JUSTIFYING BELIEFS ACQUIRED THROUGH TESTIMONY:

A DEVELOPMENTAL ACCOUNT

By

DOROTHY CHEN

April 1, 2013

A senior thesis submitted to the Philosophy Department of Columbia University for the degree of Bachelor of Arts

Supervisor: Professor Philip Kitcher, Columbia University
There are some things we know that we would not normally doubt. Take my birthdate for example. I was born on the 17th of July 1991. I think I know this, and am fairly certain about it. But do I really know my birthdate? And, if so, how do I know?

Surely, it would be too much to ask of my newborn self to check the calendar on the delivery room wall and say to itself, “I must remember this date, for it is my birthday.” Nor is it likely that I was simply born knowing my own date of birth. So I must have learnt it from some secondary source, such as one of my official documents or my parents’ testimony. Of these two options, the second one seems to be the only reasonable one, since as far as I can remember, I knew my birthday before I could even understand what an official document is, let alone decipher its content. But even if my memory serves me wrong, the birthdate specified by my official documents is likely to have been provided by my parents when they applied for those documents anyway. Hence, testimony plays an important role even in the acquisition of such mundane beliefs as that of our own dates of birth.

In fact, large portions of what we know are derived from testimonial sources. From these sources there arise not only our beliefs about the distant past and unobservable present, but also our predictions for future states of affairs. This is because everything we did not learn through direct perception we learned through some intermediate source, such as conversations, publications, and webpages. Though the types of intermediate sources we draw on differ greatly from context

---

1 I also learned from testimony that the information on my official documents can be relied upon.
2 For example, we are often led to form beliefs about future weather conditions by the testimony of weather reporters.
to context, they all have in common the fact that they contain information as presented to us by another person or group of people. This latter fact makes the intermediate sources instances or subsidiaries of testimony (broadly construed). But if this is indeed the case, then large portions of our systems of beliefs must either be acquired from testimonial sources or are in some way dependent on beliefs acquired through testimony. Therefore, it might be difficult for us to contend that we know anything at all unless we could establish testimony as a justifiable source of knowledge.

Given the threat of such general skepticism, testimony has increasingly become a topic of interest in recent philosophical discourse. But instead of studying testimony in the full variety of contexts where knowledge is acquired from others, the majority of existing discussions of testimony start with an idealized context. The setup of this context eliminates all sources of information and interlocutors except a single speaker S and an audience A, and then supposes that S wishes to pass onto A a piece of information given by the proposition p. There is, however, a restriction on the transmission of p; namely, that S must engage in a particular speech act (assertion), the conditions governing which are:

1. S makes an utterance p with the intention of getting A to believe p
2. If A understands S’s utterance p and recognizes S’s intention in 1, A could then proceed to evaluate p with regards to her system of beliefs

---

4 For a further discussion on how assertion features in testimony, see Williamson (1996), Owens (2006), Lackey (2007), and Goldberg (2009).
5 Her evaluation is based on her system of beliefs, as opposed to other sources of information such as alternate informants, publications, and webpages, because only one testimonial source enters at a time in the idealized context.
3. Once the evaluation is complete, A could either come to believe p, or she could withhold judgment.

Suppose it so happens that, for a given proposition p, A chooses to acquire the belief p as a result of this idealized situation. Then being good epistemologists, we should also ask the following set of justificatory questions:

- Is A’s belief p justified?
- Was A justified in coming to believe p?\(^6\)
- What types of justification are we concerned with in either of these cases?
- What are the natures of those justifications?
- Do the justifications depend on S’s knowing that p? If so, how?

Though seemingly wearisome, these questions are actually quite important. On a very basic level, they stimulate inquiry into the different types of justificatory frameworks that could support beliefs acquired through testimony, and perhaps even prod the inquirer in the direction of some particular framework. But more importantly, how these questions are asked and responded to will determine the strength of our anti-skeptical arguments, which are essential to defending our right to hold our existing systems of beliefs.

---

\(^6\) This question is distinct from the one above because the process of acquiring the belief p might be justified without the belief itself being so, and vice versa. e.g. Suppose that for A’s belief p to be justified, A must be aware of the justifiable means through which she acquired p. Then it might well be the case that A was justified in coming to believe p, but A’s later belief p is not justified because she has forgotten how she acquired p. Conversely, suppose that A’s belief p is justified insofar as it is produced by a reliable process (i.e. a process that has a tendency to produce a percentage of true beliefs). Now suppose that A is Laurence Bonjour’s (1980) incredibly reliable clairvoyant, then all her beliefs, including p, would be justified by the reliable process, yet she would have no idea how she comes by those true beliefs; therefore, she would not even think that she was justified in coming to believe p.
One popular attempt to answer these justificatory questions appeals to the possibility of induction on the speaker’s past accuracy. Consider a variant of the idealized context where S, instead of uttering a single proposition, makes many assertions in A’s presence over time. By the rules of the idealized context, A could respond to each of S’s assertions by either acquiring a new belief or withholding judgment. Now suppose that over time, A keeps track of all her beliefs acquired in this way, and continuously evaluates them in light of new pieces of information that arise as a result of her own observations. Then one of two possibilities might arise. On the one hand, A might discover that the majority of beliefs she acquired from S are false, in which case she might tag S as an unreliable source of information and think twice about believing S again. But on the other hand, A might also discover that S’s past utterances have been quite accurate, in which case she would tag S as a reliable informer and place more trust in S’s future utterances.

In light of these two possibilities, theorists who have defended the inductive explanation would point to the latter and say that A’s increased trust there is justified. Their reasoning is roughly as follows: if A has observed a history of correspondence between S’s assertions and facts about the world, then according to everyday inductive practice, A would be justified in believing those of S’s assertions even in instances where the correspondence could not be checked. But if, instead, A has observed a history of disagreement between the two, then A would have reason not to believe whatever S says.

