Understanding Assertion:
Lessons from the False Belief Task
Joseph Shieber

Abstract: This paper uses recent research in developmental psychology regarding the acquisition of the concept of belief in young children to explore the contrast between a disposition-based account of the principles underlying linguistic communication and the representative and highly influential intention-based accounts of assertional practice advanced by David Lewis and Donald Davidson. Indeed, evidence from recent work in developmental psychology would seem to suggest that disposition-based accounts are not only possible accounts of the acquisition of competence in assertional practice, but are in fact better than their rivals in explaining the way such competence is actually acquired.

Clearly, most of our acquisition of knowledge through verbal communication rests upon our sharing a language with others. That our interlocutors mean what we take them to mean is not something that we normally establish via argument; rather, it is an unspoken assumption upon which any information acquisition through our communications with others rests.

In order to pursue some of the implications of this fundamental fact regarding the nature of communication for our understanding of the structure of communication, it will be useful to cite a passage from the 18th century Scottish philosopher of Common Sense, Thomas Reid:

… a difficulty occurs which merits our attention, because the solution of it leads to some original principles of the human mind, which are of the
greatest importance, and of very extensive influence. We know by experience that men have used such words to express such things; but all experience is of the past, and can, of itself, give no notion or belief of what is future. How come we, then, to believe, and to rely upon it with assurance, that men, who have it in their power to do otherwise, will continue to use the same words when they think the same things? Whence comes this knowledge and belief—this foresight, we ought rather to call it—of the future and voluntary actions of our fellow creatures? (Reid [1868], 196)

According to Reid, linguistic communication requires that one trust in the foundation of a shared language, and that this trust is not one that can be grounded by means of independent argument: as Reid continues, “there is, therefore, in the human mind an early anticipation, neither derived from experience, nor from reason, nor from any compact or promise, that our fellow-creatures will use the same signs in language, when they have the same sentiments.” (Reid [1868], 196; italics mine)

Reid puts his point in terms of our inability to provide some argument capable of grounding our trust in the voluntary actions of our interlocutors to refrain from, as Reid puts it somewhat later within the same passage, “equivocation or falsehood.” The core insight behind Reid’s comments, however, is more fundamental. In order for there to be shared languages at all, there must be common dispositions to linguistic behaviors.

More recently, David Lewis—although seemingly unaware of Reid’s earlier discussion—has made strikingly similar claims concerning the behaviors underlying linguistic practice. Indeed, Reid identifies the same two core principles underlying shared

---

1 First, in his [1969], where Lewis emphasizes his debt to Eric Stenius’s [1968]. Lewis modifies his account—developing it in the form to which I refer here—in his [1975].
linguistic institutions to which Lewis also points. Reid refers to these principles as innate dispositions, *the Principle of Veracity*, “a propensity to speak truth, and to use the signs of language so as to convey our real sentiments,” (Reid [1868], 196) and *the Principle of Credulity*, “a disposition to confide in the veracity of others, and to believe what they tell us.” (Reid [1868], 196) In contrast, Lewis treats these principles as conventions of truthfulness and trust, where, as Lewis puts it, “to be truthful in \( L \) is to … try never to utter any sentences of \( L \) that are not true in \( L \) [and] … to be trusting in \( L \) is to … tend to respond to another’s utterance of any sentence of \( L \) by coming to believe that the uttered sentence is true in \( L \)” (Lewis [1975], 167)

In this paper we will contrast the Reidian disposition-based account of the principles underlying linguistic communication with representative intention-based accounts of assertional practice, influenced by the work of H. P. Grice, in light of recent research in developmental psychology regarding the acquisition of the concept of belief in young children. In the first two sections, we will introduce two such intention-based approaches: David Lewis’s convention-based account, and the interpretationist approach of Donald Davidson. In the third section, we will present evidence from recent work in developmental psychology to the effect that the Reidian disposition-based approach is not only a possible account of the acquisition of competence in assertional practice, but is better than its rivals in explaining the way such competence is actually acquired. Finally, in the fourth section, we will consider the extent to which—far from supporting intention-based approaches, as has been suggested by Bennett [1976]—rationality considerations in fact support the disposition-based account of assertional practice introduced by Reid.

