropriate* to disapprove of those who do not A in C ("Can you believe that jerk didn't help with the door?!" "I know, just who does he think he is?!"), *acknowledging* the legitimacy of such criticisms and demands when received from others ("Oh, yes, you are right. I was in the wrong. My bad. I apologize."), and *judging* that one is *entitled* (and so are others) *to sanction* violations ("And so that's why I wouldn't help them bring down that heavy box!" "Right you are! I even walked away when they asked for help in the parking lot, pretending I was too busy to help.").

When enough people in a group have these normative attitudes, and when enough people know that enough people in the group have these normative attitudes, then the group has embraced the normative principle or rule. It's then a norm for the group that one ought to A in C. Social norms are widely embraced normative principles.

Which brings us to our third component. We are often moved to act in accord with these principles when socially embraced. Social norms motivate.

This may be because we've internalized the normative attitudes ourselves; we believe, of ourselves, that this is how we *ought* to act. We then often experience guilt or shame when we fail to act this way. Even thinking about failure might evoke strong feelings of anxiety, guilt, or shame (Gintis 2003; Henrich and Henrich 2007). "Social norms," Jon Elster writes, "have a strong grip on the mind that is due to the strong emotions they can trigger" (Elster 1989, 100). He goes on to say that "shame is not only a support of social norms, but *the* support" (1999, 145–146). We thereby punish *ourselves* for failure to conform. That possibility often motivates us to conform in the first place, even if we have a reason to do otherwise. We deter ourselves from breaking the rules in the first place.

Bicchieri (2014, 2017), on the other hand, thinks third-party sanctions play the largest role. She thinks we are primarily concerned with what others will think about our actions, and what consequences that may have for us. The possible punishment from others often motivates conformity. We deter others from breaking the rules.

One of the biggest third-party sanctions is just the loss of social approval or esteem that comes from not doing what others think we ought to do, for failure to live up to their normative attitudes. As a result, we often moved to act in accord with a socially embraced normative principle for fear of disapproval, for fear of loss of esteem. Adam Smith's 1759

classic *The Theory of Moral Sentiments* emphasizes the motivating force of our desire for esteem. This desire is also emphasized by Philip Pettit (1990, 2002) and Brennan et al. (2013), among others, in their work on the motivating power of normative attitudes.[10]

Shame and guilt, on the one hand, where we punish ourselves, and disapproval (and so loss of esteem), on the other, where we punish others, are both automatic (Brennan et al. 2013). For the most part, you can't avoid feeling bad when breaking one of your own norms, and you can't avoid feeling disapproval of another when you observe them breaking a norm, and so you can't but anticipate others' disapproval when they observe you breaking a norm. You sometimes can't avoid punishing yourself or others. Even if these mechanisms are costly, they can't be avoided. And that's why, in part, they are so powerful at motivating compliance. Though they are not the only forms of punishment—there are many—they are two of the most discussed.

Bicchieri (2006, 2017) emphasizes the social pressure to conform. "With social norms the normative influence is strong and plays a crucial role in driving compliance. It matters to us that most people in our reference network believe we ought to conform to a certain behavioral [rule] … [T]he social pressure to conform, expressed in social [normative] expectations that one is to conform, is a powerful motivator" (Bicchieri 2017, 34–35). This pressure to conform "must be emphasized," Bicchieri writes, because left to our own devices we'd often rather do something else. "With a [social] norm, there is often the temptation to transgress it—this is precisely why social norms must be socially enforced" (Bicchieri 2017, 39). Surely the social norm to wait in line explains a good deal of the daily grind. Our normative attitudes and our beliefs about the normative attitudes of others then make a difference in explaining why we do what we do, especially when we might have a reason to do something else instead.

5.3. Solving Mixed-Motive Games

That social pressure to conform can move us to act in ways we might otherwise not choose is partly why social theorists are interested in social norms, for social norms have the potential to solve mixed-motive games (what are also called cooperation games). In a mixed-motive game there is a good to be attained by cooperating—a good in which all members of the groups can partake to the extent that it is produced—and an individual cost to be paid in cooperating. In these games, the marginal benefit gotten by the individual from the individual's own contribution (cooperating) is less than the cost of cooperating. The individual then has an incentive to free-ride instead. But so does everyone else, and cooperation as a result collapses. Everyone is worse off in the end.

