

## *Testimonial knowledge through unsafe testimony*

SANFORD GOLDBERG

Frank is a writer with a strange habit. Every morning, at precisely 7:30 a.m., he wakes up and dumps out whatever is left of the pint of milk he purchased the day before, but places the empty carton back in the fridge until noon. Then, throughout the interval from 7:30 to noon, he always remains in the kitchen, as that is where he writes every morning like clockwork. Finally, at exactly noon, he takes the now-empty milk carton out of the fridge and throws it away – an act which to him symbolizes the end of his day's writing. Now Mary is unaware of Frank's milk-dumping practice. One morning, having spent the prior evening at Frank's house with Frank and her son Sonny, she awakens at 7:40 and goes to the kitchen with Sonny. Upon entering (Frank is already there) she immediately goes to the fridge for a glass of OJ, and as she reaches for the OJ she casually observes a small carton of milk. She goes on to tell Sonny (who always has cereal with milk for breakfast) that there is milk in the fridge. As luck would have it, there is indeed milk in the carton on this day (Frank failed to remember that he had bought milk yesterday). When Frank observes Mary's testimony, he realizes that he forgot to dump the milk; when Sonny observes her testimony, he forms the belief that there is milk in the fridge.

In what follows I want to argue that the following two claims hold:

- (1) Mary's testimony (to the effect that there is milk in the fridge) is unreliable, insensitive, and unsafe.
- (2) Sonny comes to know through Mary's testimony that there is milk in the fridge.

Like the examples offered in Lackey 1999 and Graham 2000, I offer this example as a case in which testimony *generates* knowledge that the source herself lacks. Unlike the examples offered in Lackey 1999 and Graham 2000, however, in the present case knowledge is generated through testimony which itself is unreliable, insensitive, and unsafe.

I begin with (1). Mary's testimony is unreliable: it is not the case that she would have testified that there is milk in the fridge only if there had been milk in the fridge. This is because the basis of her testimony is her visual experience of the carton of milk in the fridge. But the carton is not transparent; she did not actually see milk in the container. Rather, she just assumed that it contained milk. Given Frank's habit, she would have had

precisely the same evidence, and would have made precisely the same assumption, even in a situation in which there was only an *empty* milk carton in the fridge (owing to Frank's having performed his habitual act). In that case, she would have testified that there is milk in the fridge, in a situation in which there wasn't any milk in the fridge. Since that scenario occupies a nearby world, we can conclude that Mary's testimony is unreliable. Indeed, we can also conclude that her testimony is insensitive: even if there hadn't been any milk in the fridge she would have testified that there was. And we can conclude that her testimony is unsafe as well: it could easily have been the case that there was no milk in the fridge but she testified that there was. (The same points which I just made regarding Mary's testimony hold, *mutatis mutandis*, for Mary's belief in what she attested.)

I take (1) to be uncontroversial; more controversial will be (2). For the hypothesis that unsafe testimony can give rise to testimonial *knowledge* seems as unlikely as the hypothesis that one can have *unsafe knowledge* – which, arguably, is something that we can know a priori to be false. But far from being knowably false a priori, the claim that there can be testimonial knowledge through unsafe testimony is true: it is instantiated in the case of Mary and Sonny. After arguing for (2), I will defend the argument against several objections.

I begin with a *prima facie* case for the claim that, while *the testimony* Sonny consumed was neither reliable nor sensitive nor safe, *Sonny's belief* in what was attested is reliable, sensitive and safe. In particular, the following counterfactuals are all true of his belief:

- (Rel) In situations similar to the present one, Sonny would form the testimonial belief that there is milk in the fridge only if there were milk in the fridge.
- (Sen) If there hadn't been milk in the fridge, Sonny wouldn't have formed the testimonial belief that there is milk in the fridge.
- (Saf) It couldn't have easily been the case that Sonny formed the testimonial belief that there is milk in the fridge, and yet there is no milk in the fridge.