1. Lewis on Linguistic Conventions

   Above, we cited the work of David Lewis on the conventions governing linguistic practice and noted the significant overlap between the content of Lewis’s conventions of truthfulness and trust in \( L \) and Reid’s principles of Veracity and Credulity. Unfortunately,
however, Lewis’s reliance on the notion of convention makes his program unsuitable as a foundation upon which to build a tenable account of the nature of assertional practice. In order to see this, it is necessary to look more closely at Lewis’s work on the nature of conventions.

According to Lewis, a regularity $R$ in action (or action and belief) is a convention in some population $P$ if, and only if, that population conforms to the following six conditions:

1. All members of $P$ conform to $R$.
2. All members of $P$ believe that all other members conform to $R$.
3. The belief that others conform to $R$ gives each member of $P$ a decisive reason to conform to $R$ herself.
4. There is a preference, among the members of $P$, for general conformity to $R$ rather than slightly-less-than-general conformity to $R$.
5. There is at least one other alternative regularity $R'$ such that: (a) the members of $P$ cannot normally conform to both $R$ and $R'$, and (b) if there were general to $R'$, then conditions (3) and (4) would be true of $P$ and $R'$.
6. It is common knowledge among the members of $P$ that conditions (1) to (5) are true of $P$ and $R$. That is, “everyone must potentially know that (1) to (5) hold; potentially know that the others potentially know it; and so on.” [Lewis (1975), 7; my italics]

Given this understanding of convention, we may assess Lewis’s thesis that “a language $L$ is used by a population $P$ if and only if there prevails in $P$ a convention of truthfulness and trust in $L$, sustained by an interest in communication.” [Lewis (1975), 11] To the extent that

---

2 Cf. Lewis [1969] and Lewis [1975].
this account of language use is intended to capture the conventions governing the speech act of assertion, this account differs in fundamental ways from the disposition-based account sketched by Reid.

Before presenting ways in which Lewis’s account differs from that of Reid, it is important to underscore one way in which the two accounts are not opposed. According to Lewis, the linked conventions of truthfulness and trust in L hold in a population P only if it is common knowledge within the population P that—to take condition (3) as a representative one—the belief that others conform to the conventions of truthfulness and trust gives each member of P a decisive reason to conform to the conventions of truthfulness and trust herself. Thus, on Lewis’s account, possession of such common knowledge is a necessary condition for conformity to the conventions of truthfulness and trust, and it may seem that no such necessary condition would be a component of Reidian disposition-based account. The rationale behind the requirement of such a condition, according to Lewis, is that it “ensures stability. If anyone tries to replicate another’s reasoning, perhaps including the other’s replication of his own reasoning, …, the result will reinforce rather than subvert his expectation of conformity to R.” (Lewis [1975], 7) This is indeed a powerful rationale, and it would be a mark against Reid’s account if it were to diverge from Lewis’s account on this point.

Happily, there is no reason for Reid’s account to deny that it is common knowledge within a population that the principles of veracity and credulity hold within that population. For Lewis’s notion of common knowledge is of a species of potential knowledge, “knowledge that would be available if one bothered to think hard enough.” (Lewis [1975], 7) If Reid’s account is correct, however, then it would be reasonable to suppose that all participants in a practice of assertion would implicitly recognize the principles of veracity and credulity: they would be disposed to sanction failures to

---

comply with the principles—even their own such failures—and to consider compliance with the principles as appropriate, and indeed natural, behavior. If one were then to attempt to replicate the reasoning of another, e.g., by putting oneself in their shoes, one would automatically reason, from the point of view of the other, as if such compliance were natural for her as well. Insofar as this is an accurate account of the way such reasoning would be used in replicating the thought processes of the other in making and/or attending to assertions, knowledge that others act in compliance with the principles of veracity and credulity is available to those who think hard enough. If this is the case, then such knowledge can also be used to reinforce one’s own compliance with the principles of veracity and credulity.