The Prisoner's Dilemma is the classic example. In the Prisoner's Dilemma, the dominant strategy for each player is defection. If the other player cooperates, you are better off by defecting. If the other player defects, you are better off defecting. Thus, no matter what the other person does, you should defect. But since this reasoning applies to both players, both should defect. But—and this is why it is a cooperation game—it would be better if both players cooperated instead of both defecting.

To “solve” a mixed-motive game then requires adding another motive to the game—another payoff—that makes cooperation the best choice for all players. In the Prisoner’s Dilemma, if there were a social norm against defection (“don’t be a rat!”), then the players would have different preferences; they would have the conditional preference to conform to the norm, either perhaps because they endorse the legitimacy of the principle (“friends don’t rat on friends”) or because they fear sanctions, as we’ve just discussed. Once the norm is in place—once players have the “right” social conditional preferences and empirical and normative expectations about others in their reference network—we’ve changed the incentive structure, and thereby “solved” the game by changing the game (changing the payoff structure) so that cooperation (not defection) is the best strategy for all players. The temptation payoff no longer looks so rational after all. The players are now deterred from defecting, despite their temptation. The cooperative outcome emerges. Social norms are game changers.[11]

5.4. Tell the Truth

A social norm for truth telling is exactly one of those norms. Though speakers often have a motive to misinform, that motive is counterbalanced by the motivation to conform with the social norm to inform. Communication that often starts as a mixed-motive game transforms into a cooperation game by a social norm for truth telling.

Is truth telling a social norm? Yes, indeed. It’s a textbook case of a social norm: open any textbook discussion of social norms and *tell the truth* will be on the shortlist of examples. In her contribution to a recent introductory anthology of social science, Bicchieri proclaims that “[e]ach group has its own norms, and some, like reciprocity or truth telling, [are] very general, spanning all groups” (2014, 208). When providing examples of social norms in his book *Social Action*, Seumans Miller (2001) lists refrain from violence, remain faithful to one’s spouse, avoid incest, keep promises, and tell the truth. Bowles and Gintis (2011) emphasize the social norm that one ought to tell the truth a number of times in their book *A Cooperative Species*. Philip Pettit (1990, 2018) treats truth telling as the central case in his work on social norms. As I said, it’s a textbook case.

I think it’s clear that we *prescribe* providing true and relevant information; we believe, presume, presuppose, or are disposed to believe that we *ought* to be good informants. Many of us internalize these attitudes and thereby care deeply about providing true and relevant information.

We express our commitment to this on many occasions; our evaluations and assessments reveal our commitment to the norm. Our evaluations of what others say show we prescribe providing true and relevant information. We teach the importance of truth telling. We praise people for their honesty. We admire those who (seem to) know a lot and can inform us about items of interest. We also regularly criticize people who deceive or otherwise mislead.

Our commitment is also revealed in our social emotions. We sometimes feel horribly let down or betrayed by those who deliberately fail to tell us the truth. We even judge ourselves. If I realized I gave you bad directions, I might feel embarrassed. If I lie, I might feel guilty. If I get called out for lying or speaking incorrectly, I might feel ashamed.

Our commitment drives our third-party evaluations and punishments. I'll negatively evaluate others who provide poor information. I'll feel punitive attitudes toward those who lie to others. I may even punish them. One lie can sever a relationship, destroy a reputation, or lead to serious third-party punishments. These prescriptions are clearly experienced categorically, as not depending on what we want or desire. Even if there's nothing in it for me, I know I'm supposed to provide true and relevant information.

We do not embrace the principle *tell the truth* in the literal sense of requiring people to utter truths nonstop. That is not what is meant when we prescribe truth telling. Here are two approximations of normative principles we have in mind:

- If someone needs information whether P, and you have it, provide it.
- If you are going to communicate anything at all, do not provide false or misleading information.

