

## TESTIMONY, TRUST, AND SOCIAL NORMS

Peter J. Graham

Paul Faulkner's book revolves around "the problem of cooperation." According to Faulkner, when we communicate with another as to the facts, we face a situation akin to a prisoner's dilemma (2011: 6). In a prisoner's dilemma, our aggregate well-being will be maximized if we both cooperate. However, given the logic of the situation, it looks like the rational thing to do is defect. We're then faced with a problem: how to ensure the cooperative outcome? A "solution" does just that. Faulkner calls the analogous problem the "testimony game." In the testimony game, there is a sender and a receiver. The receiver wants the truth, but the sender wants to influence what the hearer believes. Given the logic of the situation, it doesn't look like it is always rational for the sender to provide what the receiver requires. The receiver is then faced with a "problem."

Faulkner argues that this problem entails that the hearer must have a reason for thinking the speaker's "purpose in utterance is informative" for "warranted uptake" (2011: 133), no matter how knowledgeable and informative the speaker happens to be (2011: 128). The hearer would need a reason for thinking the speaker, at least on this occasion, is one of the "good guys." The "solution" to the "problem" would then consist in the hearer's reasons for uptake.

Faulkner argues for a "trust-based" solution, where the hearer's attitude of affective trust warrants uptake. When the hearer affectively trusts the speaker, Faulkner argues, the hearer's trust both rationalizes the hearer's reliance and moves the speaker to choose the cooperative outcome, and thereby solves the problem. Hence the title, *Knowledge on Trust*.

I will argue that *the hearer's* attitude of affective trust isn't fundamental. A hearer can be warranted, and even acquire so-called testimonial knowledge, without affectively trusting the speaker in Faulkner's sense. What solves the problem isn't *the hearer's* trust; it's *the speaker's* trustworthiness.

### The Problem of Cooperation

Let's begin by saying more about the problem of cooperation. Since the problem concerns the logic of communication, I shall sometimes call it the *problem of communication*. Faulkner puts the problem in terms of rational choice explanations of behavior:

Suppose...a subject's individual beliefs and desires—his preferences—explain...explain how the subject acts...that rational action aims at the satisfaction of the actor's individual preferences. (2011: 4)

Now the “problem” arises because speakers and hearers have cross-purposes:

[S]peakers and audiences have different interests in communication.... Our interest, qua audience, is learning the truth. Engaging in conversations as to the facts is to our advantage as speakers because it is a means of influencing others: through an audience's acceptance of what we say, we can get an audience to think, feel, and act in specific ways. So our interest, qua speaker, is being believed...because we have a more basic interest in influencing others....[T]he *commitment to telling the truth would not be best for the speaker*. The best outcome for a speaker would be to receive an audience's trust and yet have *the liberty to tell the truth or not*. (2011: 5-6, emphasis added)

So audiences want the truth, but speakers don't always want to provide it. And if we act because of our interests, speakers won't always have audience's interests in mind.

...testimony [can fail to be true] because it was not produced with the intention of getting us, as audience, to believe truly. Giving testimony is something we, as speakers, do for reasons, and *our reasons need not put our audiences' informational needs first*. (2011: 132, emphasis added)

There are two typical reactions to this in the literature. On the first, all it shows is that in many but not all cases, and probably not most, a hearer will need additional supporting evidence to defeat an occasion specific reason for thinking the speaker might actually have chosen to deceive or mislead. Or it may show instead, or in addition, that the hearer should be able to reliably tell when someone is apt to lie or mislead. On the second, this point shows that in each and every case of testimony there is a standing defeater that the hearer must overcome with positive reasons for thinking the speaker is

not likely to lie or mislead. On the first, understanding the speaker is enough for *prima facie, pro tanto* warrant, warrant that sometimes needs supplementation for *on balance* warrant. On the second, understanding the speaker is never, as such, sufficient for *prima facie, pro tanto* warrant; the hearer must always possess, in each and every case, a reason for thinking the speaker is sincere to enjoy any warrant at all, *prima facie* or on balance.

Faulkner seems to fall within the latter camp. He thinks the problem shows that:

....it is not reasonable to trust without a *supporting reason* that rationalizes trust. *The acceptance of testimony must be backed by reasons if it is to be reasonable* [and so knowledge]. (emphasis in original 2011: 6-7; 10, 20-1,114, 118-19)

In short, since communication is like a prisoner's dilemma, the hearer needs a reason for thinking or presuming that the speaker has chosen the cooperative, helpful outcome. Like the prisoner's dilemma where each party knows the other has a reason to defect, the hearer knows in the "testimony game" that the speaker may very well fail to choose the cooperative outcome, for the speaker doesn't, qua speaker, prefer being helpful (even though, as a matter of fact, he may say what's true). The very logic of communication thereby creates a standing defeater that the hearer needs to overcome with positive reasons for thinking the speaker shall choose the cooperative outcome. This is why the hearer (the audience) must have a reason for thinking or presuming that the "speaker's purpose in utterance is informative" for "warranted uptake" (2011: 133). Otherwise it's not rational, *even if the speaker has chosen the cooperative outcome*. The hearer can't simply take it for granted, as it were, that the speaker is one of the good guys, but needs positive reasons for thinking he is.

Faulkner thinks the hearer's attitude of affective trust in the speaker provides just such a reason, and thereby solves the problem. But before turning to Faulkner's "trust-based" solution, I will first sketch a paper from Pettit that I have found helpful in understanding Faulkner.

### **Trust-Responsiveness**

In 'The Cunning of Trust' Pettit first characterizes *interactive, trusting reliance*. A relies on B in this way to the extent that:

1. A relies on B to PHI
2. This reliance is manifest to B
3. A expects B to be well disposed and to attach a greater utility to PHI for the fact that it represents a way of proving reliable. (Pettit 1995: 205-6)

We rely on people in this interactive, trusting way all the time. Those who prove reliable are **trust-reliable**.