The central distinction between Reid’s account and Lewis’s account is signaled by conditions (2) and (3) in Lewis’s characterization of assertion. With respect to the conventions of truthfulness and trust, these conditions require that participants in a practice of assertion have certain beliefs regarding others’ compliance with the regularities of truthfulness and trust, and that those beliefs provide decisive reasons for participants themselves to comply with those regularities. In contrast with Lewis’s convention-based account, Reid’s account specifically denies that having beliefs regarding the compliance of participants in the principles of veracity and credulity is a necessary condition for participation in such a practice, and thus also denies that one’s reasons for complying with the principles of veracity and credulity would derive from having such beliefs. Rather, participants in the practice of assertion, on Reid’s account, implicitly acknowledge those norms by structuring their actions and behavior in accordance with those norms—i.e., without forming beliefs regarding the nature of those norms, or the compliance in those norms by others.

2. Davidson on Belief and Interpretation

Perhaps the best-known account of linguistic competence presented in the last thirty years is the interpretationist account
proposed by Donald Davidson in a series of influential articles.\footnote{Most of which are collected in his [1980].} According to Davidson, it is both a necessary and sufficient condition for having beliefs that one assume the role of a \textit{radical interpreter}. That is, according to Davidson, one has beliefs just in case one interprets the public linguistic communications of another by constructing a theory of the other’s linguistic behavior, based upon attributions of relations of holding-true between speakers and sentences, according to the following form. The theory will consist of sentences (known as T-sentences) of the form:

\[ S \text{ is true (in the object language) if and only if } p, \]

where ‘$s$’ is to be replaced by a canonical description of $s$, and ‘$p$’ by a translation of $s$ in the metalanguage. The resulting theory, a T-theory for a given speaker, provides a theory of interpretation of that speaker with respect to a given metalanguage. Thus, on Davidson’s account, one understands a given speaker only if one can construct such a T-theory for that speaker. Of course, constructing such a T-theory requires that one already understand some language—i.e., the metalanguage that figures on the right-hand side of the T-sentences; thus, understanding a language is also a necessary condition for constructing a T-theory. Thus, being able to construct a T-theory is both a necessary and sufficient condition for being a linguistic agent.

To see that Davidson’s theory, like Lewis’s, also requires of linguistic agents that they explicitly attribute beliefs to other participants in a practice of assertion, it is enough to recognize that Davidson’s theory of radical interpretation requires that one be able to attribute, to their interlocutors, attitudes of holding-true (with respect to target sentences).\footnote{Cf. Davidson [1975], 170.} Such attributions are, essentially, attributions of belief. Thus, it is a necessary condition of understanding another, on the basis of Davidson’s notion of radical interpretation, that one ascribes beliefs to one’s interlocutor. Thus, Davidson writes:
[a]nyone who understands speech can recognize assertions and knows that someone who makes an assertion represents himself as believing what he says. Similarly, someone who says to Jones that snow is white knows the truth conditions of 'Jones believes that snow is white' (even if he does not know English nor have a way of expressing belief). … Having a belief demands in addition appreciating the contrast between true belief and false, between appearance and reality, mere seeming and being.⁶

Note that Davidson’s point here is not merely that it is a necessary condition of understanding another’s linguistic performances that one attribute beliefs—or, more properly, intentions to express beliefs—to one’s interlocutor. Rather, Davidson is also claiming that, in order to have beliefs oneself—and thus, to the extent that the principle of truthfulness governs assertion, to be capable of producing linguistic performances—one must be able to interpret others via the construction of a T-theory of their performances. Thus, it is a consequence of Davidson’s position that it is a necessary condition of linguistic competence that one form beliefs about the beliefs and intentions of one’s interlocutors.⁷

3. Support for Reid’s account From Developmental Psychology

We will not attempt here to argue that it would be conceptually impossible for linguistic agents to behave in ways comporting with the view according to which—beyond complying with the principles of veracity and credulity—one demonstrates one’s assertional competence by forming beliefs regarding the beliefs or intentions of others. It seems possible that there could be linguistic

⁶ Davidson [1996], 161-2.

⁷ For further emphasis of this point, that it is a necessary condition for linguistic competence that one possess the ability to form beliefs about the beliefs of others, cf. Davidson [1985].
agents who always interpret others by forming beliefs regarding their beliefs and intentions. I wish to argue that it cannot be a necessary condition for the possession of assertional competence that one form such beliefs.