Philosophical thought experiments, experimental philosophy, and solid anthropological inquiry are sure to uncover a number of more precise normative principles, clustering around the theme of being informative when asserting to another that such and such is true, even when one might rather settle for influence without information. There might just be a whole score of norms governing assertion—*assert only what you know, retract your assertion if you cannot defend it*, and so on—that play a substantial role in stabilizing the reliability of assertion.[12]

6. Conclusion

Exploiting reasoning from Coady and Millikan—reasoning that is an instance of a more general fact about human and nonhuman animal communication—I have argued that the proper function of assertion is to induce true hearer belief. Assertion may have other proper functions. I have remained silent on that question, though I am sure it does.

But functions are not guaranteed to persist, especially when something might destabilize their persistence. When it comes to animal signals—human and nonhuman—the signal will persist only if both sides, sender and receiver, continue to benefit. Given that the interests of speakers qua speakers and hearers qua hearers do not always align, there must be some explanation for the stability of assertion. It may be, as Millikan seems to think, that our interests align often enough that the stability problem isn't much of a problem for us. We're then more like bees than predator-prey pairs. Thank goodness!

But even so, we might naturally investigate other solutions, for we are not obviously always aligned in our interests for all that. Far from it. Sperber et al. explore the various ways hearers protect themselves. But are there other ways we might deter speakers from making unreliable assertions? A loss of reputation, and so a loss of credibility that undermines the ability to influence, is one such mechanism. Speakers pay with truth today so they might influence again on another day. Socially imposed punishments arising from social norms are another such mechanism, as we've just seen. Social norms for truth telling play their part in stabilizing the reliability of assertion, and thereby help explain the persistence, and so the proper function, of assertive speech acts.

References

Alexrod, R., and W. D. Hamilton. 1981. "The Evolution of Cooperation." *Science* 211: 1390–1396.

Bach, K., and R. M. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Berg, J., J. Cickhaut, and K. McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10: 122–142.

Bicchieri, C. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.

Bicchieri, C. 2014. "Norms, Conventions, and the Power of Expectations." In *The Philosophy of Social Science: A New Introduction*, edited by N. Cartwright and E. Montuschi, 208–229. Oxford: Oxford University Press.

Bicchieri, C. 2017. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York: Oxford University Press.

Bowles, S., and H. Gintis. 2011. *A Cooperative Species*. Princeton, NJ: Princeton University Press.

Boyd, R., and P. Richerson. 1989. "The Evolution of Indirect Reciprocity." *Social Networks* 11: 213–236.

Brennan, G., L. Erikson, R. Goodin, and N. Southwind. 2013. *Explaining Norms*. Oxford: Oxford University Press.

Coady, C. A. J. 1973. "Testimony and Observation." *The American Philosophical Quarterly* 10: 149–155.

Coady, C. A. J. 1992. *Testimony: A Philosophical Study*. Oxford: Clarendon Press.

Dawkins, R., and J. Krebs. 1978. "Animal Signals: Information or Manipulation?" In *Behavioral Ecology: An Evolutionary Approach*, edited by R. Dawkins and J. Krebs, 282–309. Oxford: Blackwell Scientific.

Dessalles, J.-L. 2007. *Why We Talk: The Evolutionary Origins of Language*. Oxford: Oxford University Press.

Elster, J. 1989. *The Cement of Society*. Cambridge: Cambridge University Press.

Elster, J. 1999. *Alchemies of the Mind*. Cambridge: Cambridge University Press.

Faulkner, P. 2010. “Norms of Trust.” In *Social Epistemology*, edited by A. Haddock, A. Millar, and D. Pritchard, 129–147. Oxford: Oxford University Press.

Faulkner, P. 2011. *Knowledge on Trust*. Oxford: Oxford University Press.

Fehr, E. 2004. “Human Behavior: Don’t Lose Your Reputation.” *Nature* 432: 449–450.

Fitzgibbon, C. D., and J. H. Fanshawe. 1988. “Stotting in Thomson’s Gazelles: An Honest Signal of Condition.” *Behavioral Ecology and Sociobiology* 23: 69–74.

Fricker, E. 2006. “Second-Hand Knowledge.” *Philosophy and Phenomenological Research* 73: 592–618.

Fricker, E. 2017. “Norms, Constitutive and Social, and Assertion.” *American Philosophical Quarterly* 54: 397–418.