Pettit distinguishes two forms of trust-reliability: **trust-worthiness** and **trust-responsiveness**. Trust-worthiness implies that the trustee is antecedently disposed to prove reliable; the trust-worthy person has a trust-worthy *character*. Pettit identifies three forms of trust-worthiness: loyalty, virtue, and prudence. Each provides a reason for a trustor to interactively trust a trustee:

Suppose I believe that someone is a loving family member, a loyal friend, a devoted colleague, or whatever....Or suppose I believe that someone is virtuous; say, a god-fearing sort who can be relied upon to follow certain religious norms...Or suppose I believe that someone is a prudent sort who will see the potential long-term rewards of maintaining a certain relationship...that requires her to prove responsive to certain acts of reliance on my part....[By] manifesting the fact of relying on her...I can actually motivate her to perform accordingly. [These beliefs offer grounds] on which I may expect that if I manifest the fact that I am relying on that person to do something that person will be led to attach a greater utility to doing it. (1995: 208)

If I believe you are loyal to me, or if I believe you follow norms prescribing help, or if I believe that you can see the wisdom in helping, then I would have good reasons rationalizing my reliance on you.

Pettit then explains *trust-responsiveness*. He first notes that human beings, among other things, desire the good opinion of others. We value being loved, liked, acknowledged, respected, admired, and so on. He calls these “attitude-dependent” goods for they depend on being the object of someone else’s positive attitude.

Pettit then argues that our desire for the good opinion of others can give a trustor another reason to trust a trustee. Compactly stated, here’s his argument:

- (1) Suppose Andy would like Lindy to PHI.
- (2) Suppose Lindy desires the good opinion of Andy.
- (3) Suppose Lindy's doing PHI when Andy manifests reliance would earn Andy's good opinion.
- (4) Then Lindy has a motive to PHI when Andy manifests reliance.
- (5) Knowing or believing all of this, Andy would then have a reason to manifest reliance on Lindy.

How does the trustor make his reliance manifest so the trustee will be moved to prove-reliable? By also manifesting a *belief* or *presumption* that the trustee is trustworthy, or at least by manifesting a *disposition* to believe that the trustee is trustworthy should she prove trust-reliable. And to believe or presume that another is so trustworthy is to think well of that person, or at least to think well of that person in the event that he or she proves trust-reliable. By manifesting his reliance, Andy communicates that he believes or presumes Lindy "to be truly the sort of person who will not take advantage of someone who puts" himself at her mercy (1995: 214). This belief, presumption or disposition of the trustor then motivates the trustee to prove reliable. When the act of manifesting reliance interacts "this way with the desire of a good opinion, then the act of trust is likely to have an important motivating aspect for the trustee...It is a sort of bootstraps operation, wherein the trustor takes a risk and, by the very fact of taking that risk, shifts the odds in their own favor" (1995: 215-6). If I trust you in this way, then my trust incentivizes you to be one of the good guys. I then have a positive *prima facie* reason warranting my reliance.

### **The Straight Solution**

So suppose with Faulkner that a hearer's acceptance of a speaker's testimony must be "backed by reasons" (2011: 9-11). A common reaction among epistemologists to this is that the hearer must have a set of warranted *beliefs* that provide an argument from the premise that the speaker has asserted that P to the conclusion that the speaker's purpose is informative, and from there to the conclusion that (probably) P. This argument would overcome the standing defeater posed by the logic of the testimony game. Those beliefs

are then the reasons—or constitute the “having” of reasons—that provide the hearer’s warrant for her testimony-based belief (2011: 119).

Pettit’s paper then provides two different (though related) kinds of reasons a hearer might have that warrant uptake, reasons for thinking that the speaker, at least on this occasion, is one of the good guys. I could possess reasons for antecedently thinking you will prove reliable, reasons from loyalty, virtue or prudence. The solution to the problem of communication would then be straightforward. I would have reasons that “solve” the standing defeater posed by the “problem” of communication. I would have a reason that rationalizes my reliance that also engages your antecedently existing motives for proving reliable.

Or the hearer could possess a “trust-responsive” reason *à la* Pettit. Suppose I need information whether P and I think you might have it. I might choose to ask you whether P and even accept your answer in the belief or presumption that by manifesting my reliance on you I will motivate you to tell me the truth. I would believe that you are able to see that I depend on you proving reliable, and I would believe that you see my dependence as a reason to prove reliable, for you desire my good opinion. I would then believe I have tipped your preferences towards the cooperative outcome. I would then have a reason that rationalizes my reliance that provides you with a new motive for proving reliable.

Even though Faulkner says that both kinds of reasoning just sketched would “solve” the “problem,” and even though the latter mechanism of trust-responsiveness will help us understand Faulkner’s “trust-based” solution, Faulkner rejects both kinds for he thinks they are not *the central reason* we “trust” testimony.

[Requiring reasons like these] over-intellectualizes our relationship to testimony. We do not always base uptake on the belief that what is told is true, sometimes we merely trust a speaker for the truth. ... [These reasons miss] a central reason, arguably the central reason, why we trust testimony... An audience’s reason for the uptake of a speaker’s testimony can be no more than that the audience *believes the speaker, or trusts the speaker for the truth.* (2011: 175-6)

### **Affective vs. Predictive Trust**

To get your mind around this, you'll need to understand Faulkner's distinction between two kinds of trust. Following others, Faulkner distinguishes *affective* from *predictive* trust (2011: 144-147). Trust in both cases is a three-part relation: A trusts S to PHI.