Had there been no milk in the fridge, this would have been because Frank dumped it (and put the empty milk carton back in the fridge). As noted above, in such a situation Mary would still have testified as she did; but Frank (who is a fixture in the kitchen, and so who is in the kitchen in most or all of the nearest worlds) would have immediately spoken up against that testimony, informing his uninitiated guests of his strange practice. In that case Sonny would not have consumed Mary's testimony, and so would have *refrained* from forming the testimonial belief that there was milk in the fridge. This establishes that Sonny's testimonial belief is

*sensitive*. Now, had Sonny formed the testimonial belief that there is milk in the fridge, this would have been a case in which Frank did *not* speak up against that testimony; but, given Frank's scrupulousness, the only cases in which he would not speak up against that testimony (given that he was in the kitchen, as always) would be those cases, like the actual one, in which (upon hearing the testimony) he came to acknowledge that he failed to dump the milk from the previous evening. In all such cases, there *would* be milk in the fridge. In sum, had Sonny formed the testimonial belief that there is milk in the fridge, there would have been milk in the fridge: Sonny's testimonial belief is *reliable*. Note, too, that, any nearby world in which (a) Frank disposed of the milk and returned the empty carton to the fridge, yet (b) Mary – or someone else, for that matter – testified (on the basis of seeing the milk carton in the fridge) that there was milk in the fridge, will be a world in which Frank speaks up against that testimony, prompting Sonny to refrain from consuming that testimony. Sonny's belief is *safe*.

The foregoing is a first-pass attempt to establish that Sonny's belief is reliable, sensitive and safe. Below I will suggest how to extend this result to reach the conclusion that Sonny's belief amounts to knowledge. But before I do so, I want to address two ways in which my claim regarding the reliability (etc.) of the belief can be resisted.<sup>1</sup> The first grants the conditional thesis that Sonny's belief is reliable (etc.) *if* Frank's presence in the kitchen is held fixed, but denies that the issue of reliability should be settled by holding Frank's presence fixed. The second simply denies that Frank's presence *does* render Sonny's belief reliable, on the grounds that Frank's presence is not causally linked to the occurrence of Mary's testimony. In addressing both of these objections, I will be considering two different cases, one in which Sonny is aware of the reliability-increasing role being played by Frank's silent monitoring, the other in which Sonny is not so aware. In either case, I argue, Sonny's belief is reliable (etc.).

Suppose that Sonny *is* aware of the (silent) monitoring role being played by Frank. Then it would seem that, contrary to the first objection, Frank's presence *should* be held fixed, since in that case Sonny's acceptance of the testimony is grounded on his reliance on Frank's silence, together with his (Sonny's) own well-grounded belief that Frank would not be silent if Frank thought Mary's testimony were false. However, to this it might be responded that if Sonny is explicitly relying on Frank's silence in this way, then the case for (2) (= the claim of *knowledge through testimony*) will have been compromised, since in that case it is not Mary's testimony alone, so much as Mary's-testimony-together-with-Frank's-monitoring,

<sup>1</sup> I thank an anonymous referee for pointing out these objections.