Garcia-Carpintero, M. 2004. “Assertion and the Semantics of Force-Markers.” In *The Semantics/Pragmatics Distinction*, edited by Claudia Bianchi, 133–166. Palo Alto, CA: CSLI Publications.

Goldberg, S. 2015. *Assertion*. Oxford: Oxford University Press.

Goldberg, S. 2017. *To the Best of Our Knowledge: Social Expectations and Epistemic Normativity*. Oxford: Oxford University Press.

Graham, P. J. 2000. “The Reliability of Testimony.” *Philosophy and Phenomenological Research* 61: 695–709.

Graham, P. J. 2010. “Testimonial Entitlement and the Function of Comprehension.” In *Social Epistemology*, edited by A. Haddock, A. Millar, and D. Pritchard, 148–174. Oxford: Oxford University Press.

Graham, P. J. 2011. “Does Justification Aim at Truth?” *Canadian Journal of Philosophy* 41: 51–72.

Graham, P. J. 2012a. “Epistemic Entitlement” *Nous* 46: 449–482.

Graham, P. J. 2012b. “Testimony, Trust, and Social Norms” *Abstracta* Special Issue VI: 92–116.

Graham, P. J. 2013. “Review of Faulkner, *Knowledge on Trust*.” *Notre Dame Philosophical Reviews* **https://ndpr.nd.edu/news/knowledge-on-trust/**.

Graham, P. J. 2014. “Warrant, Functions, History.” In *Naturalizing Epistemic Virtue*, edited by A. Fairweather and O. Flanagan, 16–36. Cambridge: Cambridge University Press.

Graham, P. J. 2015. “Epistemic Normativity and Social Norms.” In *Epistemic Evaluation*, edited by J. Greco and D. Henderson, 247–273. Oxford: Oxford University Press.

Graham, P. J. 2017. “The Function of Perception.” In *Virtue Scientia: Bridges between Virtue Epistemology and Philosophy of Science*, edited by A. Fairweather, 13–31. 366: Synthese Library.

Graham, P. J. 2018. “Social Knowledge and Social Norms.” In *Knowledge in Contemporary Philosophy*, edited by S. Hetherington and M. Valeros, 111–137. Bloomsbury.

Green, M. 2009. “Speech Acts, the Handicap Principle and the Expression of Psychological States.” *Mind & Language* 24: 139–163.

Henderson, D. 2012. “Norms.” In *The Oxford Handbook of Philosophy of Social Science*, edited by H. Kinkaid, 367–382. Oxford: Oxford University Press.

Henderson, D., and P. Graham. 2017a. “Epistemic Norms and the “Epistemic Game” They Regulate: The Basic Structured Epistemic Costs and Benefits.” *American Philosophical Quarterly* 54: 367–382.

Henderson, D., and P. Graham. 2017b. “A Refined Account of the “Epistemic Game”: Epistemic Norms, Temptations and Epistemic Cooperation.” *American Philosophical Quarterly* 54: 383–395.

Henrich, N., and J. Henrich. 2007. *Why Humans Cooperate: A Cultural and Evolutionary Explanation*. New York: Oxford University Press.

Herbert Gintis. 2003. “Solving the Puzzle of Prosociality.” *Rationality and Society* 15(2): 155–187. doi:10.1177/1043463103015002001.

Kelp, C. 2016. “Assertion: A Function-First Account.” *Nous* 50: 411–442.

Krebs, J., and R. Dawkins. 1984. “Animal Signals: Mind-Reading and Manipulation.” In *Behavioral Ecology: An Evolutionary Approach*, 2nd ed., edited by J. Krebs and R. Dawkins, 381–402. Oxford: Blackwell Scientific.

Lewis, D. 1969. *Convention: A Philosophical Study*. Oxford: Blackwell.

Maynard Smith, J., and D. Harper. 2003. *Animal Signals*. Oxford: Oxford University Press.

McCready, E. 2015. *Reliability in Pragmatics*. Oxford: Oxford University Press.

McDowell, J. 1980. “Meaning, Communication, and Knowledge.” In Philosophical Subjects: Essays Presented to P.F. Strawson, edited by Z. van Straaten, 117–139. Oxford: Clarendon Press.