This primitive justificatory framework, though it successfully explains the difference between the two possibilities, is inadequate. The problem I wish to raise is not a general one about

---

7 A clear exposition of this account could be found in David Hume’s *Enquiry Concerning Human Understanding*, s???. 88. See also Mackie (1970) and Fricker (1987).
induction as a form of justification (given the familiar critiques of David Hume and Nelson Goodman\textsuperscript{8}). Rather, I will take for granted the validity of our everyday inductive practices but explain why they should not be extended to the testimonial case. First, let us recognize that there are specific paradigms governing our everyday inductive practice. For example, when pharmaceutical scientists conduct clinical trials to test the safety and efficacy of particular drugs, they place particular constraints on the size and variability\textsuperscript{9} of their trial participants in order to mirror the demographics of the drugs’ targeted populations.\textsuperscript{10} It is very plausible that these paradigms of size and variability would also carry over to inductive practices made in the testimonial case. And if so, audiences would need to monitor a good number of S’s utterances on various subjects before they could make justifiable generalizations about the accuracy of her assertions as a whole, although practically speaking it is hardly plausible to expect audiences to undergo such thorough monitoring for each of their informants. The induction theorist about testimony must thus give a practical solution to this problem, or alternatively give an argument against extending the paradigms governing our everyday inductive practices to the testimonial case.

This concern aside, the inductive framework is also unacceptable for testimonial contexts in which we cannot monitorthe speakers’ past accuracy. Suppose as an example that I am hopelessly lost in a foreign city and decide to ask a stranger for directions. In light of our present discussion, the position that I find myself in is like that of the audience in the idealized context.

\textsuperscript{8} Hume, 1942/1902; Goodman, 1955.\textsuperscript{9} There are all sorts of complications in specifying what adequate size and variety amount to. For example in the drug trial, a varied sample might ask for participants of different ages, sexes, weights, ethnicities, and disease types (or lack of disease types for the placebo group).\textsuperscript{10} For more details, see any introductory textbook on clinical research methods, such as Friedman et al. (2010).
Like A, I have no other sources of information to depend on except the testimony of a single speaker who is uttering directions to me with the intention of getting me to believe what she says. Yet unlike the possibilities just considered, I could not track the accuracy of the stranger’s past utterances.

Put more precisely, cases involving strangers pose a priority problem for the inductive theorist. The procedure that A must undergo to inductively justify acquiring a belief from S is as follows:

i) A remembers the beliefs she acquired from S in the past

ii) A monitors the accuracy of those beliefs and arrives at an estimate of S’s rate of accuracy in the sample.

iii) A induces how likely S is right in the present instance

It is clear from this characterization that i) and ii) must precede iii), for A’s recollections and past monitoring provide the information she needs to make an inductive generalization about S’s reliability. But in the example of the stranger providing me with directions, neither of the conditions i) and ii) could be fulfilled, so I could not be inductively justified in acquiring a belief about directions from the stranger.

Perhaps in this case I could be justified by a higher order induction. Suppose that I am a globetrotter and have had much experience asking strangers for directions in foreign lands, and have generalized from those successful encounters that locals are generally quite reliable sources for directions. Here again, constraints on sample size and variability might pose practical concerns, since most of us who take directions from strangers are not globetrotters. But more

---

11 I might have even developed strategies to mark out those strangers who are likely to be more accurate about directions than others. I might for example choose to ask food deliverers and people holding grocery bags over those holding cameras. More on these distinctions in the final section.
problematically, if we consider where our practice of asking strangers for directions came from, we may discover that we learned to do so through the testimony of someone else\textsuperscript{12}, which means that the practice itself also needs to be inductively justified under the current framework.

To see the full extent of this regress problem, consider the testimonial beliefs we acquired as children. In the cases of the very first beliefs that our infant selves acquire through testimony, our informants should be treated in the same way as strangers because we could not have been alive to hear any of their past utterances, let alone remember and assess them. As we grow slightly older, however, the speakers should no longer be treated as strangers but as familiar informants. But even then, it is unclear whether or not the beliefs we acquired from those familiar informants could be inductively justified. Here are some reasons why they would not be. First, depending on how young we were, we might not have had the memory, monitoring, and reasoning capacities that we needed to fulfill conditions i) through iii).\textsuperscript{13} Second, even if we had these capacities, it is unclear that we would know to use them to track a speaker’s past accuracy unless our previous beliefs acquired through testimony had led us to error. In light of these concerns therefore, the proposed justificatory framework seems unable to vindicate our initial testimonial beliefs.

Why should we care if these early beliefs are not justified? Well the shortest and most compelling answer is this: the skeptic cares, so we should too. If indeed none of the early beliefs we acquired through testimony are justified, then a skeptic about testimony might proceed in the following way. She might, for example, point to the fact that one of the conditions needed for the

\textsuperscript{12} For more on this point about justifying testimonial knowledge through collective observation, see Coady (1992), chapter 4.

\textsuperscript{13} According to Harris (2012), children as early as 16 months old are able to monitor the accuracy of their speakers’ past utterances.
inductive justification (condition ii) requires that we monitor the accuracy of beliefs we gained through testimony. And one way we might go about checking these beliefs is by verifying them with what we already know. But depending on our stage of development, we might either have an insufficient number of beliefs to ground the acquisition of new ones, or be led by our unjustified prior beliefs to mistakenly trust unreliable agents and acquire more false beliefs. Thus, minor inaccuracies in our initial beliefs might be exacerbated by the inductive process, causing more and graver errors in the systems of beliefs we ultimately end up with.

It seems crucial, in light of the preceding diagnosis, that the anti-skeptical argument includes an account of the development of testimonial justification. We might for instance start by offering the skeptic a gambit, by admitting that, indeed, the initial beliefs we acquire through testimony are not justified. But then, over time, we learn to monitor those initial beliefs and adjust them (as well as the strategies we use to obtain them) so that ultimately we end up with sets of more justifiable, if not justified, beliefs.

The anti-skeptical argument I will offer is analogous to the approach often taken on the question of whether the sciences yield truth (or approximations to truth). When people first started making scientific inquiry, they often used unreliable methods that led them to arrive at false beliefs about nature. But over time, they discover that their views are false, so they revise their methods of accepting new beliefs, and gradually develop more reliable standards and more accurate bodies of beliefs. Eventually in some areas of inquiry, they arrive at methods that are likely to yield true beliefs (or at least beliefs that are good approximations to truth). The beliefs that they acquire
using these methods are thus justified through their reliability. In the following, I argue that the beliefs we acquired through testimony could be similarly justified.

Degrees of Belief

Let us recall an instance of testimony in the idealized context. There, the speaker S makes an assertion p, which is heard and understood by the audience A. A could then choose to proceed in two ways: she could acquire the belief p or she could withhold judgment. Thus from the very setup of the problem, A is boxed into only one of two epistemic attitudes towards p: belief and disbelief.