A *predictively* trusts S to do PHI if and only if (1) A depends on S to PHI, and (2) A predicts (believes) that S will PHI. The first condition is factive; A must really depend on S to PHI. The second is not. It's the belief (the prediction) that S will PHI (2011: 145). Talk of trust in this sense "sits happily" with the "intellectual" reasons for uptake just canvassed, for "predictive trust is reasonable just when there are grounds for judging a cooperative outcome; and what makes the act of trusting reasonable is these grounds" (2011: 145).

A *affectively* trusts S to do PHI if and only if (1) A depends on S to PHI, and (2) A expects (has the normative expectation) that (1) will motivate S to PHI (where A [normatively] expects it *of S* that S be moved by the reason to PHI given by A's dependence on S) (2011: 146). Faulkner elaborates on (2):

The normative dimension of the expectation...is then that the trusted party *should be* trustworthy. Thus, in trusting S to PHI, A presumes that S *ought to PHI* and, other things being equal, that S will PHI for this reason. (2011: 147-8)

With this distinction in hand, we are on our way to seeing what Faulkner has in mind. I might *believe* that if I manifest my dependence on you for information whether P, then you will tell me the truth, where my belief is a *prediction* based on reasons or evidence. I would then have a positive reason for thinking you are one of the good guys that overcomes the standing defeater. My uptake is warranted provided I've got good reasons or evidence to support my prediction.

On the other hand, I might *presume* that if I manifest my dependence on you for information whether P that you *ought to* tell me the truth. According to Faulkner, this *presumption* constitutes, or is constituted by, affective trust, the *normative expectation* that you *should* tell the truth. Faulkner's main idea is that this expectation rationalizes reliance and thereby warrants uptake; it provides a "de-intellectualized" solution to the problem of communication.

This move, I believe, marks a significant development in the evolution of Faulkner's thinking about testimonial warrant. I recall his earliest papers using the "problem of communication" to justify a "positive reasons" requirement that predictive reasons would readily satisfy; he seemed to endorse a "straight" solution (Faulkner 2000, 2002). Now he's turned to *trust* as a new basis for warranted uptake. Faulkner's main thesis in *Knowledge on Trust* is that this presumption—this normative expectation—warrants uptake and thereby solves the problem of communication. But what exactly is this presumption, and why does it "solve" the problem? We still have a good amount of unpacking to do to understand Faulkner's "trust-based" solution.

### **The Presumption of Trustworthiness**

Faulkner argues that when the audience A believes that speaker S can see that A is relying on S for information whether P, and in addition A affectively trusts S for that information, then A will make a number of *presumptions*. These presumptions do a good deal of the same work explicit beliefs did in the account inspired by Pettit. As Faulkner sees it, the psychology of the trusting audience goes like this:

1. A *believes* that S recognizes his, A's, trusting dependence on S proving informative.
2. A *presumes* that if S recognizes A's trusting dependence, then S will recognize that A normatively expects S to prove informative.
3. A *presumes* that if S recognizes A's expectation that S *should* prove informative, then other things being equal, S *will* prove informative for this reason.
4. So taking the attitude of affective trust involves *presuming* that the trusted *will* prove trustworthy. (2011: 130)

According to Faulkner, these "presumptions" are not, or "need not amount to," beliefs (2011: 154).

The audience A's reason for believing that p, when this is what S tells him, is not the belief that S will prove to be trust-responsive, it is simply the fact that S told him that p (and A trusts S for truth on this matter). S's telling is seen to provide a reason...because in trusting S for the truth, A

accepts certain things about S and the testimonial situation which yield the presumption that S is trustworthy. (2011: 164)

How do these presumptions “solve” the problem of communication? According to Faulkner, the hearer’s *presumption* that the speaker will prove informative *rationalizes* the hearer’s uptake of the speaker testimony in the same way a hearer’s *belief* that the speaker will prove informative rationalizes uptake. These presumptions then do a good deal of the same work explicit beliefs did in Pettit’s model.

The presumption [that the speaker will prove trustworthy] like the belief with the same content, *makes it probable for A* that p is true given that this is what S tells him. So A’s attitude of trust raises the probability of p [for A], given this is what S tells him. So A’s trusting S for the truth...provides A with an epistemic reason to believe. (2011: 154, emphasis added)

Besides *rationalizing* A’s uptake, does the presumption that the speaker should prove reliable also make A’s uptake *objectively more likely to be true*? It does if, in fact, A’s attitude of trust effectively *moves* the speaker to be informative, or moves the speaker to be more likely to be informative. And, as in Pettit’s analogous model, it often does. Faulkner claims, citing Pettit approvingly, that A’s trust gives S “a reason to be trustworthy...acts of trust can create as well as sustain trusting relations” (2011: 156-7). And so Faulkner thinks that A’s trust not only makes it more probable *for A* that S is apt to prove informative in utterance and so rationalizes A’s uptake, it also *motivates* the speaker to prove trustworthy and thereby raises the objective probability that S will prove informative in utterance; A’s trust is often an *effective* reason—a reason that *motivates* the speaker to prove informative, and thereby objectively warrants an audience’s uptake. We now have a “de-intellectualized” version of Pettit’s “intellectualized” trust-responsive reason for manifesting reliance that fits Faulkner’s requirement that uptake be backed by reasons (2011: 57, 160, 167).<sup>1</sup>

---

<sup>1</sup> Unfortunately Faulkner does not explicate presumptions or show why they are not, or need not be, beliefs. All he says is that when we presume these propositions, we are “not explicitly committing to these propositions in [our] reasoning” (2011: 151). Instead he says that presuming these propositions “partly defines how it is that the attitude of affective trust involves seeing things in a certain light...in the positive light of trust” (2011: 152, 154).