Now, it is a necessary condition for possession of the concept of belief that one understand—and be able to attribute to others—false beliefs. To the extent that the Davidson quote above—i.e., that “having a belief demands in addition appreciating the contrast between true belief and false, between appearance and reality, mere seeming and being”—is plausible, its plausibility rests on a confusion between what is required for having the concept of belief and having belief, simpliciter. Of course, it is true that having the concept of belief requires having the concept of false belief, but it is false that having a belief requires having a concept of true belief and false belief. Thus, Davidson is correct to the extent that he identifies a conceptual connection between belief and false belief, but incorrect to the extent that he claims that one must grasp such a connection in order to have beliefs at all.

The reason for the conceptual connection between belief and the possibility of falsehood is that belief, unlike knowledge, is intentionally opaque. That is, it is part of the concept of belief that it is possible that one can believe $p$ without its being the case that $p$. If one cannot ascribe false belief, then one demonstrates that one does not recognize this possibility, and, thus, that one does not grasp the concept of belief itself. Thus, if it is necessary for the possession of assertional capacities that one be able to ascribe beliefs to others, then it is necessary for the possession of assertional capacities that one be able to ascribe false beliefs to others.

Given these considerations, it is clear that if it were to be a necessary condition for the possession of assertional competence that one form beliefs regarding the beliefs of others, it would be necessary for the possession of assertional competence that one possess the concept of belief. I will argue that it is not only possible that one could possess assertional capacities without one’s being able to ascribe false beliefs to others, but that it is actually the case that a class of humans possess significant assertional capacities without being
able to ascribe false beliefs to others. (We will follow developmental psychologists in describing the ability to make such ascriptions as the ability to perform the false-belief task.)

There is extremely strong empirical evidence that children, up to the age of well beyond 3 1/2 years, are able to make and receive assertions, without being able to perform the false-belief task—and thus without being able to form beliefs regarding the beliefs of others. If this is correct, then the best empirical evidence from current developmental psychology would suggest that intention-based accounts are wrong to posit the ability to form beliefs concerning the beliefs of others as a necessary condition for the possession of assertional capacities.

Evidence regarding the abilities of children to perform false-belief tasks stems from the initial work of J. Perner and H. Wimmer. In their first experiment, Perner and Wimmer had children from three age groups—ages 3-4, 4-6, and 6-9—watch a puppet show involving a boy named Maxi. While the children watch the puppet show, a storyteller describes the scene as the story unfolds. Maxi is in the kitchen with his mother. The kitchen has two cupboards, blue and green. The children watch as Maxi puts a piece of chocolate in the blue cupboard, after which they see him leave the room. As Maxi leaves, the storyteller emphasizes that Maxi remembers where he puts the chocolate, and that he is looking forward to eating it when he returns to the kitchen. While Maxi is gone, the children see his mother take the chocolate out of the blue cupboard, use a bit of it to bake a cake, and put the remaining chocolate in the green cupboard. A short time later, the children see Maxi return to the kitchen, and are told that he wants to eat his chocolate. At this point, the experimenters ask the children the BELIEF-question: “Where will Maxi look for the chocolate?” In the initial experiment, Perner and Wimmer found that none of the 3-4 year-olds, 57% of the 4-6 year-olds, and 86% of the 6-9 year-olds answer the question correctly. Furthermore, those who answer the BELIEF-question incorrectly commit a systematic error; they assume that Maxi will look for the chocolate in the green cupboard, as opposed to the blue one.

One initial suggestion for this striking result might be that the
children themselves either no longer remembered where the chocolate is, or no longer remembered where Maxi initially put the chocolate. However, to rule out such alternate interpretations, Perner and Wimmer asked the children two additional questions: (1) the REALITY-question, “Where is the chocolate really?” and (2) the MEMORY-question, “Where did Maxi put the chocolate in the beginning?” Even those children—including the 3-4 year-olds—who answered the BELIEF-question incorrectly answered the REALITY- and MEMORY-questions correctly. That is, the children know where the chocolate is and they remember where Maxi put the chocolate, but many of them—including all of the 3-4 year-olds—fail the false-belief task. This experiment has now been reproduced numerous times, with strikingly consistent results.