Yet it is quite clear from the way that propositional attitudes are used in ordinary language that A could hold a wider range of epistemic attitudes towards p than just belief and disbelief. She could for example be sure of p or be fairly certain that p. She could also find p improbable or see it as necessarily false. She might even be completely uncertain about p. In all these cases, we might say that A has some estimate of p’s plausibility. One way to formalize these estimates is by distinguishing degrees of belief, wherein each proposition entertained is assigned a numerical value between 0 and 1, with 0 representing absolute certainty that p is false and 1 absolute certainty that p is true. Under this model, if I were to say that I am fairly certain that the empire state building is more than 400m tall, then I might be interpreted as assigning a degree of belief 0.8 to the proposition “the empire state building is more than 400m tall”.

14 Kitcher, 1993.
Despite the rigor of this approach, the idea of assigning a single real value to an agent’s degree of belief is counterintuitive. It is hard to suppose that anyone seriously distinguishes for instance the value 0.8 from 0.79 or 0.81, let alone the infinite set of real numbers in between. Thus perhaps the assignment of real values to degrees of belief should only be seen as a mathematical tool used to facilitate model building and not as a realistic portrayal of epistemic attitudes. And since I do not intend to develop any mathematical apparatus in the following sections, it would make more sense to suppose that each epistemic attitude, such as “it is certainly true that…” or “it is probable that…” is assigned a subinterval in [0,1].

Depending on where the interval lies, the agent could be said to have varying degrees of belief with regards to p (where the interval’s being closer to 1 means a higher degree of belief and being closer to 0 means lower). These degrees are by no means fixed over time because epistemic agents typically change their estimates of p’s plausibility in response to new evidence. This means that the interval representing A’s epistemic attitude towards p could be shifted left or right as well as made narrower:

---

16 These intervals may vary in size and overlap one another.
Over a period of time, the resulting interval should be one of the following:

1) If \( p \) is true, the interval assigned to \( p \) shift towards 1 and narrows\(^{18}\).

2) If \( p \) is false, the interval assigned to \( p \) shift towards 0 and narrows\(^{19}\).

It should also be noted that each epistemic agent need not have an existing degree of belief for any given proposition \( p \). Since a degree of belief represents a particular epistemic attitude towards some given proposition, epistemic agents would not have intervals assigned to propositions that they do not hold any epistemic attitudes towards. For example, prior to ever considering the statement “There are aardvarks in Antarctica”, I would have no estimate of its plausibility and therefore have no degree of belief associated with it. But as soon as I have been led to think about it, I might then assign it a subinterval of \([0,1]\), determined by my newly acquired epistemic attitude, or I might remain utterly uncertain and assign it the entire interval \([0,1]\).

How does testimony lead people to new assignments of degrees of beliefs? It seems that for an audience \( A \) in the idealized context, \( A \) could go about assigning a degree of belief to \( p \) in one of three ways. First, \( A \) could make the assignment without taking into account at all the testimonial speaker or context. In this case, \( S \)’s utterance \( p \) serves only as a provocation for \( A \) to think about \( p \) in virtue of his or her system of beliefs, but does not hold any influence over \( A \)’s actual degree assignment. For instance, if in a conversation someone tells me that there are aardvarks in Antarctica, I may take that statement to be highly implausible solely as a result of my prior belief

\[^{18}\text{More precisely, the limit of the interval’s upper bound goes to 1, and the limit of the difference between the interval’s upper and lower bounds goes to 0.}\]
\[^{19}\text{The limit of the interval’s lower bound goes to 0, and the limit of the difference between the interval’s upper and lower bounds goes to 0.}\]
that aardvarks are native to Africa, and not the consequence of any factors relevant to the testimonial context. On the other hand, it might also be the case that A’s assignment of an interval is wholly dependent upon the testimonial utterance. This second type of assignment is similar to the first in that S’s utterance also induces A to think about p in virtue of his or her system of beliefs, but, unlike the first case, A could not find any prior beliefs that bear on the degree to be assigned to p (for instance, if I had not known anything about aardvarks or the typical wildlife variety of Antarctica). A’s degree assignment in this second case thus becomes wholly dependent on S’s utterance. In addition to these two cases, we should consider a third, a mixture of the first and second, where A’s assignment of degree is partly dependent on the single instance of testimony and partly on his or her prior beliefs. In the example, the mixed case might occur if I am told that fact about aardvarks by a zoologist in a bar, and because I am not too sure of her level of intoxication, I decide to check her information against my own beliefs.

It is likely that the mixed cases account for the majority of testimonial exchanges in real world contexts. But despite this observation, it is useful to start with the simpler cases in which the hearer has no prior views about p. Consider an analogy: in the empirical sciences physicists try to study a particular force. Despite the fact that in most everyday instances particular forces occur in combination with others, it is useful to start by studying the targeted force in isolation before trying to understand how it interacts with other forces. Similarly, when we consider our actual sets of beliefs, we are likely to discover that each of our beliefs is justified in several ways. So in order to discover the justifications for the beliefs acquired through testimony, it is useful to begin with an examination of the pure testimonial cases.
Evaluation Matrix In Pure Testimonial Cases

Suppose I have no prior beliefs relating to rhubarb leaves and someone tells me that they are poisonous. How would I go about assigning a degree to the plausibility of that statement? It seems to me that there are two types of evaluations I would consider: those related to the speaker and others related to the testimonial context. I no doubt make the two types of evaluations together or in light of each other in most instances, but for simplicity’s sake, let us examine each in turn.

Speaker-Related Evaluations

When presented with an utterance p, a reasonable question for any epistemically responsible audience to ask is: should I trust the person who is presenting p to me? There are several ways that the audience could go about answering this question. An obvious one is by observation of the speaker’s past accuracy. Indeed, if the person who tells me that rhubarb leaves are poisonous has told me many other things in the past, most of which are quite accurate, then I am more likely to assign a higher degree of belief to this person’s statement about rhubarb leaves. Yet it would be quite foolish of me to consider the accuracy of all of the speaker’s past utterances. It would surely not matter if my informant often gets the dates wrong, gives misleading directions, or even calls people by incorrect names, as long as she has a history of accuracy when it comes to conveying information about plants (if she is a botanist for instance).\(^{20}\) So all I need to monitor is the accuracy of the reference class of her roughly rhubarb-related past utterances\(^ {21} \). More generally then, the audience needs to only observe a history of speaker accuracy relating to the

\(^{20}\) Faulkner, 2002.