But if a hearer must have the concepts involved to have the presumptions, and so be able to think the presumptions in order to have them, exactly why is this account less “intellectualist” than the analogous solution from Pettit? Granted Faulkner’s account isn’t a goodwill account and doesn’t require

All of this, however, is only a part of Faulkner's picture. For ask yourself the following: why should the *hearer's* normative expectation that the speaker *should* prove reliable actually *motivate* the speaker *to be* reliable? Why should affective trust move the speaker to be one of the good guys? In Pettit's paper the analogous question is this: why should the fact that the trustor has manifested reliance on the trustee motivate the trustee to prove reliable? And in Pettit's paper, the answer is that the trustee desires the positive approval of the trustor. But Faulkner doesn't emphasize—and in fact at times de-emphasizes—the speaker's desire for approval from the hearer as a reason for the speaker to prove reliable.

So what is Faulkner's analogous answer to our question? His main answer—as I understand it—is this: the hearer's normative expectation that the speaker should choose the cooperative outcome is just the hearer's *internalization of the social norm* that speakers should prove trustworthy, *and* the fact that the social norm moves the speaker to choose the cooperative outcome for the speaker has internalized the norm as well. The hearer's trust—the hearer's normative expectation, which *rationalizes* uptake—then “engages,” so to speak, the speaker's internalization of the norm, which thereby *motivates* the speaker to choose the informative outcome.

By my lights, we now have another major development in Faulkner's thinking. We've moved from predictive trust to affective trust, and now from affective trust to social norms. To explain this last development, I need say a few words about social norms and their internalization.

### **Internalized Social Norms**

Social norms are an important species of social institution along with conventions, customs, and laws; they are causal structures with explanatory force. Here are some examples. Different populations and subgroups obviously regularly dress in various ways. In business contexts most people dress business professional. In Muslim societies, most women dress head to toe. But in Western societies women usually wear considerably less. The regularities in behavior are there to see. And these regularities are clearly approved. Business people approve of business professional attire; they disapprove of those who fail to dress for success. Muslim men and women clearly

---

“explicit reasoning” through these presumptions, it doesn't seem to me that Faulkner requires fewer psychological capacities for warrant than Pettit would.

prescribe complete coverage for women, whereas Western men and women care considerably less. There are countless social norms governing human life, varying in a number of ways, from group to group. The existence of social norms, like language, is a human universal (Brown 1991). Some are smart—send kids to school—some are silly—men should wear ties—and some are downright stupid—circumcise your daughters. Though universal, they differ widely. Think of norms governing what foods to eat, and when and where to eat them. But though they differ, at a higher level of abstraction human cultures prescribe and proscribe a lot of the same kinds of behavior. Most prohibit killing, assault, and incest. Most promote sharing, reciprocation, and helping.

A number of disciplines have taken up social norms: sociology, psychology, economics, anthropology, etc. Sociologists and anthropologists take up the whole range of social norms. Economics and political science, on the other hand, have mostly taken up norms that resolve free-rider problems, and in particular free-rider problems that are also many-party prisoner's dilemmas. Something that ensures universal—or even fairly general—cooperation is then a “solution” to the many-party dilemma; everyone is better off if nearly everyone cooperates. Certain social norms are then “solutions” to these dilemmas (Pettit 1990).

Abstractly characterized, social norms are (1) regularities in behavior in a group or population (so they are “norms” in the scientist's sense of what usually or “normally” happens), that (2) are prescribed by members of the population, so that they approve of conformity and disapprove of deviance (so they are “norms” in the philosopher's sense of what's prescribed or what ought to occur), and that (3) are regularities in part because they are prescribed (so that they are “normative” in the moral psychologist's sense of motivating behavior) (Pettit 1990, Miller 2001). These are general claims about most people in the population: most follow the norm; most prescribe the norm; and most follow the norm partly because prescribed.

Social norms are experienced as things we *ought* to do, that it would be *wrong* not to do. And so when we prescribe conformity, we believe each of us *ought* to conform, and when someone fails to conform we experience more than mere disapproval, but a kind of *moral* disapproval (Miller 2001: 139; cf. Pettit 1990). Think of the norm forbidding eating pork in certain communities. Sure, eating pork will win

disapproval, but many who conform believe they *shouldn't* eat pork, that eating pork is *wrong*, no matter what other people might think.

In other words, we *internalize* norms. We internalize norms through socialization, from our parents and others. When we internalize a norm, we find it intrinsically motivating; our preferences change. We conform because we think it's the *right* thing to do, because we are *supposed* to do it. We *want* to do it. We may even deeply value compliance. Many internalized norms even come to be partly “constitutive of the selfhood or identity of individual adherents” (Miller 2001: 139). Internalization then leads to compliance as an ultimate end, and not just as a means to avoid punishment or societal disapproval (Bowles & Gintis 2003: 13-14, 2011: 169). When internalized, reward and punishment may drop away, or only play a sustaining role. When internalized, I conformed to the norm *because it's the right thing to do*, because I positively value compliance, not (normally or just) because of the consequences of my actions or because of my other aims or desires. Internalized norms are then experienced as categorical, as what must be done, and not simply as what we should do given other aims or desires (Sripada & Stich 2006). Given that I've internalized a norm, I expect it of myself and others.

Internalized norms are intimately connected with the social emotions: guilt, shame, embarrassment, love, envy, pride and resentment. These emotions make a huge part of the proximate psychological mechanisms driving positive and negative evaluations of compliance, evaluations that motivate compliance; strong social emotions are a central proximate psychological mechanism ensuring conformity (Frank 1987; Ekman 1992). Even thinking about failure may evoke strong feelings of embarrassment, anxiety, guilt or shame (Gintis 2003). You might even feel sick at the very thought of breaking a norm. If I fall short of my normative expectation of myself, I will feel guilty or ashamed. If you fall short, I will feel contempt or resentment. “Social norms have a grip on the mind that is due to the strong emotions they can trigger” (Elster 1999: 100).