that generates Sonny's knowledge that there is milk in the fridge. But this objection does not withstand scrutiny. Sonny is treating Frank's silence as a factor indicating the trustworthiness of Mary's testimony.<sup>2</sup> We can easily imagine other cases in which a hearer has evidence that the presence of some worldly condition positively correlates with the trustworthiness of speaker's testimony on a given occasion. To illustrate, suppose that Ralph is aware that Nancy has a subconscious fear of dogs, and that whenever a dog is present Nancy (who otherwise shows no sign of anxiousness) will only assert things of which she is absolutely certain. (When dogs are not present she is not as reliable; dogs focus her mind, so to speak.) Then Ralph, observing Nancy to assert that *p* in the presence of a dog, might well take the dog's presence to indicate that Nancy's testimony is to be relied upon. If (having accepted her testimony) he then counts as coming to know that *p*, this would count as a case of knowledge through her testimony, where the presence of the dog is an empirical condition that grounds Ralph's belief in the trustworthiness of Nancy's testimony. In this regard the presence of a dog is analogous to the presence of, say, signs of sincerity and/or competence in a speaker on a given occasion on which she offers testimony. In both cases, the signs themselves are part of the evidence on the basis of which the hearer is justified in regarding the testimony as worthy of being trusted.<sup>3</sup> For this reason, it would be wrongheaded to hold that Ralph's knowledge is really knowledge through Nancy's-testimony-as-given-in-the-presence-of-a-dog; at least it would be wrongheaded to say this if it meant that Ralph's knowledge is *not* properly regarded as knowledge through Nancy's testimony. Similarly for the case of Sonny's reliance on Frank's silence: it would be wrongheaded to hold that Sonny's knowledge is really knowledge through Mary's-testimony-as-monitored-by-Frank; at least it would be wrongheaded if this meant that Sonny's knowledge is *not* knowledge through Mary's testimony. On the contrary, both cases *are* properly regarded as knowledge through testimony – it's just that both cases involve the hearer's recognition of non-standard reasons for regarding the testimony as trustworthy.

<sup>2</sup> Below I will explicitly distinguish between the *trustworthiness* of testimony, and the *reliability* of testimony; I will be suggesting that not all unreliable testimony is untrustworthy, and indeed that the present case is one involving testimony that is trustworthy without being reliable.

<sup>3</sup> I do not mean to be taken as implying that such evidence is *necessary* for testimonial knowledge; that would get me into a debate (discussed in a bit more detail below) which I am anxious to avoid. Rather, my claim is that, in cases in which other conditions (to be described below) are satisfied, having such evidence *suffices* for testimonial knowledge.

The foregoing analysis raises the question: how can testimony that is unreliable (insensitive; unsafe) nevertheless succeed in being *trustworthy*? Trustworthiness is that property of testimony, the possession of which renders the testimony worthy of being trusted. Perhaps all parties can agree that a piece of testimony has this property if it is reliable.<sup>4</sup> But I submit that this is a special case of a more general scheme: testimony is trustworthy when a hearer's belief in it (acceptance of it) would yield a reliable belief. Now in most cases, the more general scheme is satisfied only if the testimony itself is reliable. But I submit that the general scheme can be satisfied even in cases involving unreliable testimony. Suppose that the hearer but not the speaker is aware of the obtaining of a condition whose presence increases the reliability of an otherwise-unreliable piece of testimony. Take the following case. *S* but not *H* has perceptual access to a given situation. Stating (in *H*'s presence) what she has observed, *S* reports that *p*. Since *H* himself was not in a position to observe whether *p*, the only evidence *H* has regarding *p* is *S*'s (observation-informed) testimony, together with whatever evidence *H* uses to assess the testimony's trustworthiness. So if *H* counts as knowing that *p* at all, *H* knows *through S*'s testimony. Now suppose that *S* herself fails to count as knowing that *p*, on the grounds that there is a defeater *d* of which she is unaware and which she cannot defeat. But suppose as well that *H* himself does have a defeater for *d* (that is, a defeater-defeater). Then *H* might know full well that *S*'s attestation is not reliable in that she would have testified that *p* even if it weren't so (given that *S* couldn't defeat *d*) – and yet even so *H* might still have grounds for trusting *S*'s testimony, given that *H* has a defeater-defeater *S* herself does not have.<sup>5</sup> (*H* has adequate grounds for thinking of *S*'s testimony as *reliable-but-for-S's-inability-to-defeat-d*; so, since *H* can defeat *d*, *H* has adequate grounds for extending trust to *S*'s testimony.) Note, though, that *H*'s having a defeater-defeater does not, by itself, put *H* in a position to know that *p* directly – rather, it puts him in

<sup>4</sup> Although it can still be asked whether it is the actual trustworthiness of testimony (so understood), as opposed to a subject's *justified belief* in the trustworthiness of testimony, that renders a subject justified in accepting the testimony. I do not enter an opinion here; but see Goldberg and Henderson (forthcoming).