\(^{21}\) There is an interesting issue about just what the “right reference class” is. The specification often used by literature on prediction supposes that the appropriate reference class is the narrowest one for which reliable statistics are available. See Kahneman & Tversky (1979).
appropriate reference class of the present proposition in order to be justified in assigning a higher
degree of belief to that proposition.

In fact, the audience need not even make that observation herself. In the rhubarb example, all I
needed to know is that my speaker is a botanist, which is a piece of information I could have
learned from a certificate on her wall or someone else’s testimony. In general then, an audience
A has a reason to assign a higher degree of belief to S’s utterance p insofar as A recognizes some
indicator x in S which demonstrates S’s belonging to some class of individuals K who are likely
right about matters like p.  

This formalization points to an interesting social dimension in the initial degree of belief
assignments. First, x itself might be a socially constructed indicator. Consider for instance the
Pulitzer Prize for investigative reporting. It would not be far-fetched to suppose that this prize is
an instance of an indicator x for membership in a class of journalists to whose reporting we
should assign higher degrees of beliefs. However, if we consider how this award was made in the
first place, we would discover that it was nominated and voted for by a selection committee of
expert journalists, and it is on their recommendation that this indicator acquires any value.
Furthermore, the way through which we come to believe that x is a good indicator for
membership to K could also be testimonial. We might for instance not know anything about the
Pulitzer Prize or what it means unless someone tells us that it is a good indicator for journalistic
excellence. Hence, although information about a speaker’s credentials and competency are good

---

22 Craig, 1990.
23 Ibid.
evaluators for her reliability, it often needs to be justified through prior testimony, and thus this strategy merely delays our present problem regarding testimonial justification without solving it.

Another consideration the audience needs to use to evaluate a speaker besides her past accuracy is her likely sincerity.24 Returning to the rhubarb example, suppose my informant is one of my close friends who happens to know that I just started growing rhubarb in the garden so, by uttering the statement, is warning me against accidentally ingesting one of their leaves and being poisoned. Here I have every reason to believe in my friend’s sincerity and thus take her utterance seriously. On the other hand, one could imagine a door-to-door salesman who utters the same statement in an attempt to sell me a spray that would damage the leaves of a few plants (including rhubarb leaves), then I would be more doubtful of her sincerity and therefore be more cautious about taking up the belief.

The evaluation underlying the above phenomena seems to be the following: if the audience A’s relationship with S is such that S is unlikely to have ulterior motives to deceive or harm A, then A is more likely to assign a higher degree of belief to propositions asserted by S. This evaluation is especially powerful in cases where children are just beginning to acquire beliefs through testimony. In a 2009 experiment done by Corriveau and Harris for example, children between the ages of 3 and 5 are shown to place more testimonial trust on speakers to whom they are emotionally attached. In this study, children from two daycare centers were shown a variety of unfamiliar objects from the hardware store, and had the option of asking one of two caregivers present (one from each of the daycare centers) what the object was called. Whenever a child asked, both caregivers volunteered an answer, but each assigned a different name to the same

24 Audi, 1997.
object. For instance, one caregiver might call an unfamiliar metal bathroom hook a “linz” while the other one called it a “slod”. Then when the children from the two daycares were later asked to choose between the two names for the object, they systematically preferred the name that is given by their own caregiver and not one that is given by the unfamiliar caregiver.

The result of this experiment could be interpreted in two ways. Either, the children trusted their own caregivers more because they had observed their own caregiver’s past accuracy whereas they have not for the other caregiver, or the children trusted theirs more because they were more emotionally attached to her. There are two reasons why we should make the second interpretation. First, the behavior of children in the experiment could be placed in the greater context of infants who have been known to seek help only from familiar figures when they are faced with uncertainty and danger. Indeed, this selectivity has been widely studied\(^{25}\), and is shown to be all but universal amongst children growing up under normal rearing conditions. Harris then builds on this background and cites examples of children who hold different attachment attitudes towards their mothers and who display different degrees of trust with regards to what their mothers say.\(^{26}\) For instance, children who hold avoidant attitudes\(^{27}\) towards their mothers were shown to be systematically less likely to trust their mothers’ testimony than those children who are ambivalent\(^{28}\). What this latter experiment shows is that even though both types of children had the tools and experiences to monitor their mother’s past accuracy, they held different attitudes of trust towards her on the basis of their divergent emotional attachments.

\(^{25}\) See Bowlby (1969) for a clear exposition.
\(^{26}\) Harris, 2012.
\(^{27}\) Those who show few signs of missing their mothers during her periods of absences and ignore her on her return.
\(^{28}\) Those who cope less well with her absences and are less reassured by her return.
It should be noted that Harris’s work adds further support to utilizing a degrees of belief model to describe acquisition of beliefs through testimony. His experiments are more nuanced than simply indicating that children trust only those they have strong emotional bonds with and distrust those with whom they do not. Rather, Harris suggests that emotional attachment could be evaluated on a scale, the values of which could then be mapped onto gradated levels of trust. Therefore, he could be interpreted as saying that children are more likely to assign higher degrees of belief to information presented by people they are closed to and lower degrees to utterances given by those they are not.

Furthermore, Harris’s insight might also provide a solution to the regress problem we encountered in the first section. We noted there that the beliefs we acquired through testimony as children could not be inductively justified because we had neither the experiences nor the insights to track our informants’ past accuracies. Yet if it could be supposed that children start off merely making evaluations based on their emotional attachments with informants, then they would have an initial method of acquiring beliefs through testimony upon which the rest of their system of beliefs could later be built. Though this method of acquiring initial beliefs is not very reliable and does not justify the beliefs formed, it does yield a basis upon which we could later make adjustments and acquire more accurate beliefs.²⁹

²⁹ More on this in the final section.
Imagine a case where a botanist makes the utterance “rhubarb leaves are poisonous” at both an academic conference and the World’s Biggest Liar Contest\(^{30}\) and suppose that an audience A is only present at one of these contexts. Would the degree of belief A assigns to what she hears be the same as that which she would have assigned had she been in the other context?