### **The Ultimatum Game**

Faulkner thinks we've internalized social norm *tell the truth informatively*. And he thinks that fact somehow “solves” the problem of communication, for it rationalizes uptake and motivates trustworthiness. To explain how this works, I'll first explain how

internalized norms drive behavior in another game studied by rational choice theorists, economists, anthropologists and evolutionary psychologists: the ultimatum game.

In the ultimatum game, there are two players, the proposer and the responder. The proposer is given some money, e.g. \$100. The proposer then must propose a split of the money between the proposer and the responder, anywhere between \$1 for the responder and \$99 for the proposer and \$99 for the responder and \$1 for the proposer. The responder's job is to accept or refuse the split. If accepted, both parties receive the amount in the proposal. If refused, no one gets anything. And so both parties are better off if the responder accepts the split. These games are often played anonymously, where both parties know they won't play another round with the same person, and so tit-for-tat, among other strategies for playing repeated games, isn't an option.

According to a simple rational choice prediction where people rationally act on their preferences—especially their economic ones—the proposer should propose \$1 for the responder and \$99 for the proposer, for the proposer will know that the responder will accept the split, for the responder will reason that \$1 is better than nothing, and so it is better to accept such a split than refuse it.

Surprisingly, however, across a very wide-variety of human cultures, that is not what happens. Instead the proposer tends to propose something more like \$40 for the responder and \$60 for the proposer (the split on average varies from culture to culture), and when the proposer proposes a much smaller split for the responder, the responder tends to refuse them, which in part would explain why bigger splits are usually proposed, for the proposer usually knows which splits are apt to be refused, and which are apt to be accepted.

A now standard explanation of this behavior—which seemingly violates “rational choice axioms” of human behavior—is that the participants—and so humans generally—have internalized social norms of fair-divisions of goods (Henrich et al, 2004). Proposers know that they are supposed to propose a fair split, and responders know that too, and furthermore responders get upset and enforce the norm by punishing the proposer by rejecting obviously unfair splits (and thereby also harm themselves, for they lose out on the money offered in the split).

Describe the game to your colleagues, students, or loved ones, and you'll elicit similar reactions. Sure, some of them might take a low offer when playing responder,

but many are bound to feel punitive attitudes—to get upset—at “unfair” offers. They “know” the offers are *supposed* to be fair, that the proposer *should* offer a more even division of the goods. Because of internalized norms, we normatively expect fair divisions, and so fair divisions reliably occur.

If such “fairness” norms have been internalized, they explain why people behave as they do when playing the ultimatum game. We then have an example of where internalized social norms explain behavior that otherwise would have seemed very puzzling, at least from the point of view of rational choice theory. We have an example of behavior that goes against what rational choice theory would have predicted in the first place. Participants “see” what they are “supposed” to do in such a situation in terms of internalized norms, and those internalized norms then explain, in large part, what they do.

Now when playing the ultimatum game it’s often the proposer’s recognition that the responder expects a fair division that at least partly moves the proposer to offer a fair split. But just as often the proposer herself thinks she should offer a fair split, for she has internalized the norm, and believes anything less would be wrong. She might even, and often does, propose a fair split in games where she knows the game is anonymous and won’t be repeated. She “intrinsically values” fair divisions of the goods.

Social norms also enter explanations of actual behavior in prisoner’s dilemmas. Given a rational choice model, you’d expect everyone to defect, especially in non-repeated, anonymous games. But in fact a substantial number of people, even in non-repeated, anonymous games, choose to cooperate, and they often do so because they believe it’s the right thing to do.

### **The “Trust-Based” Solution**

Now recall our question for Faulkner: What is it about the *hearer’s* normative expectation that the speaker should prove reliable that motivates *the speaker* to be reliable? And recall our answer: the hearer’s normative expectation is the hearer’s *internalization of the social norm* that speakers should prove trustworthy, and if the speaker has internalized that norm, then the speaker’s recognition of the hearer’s reliance shall move the speaker to prove reliable. The *hearer’s* internalization of the norm then “engages” the *speaker’s* internalization of the norm, which motivates the

speaker to choose the informative outcome. The speaker is then one of the good guys, for he prefers helping. We're now in a position to say what all of that means.

Faulkner holds that when we play the testimony game we follow the internalized social norm *tell the truth informatively*.

We expect interlocutors to live up to [a standard] when...having a certain type of conversation. On this standard if another depends on you for information, then you should try to say what is true informatively. (2011: 181)

Once internalized the norm shapes our moral psychology. It shapes how we "see" the situation, our behavior, and our reasoning. Faulkner says this is a...

...quasi-perceptual matter because where the norms of trust are internalized, the subject's perception ...[of]...the Testimony Game will be structured by the prescriptions of the norms. This situation will be *seen in a certain light*. Thus, *if S can see that A depends on her telling the truth as to whether P, this will be seen by S as a reason to do so...*(2011: 185, emphasis added).

Speakers who have internalized these norms—and so intrinsically value compliance—will then often enough choose the informative outcome when they see that audiences need information; they will be "motivated to conform" because they have "internalized the norm" and so "intrinsically value" compliance (2011: 186). Speakers are then committed to telling the truth; their reasons for utterance will then often put audience's informational needs first (cp. Faulkner 2011: 5-6, 132). And so when playing the testimony game, audiences that manifest reliance will receive the cooperative outcome, for speakers will be motivated by manifest need, for they intrinsically value proving informative.

We can then explain the matter of fact reliability of our institution of testimony:

Our ability...to see testimony in the light of trust comes down the fact that we have internalized social norms of trust. ...It is these...*social norms that shape the motivations we have, as speakers, in giving testimony, and which determine that testimony is the reliable source of knowledge that it is*. (2011: 172, 51, emphasis added)

By Faulkner's lights, we have the materials in place for a "genuinely philosophical theory of testimony," for that requires an "explanation as to why testimony can be presumed to be reliable and reliable in the way that it is" (2011: 51).