The standard interpretation of the results of such experiments is that children at the age of 3-4 years are unable to attribute to Maxi the false belief that the chocolate is to be found in the blue cupboard. Furthermore, the widely-accepted implication of this result is that children at the age of 3-4 years are unable to form beliefs concerning the false beliefs of others. Thus, Wimmer and Perner write:

[A] novel cognitive skill seems to emerge within the period of 4 to 6 years. Children acquire the ability to represent wrong beliefs and to construct a deceitful or truthful utterance relative to a person’s wrong beliefs. Within this period several other related abilities also emerge: children start to understand another person’s absence of knowledge …. They become able to construct a deceitful utterance … and to infer a

---

8 H. Wimmer and J. Perner [1983].

9 Thus, Wellman writes that, “[b]eginning with Wimmer and Perner (1983) more than 20 demonstrations now exist that three-and-a-half-year-olds generally fail false-belief tasks … [and] do so even when great care has been taken to make the task easy, comprehensible, and straightforward ….” Cf. Wellman [1990], p. 260.
deceptive plan from a critical utterance in the context of conflicting goals […].

This interpretation of the implications of the experimental data is also almost universally accepted.

Although this interpretation of the implications of the experimental data is almost universally accepted, there is one significant exception to this unanimity with which we must deal. According to Gopnik, Wellman, and a few others, the results of the Wimmer/Perner experiment, and subsequent similar experiments, imply merely that children before the age of four do not attribute false belief, and hence do not attribute belief, to others. However, so the suggestion goes, children before the age of four do attribute beliefs* to others, where beliefs* are taken to be proto-beliefs that are intentionally transparent.

Thus, Gopnik and Wellman, describe the abilities of pre-4 year-olds as follows:

three-year olds’ first understanding of belief seems like their earlier understanding of perception in that it shares something of that construct’s non-representational character. Specifically, belief does not at first easily encompass a sense of misrepresentation. On this view, belief, like perception and desire, involves rather direct causal links between objects and believers. … The similar idea in all these accounts is that belief contents directly reflect the world. …

[T]hese same children do not often construe actions as stemming from false beliefs. When predicting action they typically, consistently, resistsly act as if

---

10 Wimmer and Perner [1983], 126.

11 Others expressing their acceptance of the Wimmer/Perner interpretation include Fergusson and Gopnik [1988], Flavell [1988], and Gopnik and Wellman [1995]. Some opponents of the view include Bartsch and Wellman [1988] and Chandler [1988].
the actor’s desire along with the objective facts
determine action, ignoring a role for false belief in
influencing action …. Similarly, when asked the
contents of a person’s belief they consistently,
resistantly cite the facts …. In short, when
predicting action and when diagnosing belief
contents, three-year-olds evidence largely a non-
representational desire-perception understanding.12

According to Gopnik and Wellman, three-year olds, like older
children form beliefs concerning the mental states of their
interlocutors. The difference between the younger children and the
older children lies merely in the fact that the younger children employ
the concept of beliefs* in their reasoning concerning their
interlocutors, whereas older children employ the concept of beliefs.
On the account proposed by Gopnik and Wellman, the achievement
of four-year olds—from the perspective of their three-year old
counterparts—is that of a Kuhnian scientific revolution; the four-year
olds discover that the concept of belief is a far more powerful one,
for the purposes of theorizing about others, than that of belief*.

Before criticizing the Gopnik/Wellman approach, it is
important to note that the existence of such a rival position does not
refute the possibility of the standard interpretation, according to which
children before the age of three do not form beliefs regarding the
mental states of their interlocutors at all. As long as the standard
interpretation is possible, however, this will be enough for the
purposes of our argument here. Nevertheless, it would exaggerate
the plausibility of the Gopnik/Wellman approach to suggest that
such an alternative should reduce the status of the standard
interpretation to that of mere possibility.

We will briefly sketch two chief criticisms of the
Gopnik/Wellman position. The first is that the Gopnik/Wellman
approach fails to account plausibly for the rapidity of the shift in
abilities to perform the false-belief task, and the universality of the

observed phenomenon. If, as Gopnik and Wellman suggest, children are theorizing in the way that scientists do, one would expect at least a small minority of the under-four-year olds to be able to perform the false-belief task early, reflecting the fact that some of the children presumably have better reasoning skills than others. The data, however, indicate a remarkably universal inability to perform the false-belief task among pre-four-year olds. Second, simplicity considerations suggest that, if there is an explanation of the data that does not require the assumption that children are engaging in highly sophisticated theoretical activities, such an explanation should be preferred to the Gopnik/Wellman account. But there is such an explanation: it is the standard account of why under-four-year olds fail the false-belief task.