If speaker-related evaluations were sufficient for the assignment of degrees, then the answer should be yes because the same botanist makes the assertion in both contexts. But if A were at all sensible, she would not place the same trust in a lying-contest contestant as she would a conference-attending academic. Indeed, A should reasonably recognize that the conference and the contest are two very different contexts governed by distinct speech conventions – academic integrity in the context of the conference and competitive insincerity in the context of the contest, the violations of either would result in the speaker being viewed negatively by others.\(^{31}\) Given A’s recognition of these conventions, she would thus expect the botanist to be more sincere at the conference than in the contest, which motivates her to assign a higher degree of belief to the utterance made at the conference than the one presented at the lying contest.

This argument, though it successfully accounts for the discrepancy in the two cases, could not be immediately applied to other contexts because it depends too much on explicit presentations of speech conventions. Both the academic conference and the lying contest are special contexts wherein the conventions governing testimonial behaviors are explicitly drawn, in the forms of

---

\(^{30}\) This is a contest held annually in Cumbria, England, in which competitors from around the world have five minutes to tell the biggest and most convincing lie they can.

academic dishonesty codes and contest rules, which makes it possible for participants of these contexts to recognize the existence of the conventions and conform to them. However, in most everyday testimonial contexts, the conventions of speech are not presented so clearly. Usually they are only regarded as rules of thumbs or paradigms of speech, as given by platitudes such as: one should always tell the truth, only say things one knows, and never intentionally lead others astray. These paradigms, unlike the explicit conventions, are not normally codified and are only discoverable upon reflection, so one should be careful not to assume that they are similarly recognizable and effective in dictating testimonial behavior as their explicit counterparts.

With this said, a case could still be made to treat these paradigms in the same manner as explicit conventions. First, we hold certain expectations about the sincerity and competency of speakers in contexts where the conventions are not explicitly given. For instance, we would be remiss to expect the speakers at dating events to be completely forthright about themselves or our average dinner companions to be knowledgeable about astrophysics. We hold these expectations not because we have been told them explicitly or have read it off some set of official rules about dating events and dinner parties; rather, we learn that these expectations are useful ones to hold after some experience with the contexts presently considered. Second, we expect to be similarly subjected to negative reactive attitudes when violating rules of thumb as we would when breaching explicit rules. For example when we are caught in a lie or are shown to be giving wrong directions, we would expect others to view us in the same negative manner as we would if we had violated an academic dishonesty regulation or lying contest rule. If the implicit

---

32 Unless I am having dinner with astrophysicists.
33 For another example, recall my suggestion to justify the directions we acquire from strangers through a “higher-order induction”, where I observed from my previous successes in acquiring directions from strangers that the context is one in which speakers are generally reliable.
34 Faulkner, 2007.
paradigms are not given the same weight as explicit rules, then it is hard to see how those reactive attitudes could be justified.

In fact one might even think, with David Lewis\textsuperscript{35}, that there are common paradigms, such as that of truthfulness and trust, which persist through all conversational contexts unless there are overriding reason to think otherwise. So it would not matter whether or not these paradigms are explicitly stated or merely implicit, they would nonetheless be expected to hold in the majority of testimonial exchanges. Lewis arrived at this view through his seminal account of conventions, which presents each convention as a regularity $R$ (persisting in a population $P$) that satisfies the following conditions:

1. Everyone\textsuperscript{36} conforms to $R$.
2. Everyone believes that the others conform to $R$.\textsuperscript{37}
3. This belief that the others conform to $R$ gives everyone a good and decisive reason to conform to $R$ himself.
4. There is a general preference for general conformity to $R$ rather than slightly-less-than-general conformity – in particular, rather than conformity by all but any one.
5. $R$ is not the only regularity meeting the last two conditions.

\textsuperscript{35} Lewis, 1975.
\textsuperscript{36} By “everyone” Lewis does not actually mean all the individuals in a population $P$; he only uses it to represent most individuals in $P$ under ordinary circumstances. Indeed, just because an English driver starts driving on the left side of American roads, driving on the right does not stop being a convention for the population “drivers in America”.
\textsuperscript{37} This is not necessarily equivalent to believing that every other individual in $P$ is conforming to $R$. Lewis himself indicates that he is much more flexible about how the word “others” are interpreted. It could be used to mean the others that the agent is currently dealing with, encountered in the past, or will encounter in the future. Thus the belief in the conformity of others might arise from many other sources than just the current perception of general conformity.
6. The various facts listed in 1-5 are matters of common (or mutual) knowledge: they are known to everyone, it is known to everyone that they are known to everyone, and so on. This knowledge may be merely potential: knowledge that would be available if one bothered to think hard enough.

One advantage of Lewis’s account is that it explains why negative reactive attitudes arise when conventions are violated. Indeed 4 suggests that there is a general preference for general conformity to R, which means that if certain members of P do not conform to R, they would be making the overall situation less good for everyone else. Since nonconformity results in a deviance from the equilibrium, the conforming members of P would thus have reason to motivate conformity in those deviating members through the employment of reactive attitudes.

Using the analytic framework presented by these criteria, Lewis then shows that the regularity of truthfulness and trust\textsuperscript{38} in a given language L satisfies the criteria he set out:\textsuperscript{39}

1*) Members of any linguistic community P with language L often speak sentences to one another. Whenever they do, the speaker ordinarily utters one of the sentences that she believes to be true, and the audience responds by coming to share in the speaker’s belief (if she does not hold it already) and by adjusting her other beliefs accordingly.

2*) Members of P believe that this regularity prevails among them. Each believes this because of his past experience of others’ past truthfulness and trust.

\textsuperscript{38} By truthfulness, Lewis means the practice of uttering only sentences in the language that the speaker believes to be true; and by trust, he means the practice of responding to other’s utterances by coming to believe the uttered sentences.

\textsuperscript{39} I rephrased Lewis’s argument so that it is consistent with the terminology I use in this paper. The essence of the argument should still be the same.
3*) If the speaker believes that everyone in P is trusting, then the speaker has reason to expect that by uttering certain sentences that he believes to be true, he can impart certain beliefs in the audience. Additionally, if the audience believes that everyone in P is truthful, then the audience has reason to infer that the speaker’s sentence is true and therefore to respond trustingly.