<sup>5</sup> An interesting question arises whether Mary, who fails to know *via perception* that there is milk in the fridge, can succeed in generating this knowledge *for herself* via *her own testimony* (as monitored by Frank)! One question here is whether we ought to think of a speaker as her own hearer, by default. Suppose we can. Then the case for the generation of testimonial knowledge *for oneself* (in a case like that of Sonny and Mary) might parallel the case for the generation of testimonial knowledge for another hearer (presented above). In any case, I hope to return at a later time to the question of whether testimony can generate knowledge *for oneself*, where the knowledge one acquires via one's own testimony is knowledge that one oneself did not previously have in some other way.

a position to have adequate grounds for regarding *S*'s testimony that *p* as worthy of being trusted. In that case, we appear to have unreliable testimony which, relative to *H* at least, is trustworthy.

Applying this analysis to the case of Sonny's acceptance of Mary's testimony, we can see how Mary's unreliable testimony generates reliable belief for Sonny. Sonny did not observe the fridge; if he reliably believes that there is milk in the fridge at all, he reliably believes this through Mary's testimony. To be sure, she does not reliably believe (or reliably report) that there is milk in the fridge, at least not in so far as her belief and report are based on her observation of the milk carton in the fridge. But even so, her unreliable testimony can generate reliable belief for Sonny, given his explicit reliance on Frank's silence in the face of Mary's testimony. The result is that Mary's unreliable testimony generates reliable testimonial belief for Sonny.

But why does this count as a case of testimonial *knowledge*, as opposed to merely *reliable belief*? Well, in his explicit reliance on Frank's silent monitoring, Sonny has adequate grounds for regarding Mary's testimony as trustworthy in the sense characterized above (= worthy of being trusted, conducive to reliable belief). What is more, relative to Sonny's background knowledge (e.g. regarding Frank's behaviour in the face of Mary's testimony), Mary's testimony *is* trustworthy. That Mary's testimony generates a reliable testimonial belief in Sonny, under conditions in which he has adequate grounds for regarding that testimony as trustworthy in the sense above, appears to clinch the case for saying that Mary's unreliable testimony has generated testimonial *knowledge* for Sonny.<sup>6</sup>

Consider then how we might meet the second of the objections raised above against the present argument for (2). The allegation was that, because Frank's silence is not causally linked to Mary's testimony, it cannot be seen to contribute to the reliability of that testimony. In response I say that this allegation is correct but irrelevant to (2). Precisely so, since the claim of Sonny's knowledge through Mary's testimony (= (2)) is meant to be part of an account on which it is also true that Mary's testimony is, for example, *unreliable* (= (1)). In addition to being defensible by appeal to the distinction between the reliability of testimony and its

<sup>6</sup> One might wonder how this case compares with the case (from Lackey 1999) of the teacher who, though a creationist herself, teaches the theory of natural selection to her students (not wanting to be fired), thereby generating knowledge she does not herself have. The present case is like that of Lackey's creationist teacher, in that a speaker's testimony generates knowledge she herself does not have; the difference is that the teacher's testimony is reliable, whereas (in the case of Sonny) Mary's testimony is not. (Another difference is that the teacher fails to know owing to the fact that she does not believe what she asserts, whereas Mary fails to know, despite believing what she asserts, owing to the unreliability of her perceptual belief.)

trustworthiness, this combination of (1) and (2) is also defensible by appeal to what I would regard as an uncontentious sufficient condition on knowledge through testimony. The uncontentious condition is this: *H* has the testimonial knowledge that *p* when *H*'s belief that *p* depends for its epistemic credentials on the epistemic credentials of some piece of testimony, *together with whatever grounds H has for regarding that testimony as trustworthy*. To be sure, theorists will disagree over what type of grounds are needed for testimonial knowledge – for example, whether it suffices that one not have grounds for being suspicious of the testimony, or whether it is necessary that one have positive grounds for thinking it trustworthy. But all parties should agree that testimonial knowledge involves the satisfaction of some 'grounds' condition of this sort. In that case, Sonny's reliance on Frank's silence can be accommodated within the category of 'grounds Sonny has for regarding Mary's testimony as trustworthy'. And so we see that, far from being incompatible with (2), Sonny's explicit reliance on Frank's silent monitoring can be seen as supporting (2), for satisfying the 'grounds' condition on knowledge through testimony.