This explanation then "solves" the "problem" of cooperation:

[T]he *solution* to the problem [of cooperation in the case of communication comes from our] social institution of testimony defined by the existence of social norms. (2011: 169, emphasis added)

We can now see why uptake is warranted. It's warranted because *speakers* have among their preferences the pro-social, cooperative, helping preference to prove informative in communication.

This not only solves the problem, it "dissolves" the problem. Recall how the problem was set-up. When playing the testimony game, hearers want true information. But speakers supposedly would prefer the liberty to speak as they please; they are not committed to telling the truth; their reasons for utterance do not include the audience's informational needs. It then looks like the hearer needs some reason for thinking the speaker has chosen the informative outcome before it is rational for the hearer to rely on the speaker's testimony; the hearer needs some reason for thinking this speaker is one of the good guys. But if we've internalized social norms of trustworthiness, there's no problem, for it's not true that speakers are just as apt to prove uninformative as informative. It's not true that they're out simply to influence our beliefs. Rather they are out to inform, to prove helpful; speakers *intrinsically value* informative outcomes; they are committed to informative outcomes. We're then not in a world where, from the point of view of the logic of communication, the world might as well be equally divided between the good guys and the bad guys where we need evidence telling them apart; rather we're in a world of good guys.

...the trust-based solution...*dissolves* the problem of cooperation through showing how it rests on a restricted conception of what reasons we can have for ...being trustworthy....*it shows how trust need not be conceived of as problematic...*[when] norms of trust and trustworthiness are internalized. Then the Testimony Game would simply be perceived by interlocutors as a situation wherein each had a reason to trust and be trustworthy....*the trust-based solution...is not a strategy...for removing a background of distrust...*Rather, the claim that we intrinsically value

trust and trustworthiness is part of a philosophical explanation as to why such distrust is not pervasive. (2011: 199, emphasis added)

Like the ultimatum game, when we play the testimony game we don't simply have the kinds of motivations rational choice theory would predict. Behavior that had seemed problematic in the first place isn't so problematic after all. The existence of internalized social norms "solves" the "problem" of communication.

We should now be able to appreciate exactly why, on Faulkner's view, the hearer's attitude of affective trust warrants uptake:

...ultimately what determines that the attitude of trust provides an epistemic reason is *the existence of the causal structures* [the social norms prescribing truth telling and trusting] that *ensure* the attitude of trust can be potential evidence [in the sense that given the attitude, it is objectively more likely that what the speaker says is true]...Thus, ultimately, the attitude of trust provides an epistemic reason because there are norms of conversational trust, shaping the nature of the reasons we have, for utterance and for belief, in conversations as to the facts. (2011: 159, emphasis added)

The "ultimate," fundamental basis for the solution to the problem of communication—and so the ultimate, fundamental basis for warranted uptake—rests on the internalization of social norms.

There is much to admire in Faulkner's "dissolution." As Faulkner is aware, and as the discussion of the ultimatum game conveys, the appeal to social norms is one of the mechanisms social scientists now appeal to when explaining why humans cooperate—why they help one another—to the extent that they do, especially when a prediction based on narrow self-interest would have predicted something else instead. Faulkner's application of the apparatus of social norms to the case of informative cooperation is then an instance of a more general strategy for explaining why humans are as "pro-social" in their behavior as they are. I've pursued the very same strategy in some of my own work on testimonial warrant (see Graham, forthcoming, in preparation; cp. Graham 2010).

### **Pulling the Rug Out**

I shall make one major criticism. Given the way I understand it, to dissolve the problem *speakers* must have the hearer's interests in getting the truth as one of their interests. If *speakers* don't have that interest, then we have a problem. If they do have that interest, then we don't. If *speakers* have internalized the social norm *tell the truth informatively*, then they intrinsically value the informative, cooperative outcome, and then have providing truth to others as one of their interests; they then have "pro-social, cooperative" preferences. So suppose, with Faulkner, that we have internalized that norm. Then when playing the role of speaker in a testimony game, we'll have among our preferences telling the truth informatively. Then when the hearer asks whether P, we will provide the truth informatively, insofar as we are able. We've thereby "dissolved" the problem.

I have identified two major roles the hearer's affective trust—the hearer's normative expectation that the speaker should prove reliable—plays in Faulkner's account. The first is *rationalizing*. The hearer's belief that the speaker can see the hearer's reliance and can recognize the hearer's expectation, plus the nested presumptions that the speaker will then prove reliable, all rationalize or render reasonable for the hearer the hearer's uptake of the speaker's testimony; it overcomes a standing defeater. The second is motivational. Like Pettit's mechanism of trust-responsiveness, the hearer's normative expectation of the speaker can play a motivating role; it can motivate the speaker to prove reliable. It then renders objectively probable the hearer's uptake.

But if we have *dissolved* the problem through the speakers internalizing the norm *tell the truth informatively*, why must the hearer also internalize the norm? Why is the hearer's internalization of the norm fundamental in explaining warranted uptake?

Consider the rationalizing role. The demand for reasons in each and every case was motivated by thoughts about speaker preferences: speakers act on their self-interested preference to influence the hearer's beliefs, not on any other-regarding preference. But now we know that's not true, for we know speakers have other-regarding, pro-social preferences as well, for they internalize the norm *tell the truth informatively*, and thereby intrinsically value proving informative. They are committed to telling the truth and have the audience's needs among their reasons. So the basis for

the “problem” of communication has been removed, thereby “dissolving” the problem. Sure, we may need reasons to rebut occasion specific reasons for thinking this or that speaker might actually lie or mislead, but we don’t need a reason in each and every case to rebut a standing defeater, for the very reason for thinking there was such a defeater falsely assumed speakers only care about influencing what hearers believe and never care about what hearers want when playing the testimony game. Once speakers internalize the norm, we’ve removed a basis for “distrust.” Once dissolved, Faulkner’s motivation for requiring reasons for warranted uptake goes by the board.