4*) Speakers and audiences in P have a general preference for truthfulness and trust. This is sustained by a common interest in communication. Everyone wants to be able to impart correct beliefs and bring about appropriate action in others by means of testimony, and everyone wants to be able to learn through testimony about things that he does not have perceptual access to.

5*) There is another language L* that does not overlap with L (language spoken by P) in which the regularity of truthfulness and trust also prevails.

6*) These facts are known, or could potentially be known, by everyone in P, and everyone knows that everyone knows them, etc.

If the population P is interpreted as everyone we have testimonial exchanges with, then Lewis’s argument is sufficient to establish truthfulness and trust as a general paradigm persisting in the vast majority of our testimonial exchanges. There are, however, exceptions to the general case. In contexts such as the lying contest and dating event, the paradigm of truthfulness and trust is suspended and replaced by another convention. This process of substitution is fairly straightforward when the alternative convention is explicitly stated. For instance when the organizers of the lying contest explicitly state the rules of the contest, they in effect suspend the pre-existing norm of truthfulness for all participants of the contest and replace it with one of
insincerity. So whenever contestants enter the contest and understand its rules, they withhold their adherence to the truthfulness paradigm and conform to insincerity instead.

However, the presence of an alternative convention need not always suspend the pre-existing one. In cases where the two conventions do not contradict, they may very well be combined in a single context. For example, one might think it is not enough for speakers at academic conferences to present only findings they believe to be true; the findings need to be actually true. If this intuition were right, it would suggest the existence of an additional convention of reliability in that context, which prompts speakers to make only true statements. Yet the mere fact that this alternative convention exists does not mean that it replaces the truthfulness paradigm, because it is perfectly reasonable for speakers to utter only statements that they believe to be true and are actually true. Hence it might well be the case that both the conventions of reliability and truthfulness hold for utterances made in academic conferences.

The analysis could analogously be made in contexts where alternative conventions are implicit, baring one complication. Indeed, the alternative convention could similarly either suspend the pre-existing paradigm or be added alongside it. For examples of the first possibility, consider instances where we speak in metaphors and hyperbole, or make sarcastic comments. In these uses of the language, the convention of truthfulness is suspended and replaced by such norms as humor and poetic elegance; if it were not, we would no doubt take jokes and exaggerations all too seriously. For an example of the second possibility, recall the case where I ask a stranger for directions. Suppose that a convention of reliability governs the utterances given by the stranger in this context, then it is unlikely that this alternative convention would supersede the usual
truthfulness paradigm. In fact, the sincerity of the strangers is likely part of the reason they are reliable, because they would not want to lead the already hapless foreigner astray. One should be careful to note however, that implicit conventions, unlike the explicit ones, are discoverable only through experiences of and reflections upon given contexts.\textsuperscript{40} This means that as we encounter more and more testimonial contexts, we discover new conventions that could either replace or be added to general paradigms, so that over time, our understanding of the implicit conventions of particular contexts become increasingly nuanced. This development would no doubt help us make more accurate degree of belief assignments\textsuperscript{41} in the long run.

\textbf{Justifying Initial Belief Assignments}

A good amount of energy was spent in the last section on how degree assignments in different contexts could be made more accurate over time. I chose this strategy of exposition because, when it comes to initial degree of belief assignments, I am inclined to agree with the skeptic and say that those assignments are not justified.

Let us first return to the beliefs we acquired through testimony as young children. When we were very young, we simply did not have the conceptual resources to make most of the evaluations outlined above. In fact, it would be a stretch even to suggest toddlers assign degrees of belief

\textsuperscript{40} Lewis himself was aware of this latter fact, for he made room in 2* for the belief in general conformity to be derived from past observations of the regularity in others.

\textsuperscript{41} The degree of belief apparatus does not particularly advance my inquiry in this section, but its advantages will become evident in the next sections.
because they tend to accept (and only rarely reject) everything they hear. It is only later on that they become more discriminating in their acceptance of information presented through testimony, and perhaps only accept information presented by those to whom they are emotionally attached. But even then, they are not doing much better because they would accept everything such people tell them.

Maybe we would like to say that these problems do not really concern us as adults because, unlike our childhood selves, we now have the capacities to make relevant evaluations. But this fact alone does not guarantee the accuracy of our degree of belief assignments. For instance we could we easily led astray if there is some piece of relevant information that our evaluations fail to take into account. This could be a hidden motivation on the part of the speaker, as when a typically reliable science researcher deliberately misconstrues her findings as a result of her secret desire to further the agenda of some pharmaceutical company. Or it could simply be some implicit convention that is not yet made known to us due to our lack of experience with the context where it occurs. But even if we are aware of all the relevant information, we might still give inappropriate weights to the different evaluators so that our resulting degree assignments become both unjustified and inaccurate.

Perhaps we could at least rest assure that it was some deficiency on our part which resulted in each of the above misassignment of values, so that if we were to correct those deficiencies, our initial degree assignments would be justified. But again we are disappointed, because we can still

42 There is a long-standing philosophical tradition that conceives of children as credulous. See for instance, Reid (1764/2000), Russell (1921), and Wittgenstein (1969).
43 Harris, 2012.
44 For example, if we think that the general reliability of the botanist is enough to guarantee the truth of her statement “rhubarb leaves are poisonous” even in the lying contest.
end up with inaccurate assignments in some instances even if we do the best we could as epistemic agents. Consider for example the Bohr model of atomic physics. When I was a student in high school, my chemistry teacher taught me the Bohr model, which construes atoms as combinations of small, positively charged nuclei and negatively charged electrons that travel in circular orbits around them. Suppose that when my teacher presented me with that model, I, like any good epistemic agent, evaluated it using the evaluation matrix. In terms of speaker-evaluations, I understood that my teacher was certified and was therefore quite a competent speaker on the subject. I also monitored the accuracy of his past utterances and discovered that he was generally quite reliable. Also in terms of the context-related evaluations, I was aware that the information was being presented to me in a formal educational context, which meant that my teacher was likely truthful in his presentation of the model. I therefore assigned it a high degree of belief, say [0.8,1].

Given that the Bohr model is actually an inaccurate model of atomic physics, it is obvious that my initial degree of belief was misassigned. However, there is still a clear sense in which my (mis)assignment was justified, for I did the best I could to evaluate the speaker and the context and had no reason to suspect that my teacher’s utterance might have been false. Indeed, from the point of view of my limited epistemic resources, I had every reason to believe that I was making a responsible and justified degree assignment. So let us call this assignment weakly justified, for it is the best I could have done at the time.