I just argued that both (1) and (2) hold in the case described, in which Sonny explicitly relies on Frank's silence. But we might wonder what can be said in the case in which Sonny is *unaware* of the reliability-enhancing role being played by Frank's silent monitoring.<sup>7</sup> In what follows I will be addressing my remarks to those who already accept some or another externalist (e.g. reliability) account of knowledge; with this as background I will suggest reasons for thinking that, even where the hearer is not aware of the role being played by a third party's silent monitoring, there can be testimonial knowledge through unsafe testimony. However, it must be borne in mind that, even if I am wrong about this, the foregoing defence of (1) and (2), and with it my case for the thesis that there can be testimonial knowledge through unsafe testimony, still stand. (I say this since the case to follow is admittedly somewhat programmatic.)

Suppose that Sonny is unaware of the (reliability-enhancing) role being played by Frank's silent monitoring; even so, are there conditions under which Sonny can still be said to acquire knowledge through Mary's unsafe testimony? I begin by describing what I take to be an instructive parallel case. Imagine that in some remote part of the Saharan desert there is a cave, inaccessible to humans, some of whose interior region is partially visible from a point outside the cave. The cave itself contains a skeleton, resting on a ledge in a part of the cave that is visible from the outside.

<sup>7</sup> Or at least we might wonder this in so far as we are inclined towards anti-reductionist views of testimonial justification, on which (roughly) the absence of reasons not to trust suffices to justify one's acceptance of testimony. See e.g. Goldberg and Henderson (forthcoming).

Now it just so happens that, during the interval when the sun is highest in the sky (near noontime), there is an optical illusion, whereby sunrays reaching the cave reflect off an otherwise unilluminated rocky ledge very near where the skeleton is, with the result that at that location there appears to be a skeleton. But it also just so happens that the duration of this illusion falls within an interval during the day when the heat generated by the noontime sun in this part of the Sahara is unbearable – so unbearable that no normally-functioning human being can withstand being at or near the cave during such an interval. (Since the cave itself is inaccessible, no one can take refuge in its constant 58°F.) Now, since a person who reaches the cave during the day does so during a time at which the illusion does not occur, the result is that when such a person has a visual experience as of a skeleton, the experience is veridical: the person actually sees the skeleton. Does the fact, that such a person would have been tricked by the illusion had she reached the cave at the noontime interval, render her perceptual belief (that there is a skeleton in the cave) less reliable? Presumably not: since the heat makes it practically impossible to remain near the cave during the time at which the illusion takes place, a possible world in which a subject is fooled by the illusion is not relevant to assessing the epistemic credentials of the subject's perceptual belief in the actual world (where she reaches the cave during a time not in that interval). The result is that, despite this illusion, the subject's perceptual belief that there is a skeleton in the cave is reliable, sensitive and safe. It is reliable, since in any nearby world in which she forms this perceptual belief, it is true – as the conditions for the illusion do not materialize in any of the situations in which she reaches the cave (and so forms the perceptual belief in question). It is sensitive, since in the nearest world in which there is no skeleton in the cave, she does not form the perceptual belief that there is, as, given the heat, she is nowhere near the illusion-generating but skeleton-free cave. And her belief is safe: there is no nearby world in which the cave has no skeleton but she forms the perceptual belief that it does.