Consider next the motivating role. If speakers have internalized the norm, then they are already motivated to prove informative in communication. They don’t need, at least not in each and every case, further motivation from the hearer’s normative expectation that they should prove reliable to be reliable. As Faulkner says, it’s the fact that speakers have internalized the norm that explains the reliability of testimony. If they haven’t, then we’re in trouble. But if they have, hearers needn’t worry, at least not in each and every case. And so they needn’t normatively expect speakers to be reliable in every case to motivate them to be reliable. Once we’ve shifted from something like Pettit’s mechanism—where something in the trustor explains the trustee’s motivation—to internalized social norms as the mechanism motivating trustworthy behavior, then the need for something like Pettit’s mechanism in each and every case drops away for warranted uptake, uptake not backed by other reasons either. So when Faulkner turns to internalized norms to explain why the hearer’s attitude motivates the speaker, he’s effectively shifted away from a reason that *the hearer* must possess to explain warranted uptake to a motive *the speaker* must possess to explain warranted uptake.

In sum, Faulkner has shifted from the hearer’s attitude as fundamental to the speaker’s motivation as fundamental; supposedly the hearer’s attitude did a lot of the work, but now we can see that the speaker’s motive does all the heavy lifting. This shift dissolves the basis for both the rationalizing and motivating roles of the hearer’s normative expectation. As far as I can tell, Faulkner has pulled the rug out from beneath his feet.

### **The Problem of Parental Care**

We even know such a rationalizing and motivating reason based in the recipient isn't required to motivate pro-social, helpful behavior.

Consider parental investment in offspring. As any parent knows, parents invest an enormous amount of resources in the care of their children, even when the children are unlikely to repay any of it. From the point of view of narrow self-interest, parental investment in their offspring can seem puzzling. So why do they do it?

Well, again as every parent knows, parents deeply care about the well-being of their children. Parents have "internalized" the well-being of their offspring; they have very strong preferences in favor of helping their children. That in part explains why parents invest so much.

Now what do children have to do to get their parents to invest so much care? Nothing really. They just have to need the help. They certainly don't have to manifest a belief, presumption, or disposition that if their parents care for them, they will in turn pay them back, or confer some other benefit. They certainly don't have to manifest the normative expectation that their parents *should* invest so much in them, and they certainly don't need a reason for thinking that their parents will help them. Newborns, infants, toddlers and very young children are not even psychologically capable of possessing these mental states and attitudes.

And so here we have a case of helping behavior explained by helping preferences. The children don't need to normatively expect anything from their parents. And so the "solution" to the "problem of parental investment" does not require some "rationalizing" and "motivating" attitude that "warrants" children's "uptake." The solution consists in parents caring for their children.

And so when it comes to parents talking to their children, the children obviously do not need to motivate their parents, or rationalize their trust, via a normative expectation that their parents should prove reliable. Parental testimony is an existence proof of the possibility of testimonial warrant and knowledge in virtue of the speaker's motivations without a corresponding normative expectation on the part of the recipient. Parents regularly and reliably provide true information to their children because they care about the well-being of their children. Relatives, teachers, and other caregivers too. Even complete strangers often provide true information to children. But children—

especially very young children—don't possess reasons that either motivate or rationalize uptake. At best they sometimes have grounds for suspending judgment (Graham 2010). They certainly don't possess the normatively loaded attitudes Faulkner has in mind. Nonetheless, children learn (come to know) a great deal by believing what their parents, relatives, teachers, caregivers and even strangers tell them. Children thus don't need to internalize social norms of trust and truth telling to rationalize their reliance or motivate other people to tell them the truth. If parental concern can explain childhood knowledge through testimony without a corresponding normative expectation on the child's part, why can't a speaker's concern explain adult knowledge through testimony without a corresponding normative expectation on the hearer's part?

On the last two pages of his book Faulkner addresses the problem posed by children. In total, he devotes less than a page to the problem. And what he says strikes me as unsatisfactory.

Firstly, he says that children are not entirely unsophisticated, in that in communication they are sensitive, to at least some degree, to the speaker's communicative intention (which, on standard accounts of linguistic comprehension, is required to understand speaker meaning) (2011: 203, n. 3). But sensitivity to communicative intentions is not the same as affectively trusting a speaker for the truth; at best it's a prerequisite for that. He does not say children are capable of affectively trusting speakers in his sense. He does not say they are capable of believing that the speaker can recognize their dependence. Nor does he say they are capable of presuming that the speaker *should* prove trustworthy in communication. And so he does not say that children possess the reason he is so at pains to locate as *the central basis* for our rational reliance on testimony. Furthermore, he does not say that children possess either kind of predictive reason that Pettit isolates.

Secondly, even if children are sensitive to the speaker's communicative intentions, Faulkner says nothing as to why that should warrant uptake. Indeed, if it did, Faulkner's own account would be otiose, as every single hearer must recognize communicative intentions in the first place to understand the speaker, and then every single hearer would at least enjoy a *prima facie, pro tanto* warrant via understanding. Furthermore, Faulkner seems to repeatedly reject this position, as sensitivity to communicative intentions is already present in the testimony game. Mere sensitivity to

communicative intentions, Faulkner thinks, isn't enough to warrant uptake. Faulkner barely mentions what children can do, and he says nothing about why it would be adequate. And the one thing he does say they can do he rejects as inadequate.