How did I get from this weakly justified, but misassigned value to a more accurate and strongly justified one? Well, first some doubt needs to be raised about my initial assignment, so that I
could realize that it might be mistaken. Therefore let us suppose that I had a friend who was a few years older than me and also learned the Bohr model in school. Suppose she says to me, “that model is not actually accurate. It is only an approximation designed to help you get started on thinking about the structure of atoms.” How would I react to this remark? The first thing I would do, if I were a responsible epistemic agent, is to evaluate it using the evaluation matrix. In doing so, I realize that my friend had been quite reliable about school-related topics in the past and makes the utterance in a sincere conversation between friends, which would lead me to assign it a high degree of belief. Although my friend’s statement contradicts what my teacher said previously, I am more inclined to believe her because she also provides a reasonable explanation as to why my teacher might provide false information. Hence I am most likely to respond to her statement by shifting the degree of belief I assign to the Bohr model towards 0 and assigning a high value to the additional statement “The Bohr model is a useful idealization of atomic structure”.

One could imagine many such instances of conflicting testimony as I make advances in the study of atomic physics. I will come across a range of testimonial accounts in lectures, books and publications that either speaks for or against the Bohr model, and for each account I will analyze it in accordance with the evaluation matrix. Over time, I will realize that the number of sources that speak against the Bohr model increases. If I find that each instance the objection to the Bohr model is believable, then I would be justified in moving my degree of belief towards 0. If I do not find it convincing, then I might put it in a separate mental file that I can refer to each time new evidence arises. Although the individual contents of this separate file might not be very believable in themselves, over time the contents might add up to collectively move my degree of belief...

\footnote{Especially after I start learning about quantum mechanics.}
belief towards 0. Thus these two processes would ultimately lead me to conclude that the Bohr model has a very low possibility of being true.

Justifying Final Belief Assignments

Our analysis of the Bohr model example provides important insights that could be used to construct a response to the skeptical worry presented at the beginning of this paper. The skeptic was concerned that if our initial assignments of the degrees of belief were unjustified, they would lead us to further errors in our learning so that we might end up with grave errors in our systems of beliefs. But in the case presented by the Bohr model, it seems that the result was the exact opposite: my initial misassignment was essential for the development of testimonial strategies that ultimately led to more accurate and justifiable beliefs.

Let us apply this insight to a more general case. As discussed several times, infants start off relatively receptive to the information presented to them through testimony. They, being so alien to their new environments, have a biological tendency to acquire as much information as they can from others around them, so their initial strategy might just be to accept everything they hear and understand. But as they receive more information, conflict arises either among different testimonial sources or between what they are being told and their existing beliefs. In light of these conflicts, children develop strategies to discriminate between sources of information so that they can simultaneously resolve the conflicts at hand and learn something about how similar conflicts could be resolved in the future. Over time, an increasing number of these strategies will be developed and perfected, so that they will end up with more reliable sets of beliefs as adults.
What are some primary strategies children use to make discriminations amongst conflicting testimonial sources? Corriveau and Harris had already pointed to one of these. In the case where children from two daycare centers were told conflicting names of objects by their own caregiver and the alternate caregiver, they overwhelmingly trusted the testimony of their own caregiver with whom they are more familiar. There is relatively little to say about this selectivity because it simply arises out of the biologically ingrained survival tendency to trust only those who are least likely to do us harm.

Interestingly, children’s familiarity-bias when it comes to choosing between sources of information diminishes over time, and is replaced by considerations about the speaker’s accuracy. Enlisting the help of the same children and caregivers we have seen above, Corriveau and Harris did an additional experiment. In this new experimental set up, the group is split into two, so that half the children in each center watched the familiar caregiver name some well-known objects accurately and the unfamiliar caregiver name them inaccurately; whereas the remaining children saw the reverse arrangement. When the participants were asked to choose between the two names again, children of different ages reacted differently. The three-year-olds were scarcely affected by the new setup: they continued to direct the majority of their choices to the familiar caregiver, regardless of whether or not she had been accurate. Four-year-olds, however, increasingly preferred the familiar caregiver when she had been accurate than when she had been inaccurate. This trend is exaggerated in the five-year-olds, who, although still preferring the familiar caregiver when she had been accurate, lost the preference altogether when

---

46 Harris, 2012.
47 That is, the familiar caregiver named the objects inaccurately, whereas the unfamiliar caregiver named them accurately.
she had been inaccurate and preferred the unfamiliar caregiver instead. The precise results are shown in the figure below:

The results of this additional experiment indicates that although younger children make discriminations between conflicting testimonies based on their familiarity with the speaker, it is supplemented by an accuracy-based strategy as they grow older. This adjustment of the strategy could be the result of several factors. First, children might discover through experience that even the testimony of familiar speakers could lead them astray. For example, a child who has lost her favorite toy might ask her caregiver where the toy is, to which the caregiver responds that it is on the sofa. But when the child accepts the testimony and actually goes to the sofa to get it, she may find that the toy is not there. The child might thus learn from this experience and others like it that the information provided by familiar speakers is not always accurate, and that certain familiar figures might provide more reliable information than others.48 Secondly, as children get

48 I take for granted in this essay children’s capacity to engage in autonomous reflection, a view which is supported by a long tradition in education research from Rousseau (1762/1999) to Piaget (1926).
older, they spend more time each day with unfamiliar people and as a result are exposed to more unfamiliar informants. With this expansion of testimonial horizons, a need might arise for children to listen to and learn from people they have no or little previous experience with, which makes it impossible for young children to discriminate based on familiarity alone. Hence they are forced to try out new strategies of discrimination such as monitoring of their speakers’ past accuracy.

As children are exposed to an increasing number of unfamiliar speakers, they might also realize the distribution of human knowledge is uneven and localized.\textsuperscript{49} They discover that although their immediate family members are the best people to ask about matters relating to the home, their teachers are better sources of information relating to things that go on at school. This is a process that continues into adulthood as they encounter more informants with varying areas of expertise, who they learn to recognize using increasingly specified markers such as the speakers’ levels of education, place of work, and even certain of their physical characteristics. As children become proficient in recognizing these competency based markers, they might also discover conventions governing speech in various testimonial contexts. These conventions could be explicitly told, as when parents teach children to always be sincere, or they might be implicitly discovered, as when children become convinced that whatever they are taught in school is always true. Through these processes, children learn to direct more attention to evaluations such as speaker competency and conversational conventions, which help them further discriminate between sources of information.