I submit that the cave case establishes that contingent features of the environment, of whose presence and role in the formation and sustainment of reliable belief the subject herself is unaware, can nevertheless be held fixed for the purpose of determining the relevant set of possible worlds to consider when evaluating a relevant belief's epistemic standing. To a first approximation (and I won't attempt more here), the principle involved is this: given a case in which a subject forms the true belief that *p* in environment *E*, if an environmental feature *f* of *E* reflects an (enduring) environmental regularity, which regularity plays (or would play) a regular causal role in 'preventing' the subject from acquiring the false belief that *p* in circumstances that were otherwise like the actual one save



for the falsity of  $p$ , then  $f$  itself can be held fixed for the purpose of evaluating the reliability (etc.) of the subject's actual belief that  $p$ . This approximation is no doubt crude. But it is clear enough to see that something like it is what accounts for the reliability of, for example, our perceptual beliefs about the three-dimensional world. After all, it is widely recognized that these beliefs are formed on the basis of sensory evidence that underdetermines the set of possible physical arrangements that could have given rise to the relevant sensory display;<sup>8</sup> so if despite this we count our perceptual beliefs reliable, it is because we hold fixed those features  $f$  that amount to *perceptual invariances*, and so (in evaluating the reliability of our perceptual beliefs) do not consider possible worlds in which these invariances do not hold.

Now this invariance principle (as we might call it) can be brought to bear on the case of Sonny and Frank. My claim is that Frank's presence in the kitchen should be regarded as an invariance – with the result that his presence in the kitchen should be held fixed for the purpose of evaluating the reliability of Sonny's belief. Admittedly, Frank's everyday presence in the kitchen is nowhere as 'deep' a regularity as the perceptual invariances mentioned above: it is a merely 'local' invariance. But given the role that this merely 'local' invariance plays with respect to Sonny's acquisition of a belief through Mary's testimony, and given that this 'local' invariance is in fact an enduring (if not very deep) regularity, I submit that it falls within the scope of the invariance principle above. In that case, Frank's presence in the kitchen can be held fixed for the purpose of evaluating the epistemic standing of Sonny's testimony-based belief. But then, as we have seen above, Sonny's testimony-based belief is reliable, sensitive and safe.

In sum, given, for example, reliabilism about knowledge, cases are possible in which a hearer acquires knowledge through another's unsafe (insensitive; unreliable) testimony, even under conditions in which the hearer is not aware of the reliability-enhancing role being played by the environmentally invariant factor. And since even stronger (more 'internalist') requirements on knowledge are satisfied in the case in which Sonny *is* imagined to be aware of the reliability-enhancing role being played by Frank's silent monitoring, we can reach the desired conclusion – testimonial knowledge through unsafe testimony – no matter one's views regarding the conditions on knowledge.<sup>9</sup>

<sup>8</sup> This point is made, in alternative formulations, in Gibson 1966: 3, 52–54, 81–84, 156–63, 186–223, 250–65, 284–86; Shepard 1987 and 1992; Hubel 1995 Chapter 8; and Zeki 1993: chs. 23, 25.

<sup>9</sup> This paper has benefited from discussions with Brad Monton, Duncan Pritchard, and Jim Pryor; and thanks to an anonymous referee for comments on an earlier version.

University of Kentucky  
Lexington, KY 40506-0027, USA  
scgold@uky.edu

### References

- Lackey, J. 1999. Testimonial knowledge and transmission. *The Philosophical Quarterly* 49: 471–90.
- Gibson, R. 1966. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.
- Goldberg, S. and D. Henderson. Forthcoming. Monitoring and anti-reductionism in the epistemology of testimony. *Philosophy and Phenomenological Research*.
- Graham, P. 2000. Transferring knowledge. *Noûs* 34: 131–52.
- Hubel, D. 1995. *Eye, Brain, Vision*. New York: W. H. Freeman.
- Shepard, R. 1987. Evolution of a mesh between principles of the mind and regularities of the world. In *The Latest on the Best*, ed. J. Dupré, 251–76. Cambridge: MIT Press.
- Shepard, R. 1992. The perceptual organization of colors: an adaptation to regularities of the terrestrial world?. In *Adapted Mind*, ed. J. Barlow, L. Cosmides and J. Tooby. New York: Oxford University Press.
- Zeki, S. 1993. *A Vision of the Brain*. Oxford: Blackwell Scientific.