McGlauglin, P. 2001. *What Functions Explain*. Cambridge: Cambridge University Press.

Michaelian, K. 2013. “The Evolution of Testimony: Receiver Vigilance, Speaker Honesty and the Reliablity of Communication.” *Episteme* 10: 37–59.

Miller, S. 2001. *Social Action: A Teleological Account*. Cambridge: Cambridge University Press.

Millikan, R. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, MA: The MIT Press.

Millikan, R. 2004. *Varieties of Meaning*. Cambridge, MA: The MIT Press.

Millikan, R. 2005. *Language: A Biological Model*. Oxford: Oxford University Press.

Nowak, M. 2006. “Five Rules for the Evolution of Cooperation.” *Science* 314: 1560–1563.

Pettit, P. 1990. “Virtus Normativa.” *Ethics* 100: 725–755.

Pettit, P. 2002. *Rules, Reasons and Norms*. Oxford: Oxford University Press.

Pettit, P. 2018. *The Birth of Ethics: Reconstructing the Role and Nature of Morality*. Oxford: Oxford University Press.

Pelling, C. 2013. “Assertion and the Provision of Knowledge.” *The Philosophical Quarterly* 63: 293–312.

Reynolds, S. 2017. *Knowledge as Acceptable Testimony*. New York: Cambridge University Press.

Scott-Phillips, T. 2008. “On the Correct Application of Animal Signaling Theory to Human Communication.” The Evolution of Language, Proceedings of the 7th International Conference (EVOLANG7): 275–282.

Searcy, W., and S. Nowicki. 2005. *The Evolution of Animal Communication*. Princeton, NJ: Princeton University Press.

Simion, M. 2018a. “Saying and Believing: The Norm Commonality Assumption.” *Philosophical Studies*, doi:10.1007/s11098-018-1105-8

Simion, M. 2018b. “The Explanation Proffering Norm of Moral Assertion.” *Ethical Theory and Moral Practice* 21: 477–488.

Simion, M., and C. Kelp. 2019. “Assertion: The Constitutive Norms View.” In *The Oxford Handbook of Assertion*, edited by S. Goldberg Oxford: Oxford University Press.

Sober, E. 1994. “The Primacy of Truth-Telling and the Evolution of Lying.” *From a Biological Point of View*. Cambridge: Cambridge University Press.

Sperber, D. 2001. “An Evolutionary Perspective on Testimony and Argumentation.” *Philosophical Topics* 29: 401–413.

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G. & Wilson, D. 2010. “Epistemic Vigilance.” *Mind & Language* 25: 359–393.

Sperber, D. 2013. “Speakers are Honest Because Hearers are Vigilant: Reply to Kourken Michaelian.” *Episteme* 10: 61–71.

Trivers, R. 2011. *Deceit and Self-Deception: Fooling Yourself to Fool Others*. London: Allen Lane.

Turri, J. 2016. “The Point of Assertion Is to Transmit Knowledge.” *Analysis* 76: 130–136.

Wright, L. 1973. “Functions.” *The Philosophical Review* 82: 139–168.

Wright, L. 2012. “Epilogue: Revisiting Teleological Explanations: Reflections Three Decades On.” In *Functions: Selection and Mechanisms*, edited by P. Huneman, 233–243. 363 Synthese Library.

Notes:

(1.) I first argued for this and defended it against a variety of objections in “Testimonial Entitlement and the Function of Comprehension” (Graham 2010, §§4–5). Though I shall present the argument again here, I shall not review the objections and reply.

John McDowell (1980) claimed conveying knowledge is the function of assertion. I discuss McDowell’s paper briefly in Graham (2010). Manuel Garcia-Carpintero (2004) argued that assertion is governed (constituted?) by the norm “one must (assert P) only if one’s audience comes thereby to be in a position to know P”. Charlie Pelling (2013) asserts without argument that the function of assertion is to transmit knowledge. John Turri (2016) likewise asserts without argument that the point of assertion is to transmit knowledge. Chris Kelp and Mona Simion (Kelp 2016; Kelp and Simion 2019) argue that the function of assertion is to induce knowledge in the hearer. Though worthy of attention, I shall not discuss the “knowledge or truth?” issue here.