\textsuperscript{49} Harris, 2012.
In general, we could represent a child’s epistemic progress with regards to testimony as follows. She starts off with a poor strategy $S_0$ that assigns a degree of belief 1 to everything she hears and understands. Then she realizes that the beliefs she acquired using $S_0$ are inadequate, because they often contradict one another and lead her astray. So she responds in two ways: she modifies the degrees already assigned to those beliefs, and she adjusts her assignment strategy to prevent such problems from occurring in the future. In terms of the latter adjustment, the child might for example switch to a new strategy $S_1$ that assigns higher degrees of belief to utterances made by those she is familiar with, and lower degrees of belief with whom she is unfamiliar. This new strategy is better than $S_0$ because it resolves conflicts arising from contrary statements made by familiar informants and unfamiliar ones, and it prevents such conflicts from arising again. Indeed, when the child substitutes $S_1$ for $S_0$, she undergoes a justified revision of her strategy because the degrees of belief assigned using $S_1$ are more likely to be accurate than those assigned by $S_0$.

Despite $S_1$’s being a better strategy, it too will soon prove to be inadequate. The child is almost sure to discover through observation that the information provided by familiar speakers is fallible, and also learn from her experiences that the beliefs acquired from them are not comprehensive enough for all her epistemic needs.\(^5\) With this discovery, the child might then decide to employ another strategy $S_2$, which assigns higher degrees of belief to utterances made by speakers whose past accuracy she has observed and lower degrees to those made by inaccurate speakers. Note that, in this case, the addition of $S_2$ does not automatically nullify $S_1$, for children who have started monitoring speaker accuracy still show preferential treatment to

---

\(^5\) For example if she were alone and lost, she would not be able to draw on the testimony of familiar figures for directions.
familiar speakers.\textsuperscript{51} What changes with the addition of $S_2$ is that it lessens the weight that the child attaches to $S_1$, so that she has the freedom to reject statements made by familiar, yet inaccurate speakers. Again, the change in strategy from $S_1$ to $S_2^*$ ($S_2^*$ being some weighted combination\textsuperscript{52} of $S_1$ and $S_2$) is a justified revision because $S_2^*$ allows the child to assign more accurate degrees of belief.

Once the child begins to recognize competency-related markers and context-specific conventions, she will likely revise her strategy in exactly analogous manners. Suppose she starts by assigning and modifying degrees of belief using some strategy $S_i$, and discovers through personal experience\textsuperscript{53} that $S_i$ sometimes leads her to assign incorrect degrees of belief. The child is thus motivated to make revisions to her existing strategy, which leads her to replace $S_i$ with another strategy $S_{i+1}$ (this is a weighed combination of $S_i$ and some new method of assignment that rectifies the incorrect degrees assigned by $S_i$). If it were discovered over time that the child’s use of $S_{i+1}$ is more likely to generate accurate degree of belief assignments than her previous use of $S_i$, then her change of strategy would be justified.\textsuperscript{54}

Suppose the child continues to make justified revisions of strategy every time she discovers a fault in her existing strategy, then over time her revised strategies are likely to yield increasingly

\textsuperscript{51} Such as the four-year-olds in Corriveau and Harris’s latter experiment.
\textsuperscript{52} The weights attached to $S_1$ and $S_2$ respectively differ from person to person, and could change over time. For instance, the four-year-olds seem to still attach quite some weight to the familiarity evaluation whereas the five-year-olds seem to attach less.
\textsuperscript{53} Such as perceptual verification and the testimony of others.
\textsuperscript{54} i.e. it is a justified revision of strategy. Note that this justification is distinct from the justification of the child’s beliefs acquired using her strategies, because the latter depends on the reliability (or justifiability) of the strategies.
reliable assignments.\textsuperscript{55} Initially the child would have only encountered certain testimonial situations and not others, so this heightened reliability might be unevenly distributed across utterances made by different speakers and in different contexts.\textsuperscript{56} \textsuperscript{57} But by the time the child becomes a mature adult, she would have gone through a lengthy process of justified revisions, the resulting strategy of which has a very high reliability of yielding accurate assignments in a vast variety of testimonial situations. Therefore the beliefs that mature adults acquire through testimony are justified (and likely to be true) not only because they are arrived at through reliable assignment strategies, but also because the strategies themselves are the results of lengthy processes of justified revisions through which the reliability of the final strategies are guaranteed.

The preceding account constitutes a full response to the skeptical challenge in two ways. First it suggests, contrary to the skeptic, that at least some of the beliefs\textsuperscript{58} we acquired through testimony are justified, which mitigates the worry that unjustified foundational beliefs acquired through testimony are going to infect our entire systems of beliefs. In addition, the developmental account proposes a revision process that could only be made possible by the initial acquisition of false beliefs\textsuperscript{59}. So the very beliefs upon which the skeptical challenge is based becomes tools that epistemic agents must use to ultimately arrive at more reliable

\textsuperscript{55} Compare this with the process of second-language acquisition. A language learner usually starts out making many errors in the target language, but she adjusts her strategy in accordance with those errors so she makes fewer errors. Eventually, she develops her intermediate linguistic strategy (what linguists call “interlanguage”) so that it increasingly approximates the target language. When this approximation becomes close enough, the language learner could then be said to be proficient in the target language. For more on this see Pitt Corder’s “The Significance of Learners’ Errors” (1967) and Larry Selinker’s “Interlanguage” (1972).

\textsuperscript{56} Recall that revisions are made in response to inaccuracies that the child discovers in her existing strategy, so they could only affect testimonial situations similar to those the child has encountered before.

\textsuperscript{57} This phenomenon is similar to that of the collective process of the sciences solely yielding truths in select areas of inquiry at each time. See p.9

\textsuperscript{58} That is, those made in testimonial situations similar to what we have encountered before.

\textsuperscript{59} Or more precisely, incorrect degree of belief assignments.
assignment strategies and justified beliefs. For these reasons therefore, I am inclined to say that the skeptic about testimony is misguided. I do know my date of birth after all.

DOROTHY CHEN
BIBLIOGRAPHY