Thirdly, he then makes a familiar move when pressed by such cases: deny the data. If children are not sophisticated enough to have the reasons Faulkner requires, then, Faulkner claims, their testimony-based beliefs really aren't warranted, and they don't really acquire knowledge from testimony, no matter how informative their parents and caregivers happen to be. But where is it written that children cannot acquire knowledge from their parents and caregivers unless they've matured enough to have the kinds of attitudes Faulkner requires? "There is a vague domain here," he writes, "and the objection should not illicitly exploit it" (2011: 203). The vague domain is the point at which they go from unsophisticated in the relevant domain to sophisticated. The vague point instead, Faulkner would have us believe, is when they go from ignorant recipient of highly reliable information to knowing recipients of that very same information. The objection does not illicitly exploit the first. Rather, Faulkner (at least it seems to me) illicitly conflates the former with the latter.

The case of childhood testimony is then an existence proof that the speaker's motivations are often sufficient for testimonial knowledge and warrant, that a hearer does not always need positive supporting reasons to warrant uptake. Pro-social, helping preferences are often sufficient to explain warranted uptake; reasons for uptake are not required in each and every case.

### **Why not Reliabilism?**

My objection to Faulkner is that I don't see why we, qua hearers, must internalize the norm (though clearly, in general, we have). And so even though obviously present, I don't see that Faulkner has explained why it's fundamental, when it seems, on his account, that it's the speaker's internalization of the norm that is both necessary and sufficient. It can't be that, qua hearers, we need a reason to think that the speaker has chosen the informative outcome, for once the norms are in place, the problem of communication has been dissolved, and so internalization qua hearer isn't required to overcome a standing defeater; it is simply not true that speakers are on the fence.

On the other hand, I don't deny that we, qua hearers, have internalized the norm, and I don't deny that, given internalization, it makes sense "from the inside" for hearers to rely on speakers to prove informative, and even to normatively expect speakers to prove reliable. What I deny is that Faulkner has established the requirement that, in every case, the hearer must have internalized such a norm, and so that the hearer must, in every case, presume that the speaker ought to prove reliable and so will be reliable, in order for the hearer to enjoy a *prima facie, pro tanto* warrant to rely on the speaker.

Turning to the issue of motivation, it can't be that, qua hearers, we need to internalize the norm in every case to motivate others, qua speakers, for, qua speakers, we will choose the cooperative outcome, provided we have internalized the norm; the norm in speakers, as it were, doesn't get turned on when and only when they recognize the internalization in hearers; it gets turned on when they recognize that others need information. A trust-responsive mechanism isn't required, for speakers are antecedently motivated to prove informative.

For instance, I might think you need information whether P and provide it. You might, however, not need it, or not care whether I provide it. You might even suffer from brain damage that left you intelligent but affectless, so that you no longer experience the emotions characteristic of internalized norms. As a result you could care less whether anyone provides true information informatively. I might even know that you don't care. Even so, given I've internalized the norm, I still tell you the truth informatively; it's the right thing to do.

So once we've internalized the norm *tell the truth informatively, qua speakers*, why must we internalize the norm, *qua hearers*, to either motivate the speaker or rationalize our reliance, or both? It's not required to motivate speaker compliance, and it's not required to rationalize reliance, even if it does. And so why not choose the reliabilist, externalist account where hearers are *prima facie, pro tanto* warranted in believing what they are told because testimony is reliable, reliable because, qua speakers, we've internalized the norm *tell the truth informatively*?

Faulkner's own thinking has evolved from a straight solution to something modeled on Pettit's reasoning to the social norms account. As I've argued, once he's moved to the social norms account, he's pulled the rug out under his feet; he no longer

has a basis for insisting that the hearer must have, in each and every case, a reason warranting uptake. Speaker reliability, grounded in pro-social preferences, is enough.<sup>2</sup>

**Peter J. Graham**

*University of California, Riverside*

## References

- Bowles, S. & H. Gintis (2003) 'The Origins of Human Cooperation' P. Hammerstein, ed. *The Genetic and Cultural Origins of Cooperation* (The MIT Press).
- Brown, D. (1991) *Human Universals* (Philadelphia: Temple University Press).
- Ekman, P. (1992) *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage* (New York: W.W. Norton).
- Elster, J. (1999) *Alchemies of the Mind* (Cambridge University Press).
- Faulkner, P. (2000) 'The Social Character of Testimonial Knowledge' *The Journal of Philosophy* 97.
- Faulkner, P. (2002) 'The Rationality of Our Reliance on Testimony' *Synthese* 131.
- Faulkner, P. (2011) *Knowledge on Trust* (Oxford University Press).
- Frank, R. (1987) *Passions within Reason* (New York: W.W. Norton).
- Gintis, H. (2003) 'Solving the Puzzle of Human Prosociality' *Rationality and Society* 15: 155-187.
- Graham, P. J. (2010) 'Testimonial Entitlement and the Function of Comprehension' A. Haddock, A. Millar, D. Pritchards, eds., *Social Epistemology* (Oxford University Press).

---

<sup>2</sup> I have benefited from comments from Coleen Macnamara and Zachary Bachman.

- Graham, P.J. (forthcoming) 'Epistemic Normativity and Social Norms' D. Henderson & J. Greco, eds, *Epistemic Evaluation* (Oxford University Press).
- Graham, P.J. (in preparation) 'The Reliability of Testimony and Social Norms'.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr & H. Gintis (2004) *Foundations of Human Sociality* (Oxford University Press).
- Miller, S. (2001) *Social Action* (Cambridge University Press).
- Pettit, P. (1990) 'Virtus Normativa: Rational Choice Perspectives' *Ethics* 100: 725-755.
- Pettit, P. (1995) 'The Cunning of Trust' *Philosophy and Public Affairs* 24: 202-225.
- Sripada, C. & S. Stich (2006) 'A Framework for the Psychology of Norms' P. Carruthers, S. Laurence, & S. Stich, eds., *The Innate Mind*, Volume 2 (Oxford University Press).