I shall use “assertion” somewhat loosely to cover an unspecified subset of the constative category of speech acts. For an informative taxonomy of speech acts, especially constatives, see Bach and Harnish (1979).

(2.) I discuss some of the connections between assertion, the epistemology of testimony, epistemic evaluations, and social norms in Graham (2012b, 2013, 2015, 2018) and work in preparation. David Henderson and I discuss prescriptive epistemic norms for belief production as social norms in Henderson and Graham (2017a, 2017b). Other authors who pursue a social norms approach to the reliability of assertion (or rather the reliability of testimony or telling) include Paul Faulkner (2010, 2011) and Steven Reynolds (2017), among others. See also Fricker (2006, 2017).

(3.) I have discussed functions in a number of places, including Graham (2010, 2011, 2012a, 2014, 2017). For discussion of our working definition, see Graham (2014). It is inspired by McGlaughlin (2001), among others, including Wright (2012). Because of the feedback condition, it is an etiological, historical account.

It is not, however, a selection-based account. For a time, I followed Millikan (and others) in assuming that functions required a history of selection (whether natural selection, individual selection [learning], or cultural selection), for example, Graham (2010, 154; 2011, 60; 2012a, 450). But then in "Warrant, Functions, History" (Graham 2014), I expressed doubts. Functions require history, but not necessarily a history of selection over variants. Our working definition leaves the role of selection open. It depends on the feedback mechanism.

There is a lamentable tendency in philosophy to assume that only God or Mother Nature can be the source of proper functions. Atheists and admirers of Swampman then shy away from proper functions altogether. But this assumption is demonstrably mistaken. For one, God's design, as such, does not create etiological functions anyway. And second, Mother Nature is but one of many feedback mechanisms generating etiological functions.

One reason to prefer an etiological account is that it so easily explains what counts as normal conditions and normal functioning. Normal conditions are those conditions (and conditions similar in type) where the item produced the effect that (partly) explained the persistence of the device. Normal functioning (the normal operation or working of interconnected parts of the device) is then operating or working the way the item worked or operating so as to produce the effect that (partly) explains the persistence of the item (Millikan 1984, 33–34; Graham 2012a, 459–460).

Another reason to prefer an etiological account is that it so easily grounds the difference between functional and nonfunctional ("coincidental" or "accidental") effects (Wright 1973). The heart pumps blood and makes a highly regular, rhythmic thumping noise. But making that noise does not explain why hearts persist.

Simion (2018a, n. 1; 2018b, n. 3) says I use "function," "telos," "aim," and "purpose" talk interchangeably. I don't. I rarely if ever use "telos." I use "aim" sparingly. When I do, I either use it literally for an agent's intention or goal, or I use it metaphorically for function. When I use it metaphorically, I say I am (e.g., 2011, 56–57, 60, 62) or I put it in scare quotes (e.g., 2012a, 461–464).

(4.) This argument appears throughout Millikan's writings. See Millikan (1984, 30–32, 52–60; 2004, 25–26; 2005, 14, 16, 44–45, 57–58, 90, 94–95, 153, 157, 169, 172). Millikan does not restrict her reasoning to the indicative mood. She makes similar arguments for the imperative (we use it to make commands; its function is hearer compliance) and the interrogative (we use it to ask questions; its function is informing the speaker). Consistent with her overall program, she argues there are functions wherever there are linguistic devices.

(5.) For the locus classicus of “influence over information” as the driving force in animal communication, see Dawkins and Krebs (1978) and Krebs and Dawkins (1984). My understanding benefitted from Sober (1994); Sperber (2001); Maynard Smith and Harper (2003); Searcy and Nowicki (2005); Scott-Phillips (2008); and Green (2009).

(6.) In Graham (2012b) I stressed that parental concern explains why so much of what parents say to their children is true, even if all of it isn’t. For similar reasons the circle of reliable testimony extends to the extended family.

(7.) Birds and antelopes demonstrably stare at cats, large and small. Gazelles “stott” on the savannah to signal to their predators they’ve been seen. See Fitzgibbon and Fanshawe (1988).

(8.) Reputation sometimes seems to stabilize reliable signals in the animal kingdom, though it is unlikely to be common because of the mental abilities required. For discussion and references, see Maynard Smith and Harper (2003). Should we think that reputation plays a large role in stabilizing human assertive communication? As I said in the text, which mechanism is at work is surely an empirical question. For evidence in favor of its effectiveness, see Alexrod and Hamilton (1981), Berg et al (1995), Fehr (2004). But for some doubts about the scope of reciprocity as a stabilizing mechanism, see Boyd and Richerson (1989) and Henrich and Henrich (2007).

(9.) I have learned from Elster (1989, 1999); Pettit (1990, 2002); Miller (2001); Bicchieri (2006, 2014, 2017); Henderson (2012); and Brennan et al. (2013).

(10.) I discuss social norms and punishment mechanisms somewhat further in Graham (2015: 252–255).

(11.) Why do social norms emerge that help solve mixed-motive games? They do not simply emerge because it would be better if they did. “The good” is not, pace Plato, the reason for all things. Armchair, a priori state-of-nature reasoning may be illuminating in many ways, but it doesn’t itself explain the emergence of social norms that ensure cooperative outcomes. If only it did; the world would be a much better place. I won’t say anything here myself about the empirical emergence of such norms. But see Graham (2015), especially §3 and references cited therein, for empirical explanations of the emergence of social norms that encourage cooperative outcomes.

(12.) But how could *tell the truth* be a social norm? Isn’t it a moral norm? Reply: Suppose for the sake of argument that tell the truth is an objectively valid *moral principle*. Qua objectively valid moral principle, but without realization via normative attitudes, especially normative attitudes that in turn motivate behavior, especially in the face of conflicting motives, it may very well lack the degree of motivating force required to help explain what underwrites the reliability, and so the stability, of assertion. So the principle must be embraced by normative attitudes to do the work required. But once embraced by normative attitudes, we have a social norm, even if also an objectively valid moral

principle. Social norms are normative principles realized in motivationally effective normative attitudes.

But aren't social norms essentially culturally varying, where moral norms are universal? If the norm *tell the truth* is going to explain the reliability of a universal speech act type across human cultures, then how could something varying explain something universal? Reply: First, as the quotes just given emphasize, there are social norms that are cultural universals. Hence social norms are not ipso facto culturally varying, ipso facto nonuniversal. Second, a strictly universal norm isn't required to solve our problem. We were looking for various mechanisms that could contribute to the reliability, and so the stability, of assertion. These mechanisms might vary in kind, in extent, and so on, in different cultures, different groups, and different times as places. Even if they vary, that is no obstacle to our thesis.

But doesn't the idea that I am motivated to tell the truth, when I do, by social norms, by social pressure and the fear of punishment, ring false? Yes and no. Our motivations are complicated. Usually, as many authors have emphasized, by default we just tell the truth. We have not been assuming that, by default, we aim to mislead, but need some overwhelming threat hanging over our head to keep us in line. Often we just want to share. Often we want to help. Often we want to coordinate. We don't need a social norm for that. But often we do not want to share, help, or coordinate. That's when social norms enter the picture. When we internalize those norms, we won't conform, when we do, out of social pressure seen as "outside" but because of social pressure internalized. We won't be thinking of what others might think or do, but that others think and do what they do plays a role anyways in shaping our motivational psychology. And then, as we all know, there are times when we are tempted, even so, to either keep something to ourselves, or provide misleading or even false information. Avowed social pressure doesn't always carry the day, of course, but even so sometimes it does.

Peter J. Graham

Peter J. Graham is Professor of Philosophy and Linguistics at the University of California, Riverside, and formerly Associate Dean for Arts & Humanities. He specializes in Epistemology and the Philosophy of Psychology, having developed proper functionalist approaches to the epistemologies of perception and testimony. He has held visiting research and teaching positions at the Universities of Aberdeen, Cologne, Edinburgh, and Oxford, the Claremont Colleges, and Yonsei University. He is the co-editor of Epistemic Entitlement (Oxford, 2009) and The Routledge Handbook of Social Epistemology (Routledge, 2019).