If the standard explanation of the implications of the Wimmer/Perner—and subsequent—experiments is correct, and if our considerations concerning the conceptual capacities required to attribute beliefs to others are correct, then it follows that virtually all pre-four-year old children are incapable of forming beliefs concerning the beliefs of others. However, children at those ages evidence an impressive ability both to make assertions and to consume assertions.

Thus, e.g., Ninio [1995] reports that, by the age of 22 months, children are able to express an average of over 30 communicative intents, including the following:

1. Statement discussing a joint focus of attention.
2. Statement discussing a recent event.
3. Statement discussing a past or future event.
4. Agreeing with a proposition.
5. Disagreeing with a proposition.
6. Correcting an utterance.
7. Yes-no question requesting clarification of an
Foster [1986] has noted that, by the age of 1 1/2 years, children are capable of initiating conversation about absent or intangible things. Choi [1991] found that children Korean children under the age of 2 years were already capable of using different sentence-ending suffixes depending on whether the utterances they produced were statements or requests. Brown [1980] reports that some 30-month olds are capable of producing as many as 20 related utterances in sequence.

Indeed, the design of the Wimmer/Perner experiment itself provides a great deal of evidence of the linguistic competence of under-four-year olds. The children clearly understand the storyteller, respond appropriately to the questions of the experimenters (i.e., by attempting to answer them sincerely), and understand the situation presented—as evidenced by their correct responses to the REALITY- and MEMORY-questions. That is, children under the age of four both exhibit a great deal of ability to make and consume assertions and an utter inability to ascribe beliefs to others. Thus, not only it is possible that some assertionally competent individuals do not ascribe beliefs to others; indeed, a large number of assertionally competent individuals actually do not ascribe beliefs to others.

It is important to emphasize that the standard explanation of the failure of under-four-year olds in false-belief tasks has no implications for the sensitivity of those children to the sorts of norms involved in Reid’s account of assertional competence. Indeed, that those children comport with the primary norms—that of [TRUTH] and the principle of veracity with respect to asserting, and of the principle of credulity with respect to attending—is underscored by the experimental data adduced above. The sorts of linguistic behavior evidenced by the production of the statement-types (1) – (7) in the list excerpted from Ninio’s [1995] research underscores this conclusion, as do the responses of the children described in the Wimmer/Perner experiment. Indeed, the very fact that the children demonstrate linguistic competence in their respective languages

13 For list, cf. Hoff [2001].
supports the notion that they are sensitive to a host of norms governing their behavior. Thus, the results of current work in developmental psychology regarding the ability of children to perform false-belief tasks, although it has largely gone unremarked by philosophers of language, has serious implications for the dominant intention-based approach to the explanation of assertional competence. If the considerations adduced in our discussion are correct, the Wimmer/Perner experiment provides support for Reid’s account—as opposed to the dominant intention-based account—by demonstrating that the compliance of the children with the principles governing linguistic behavior has to be explained without an appeal to abilities to form beliefs regarding the beliefs of others.

4. Rationality and Assertional Practice

The empirical evidence presented in the previous section provides extremely strong support for the idea that accounts of assertional capacities that require linguistic agents to be able to ascribe beliefs to others are untenable. However, it would be possible to salvage such accounts were one able to reject evidence regarding the linguistic capacities of children as irrelevant to the question of the conditions required for full linguistic competence. Indeed, this was the strategy pursued by Hume in responding to the Reidian disposition-based account of linguistic understanding. Thus, it should be unsurprising that supporters of intention-based accounts of linguistic competence have appealed to similar considerations in support of their positions.

Lewis himself—although writing before the particular research with respect to the inability of children to perform false-belief tasks had been conducted—anticipated a potential challenge to his convention-based account resulting from considerations regarding the linguistic abilities of young children. Lewis writes:

Perhaps the use of language by young children is not a rational activity. Perhaps it results from habits which would not be overridden if they ceased to serve the agent’s desires according to his beliefs. If that is
so, I would deny that these children have yet become party to conventions of language, and I would deny that they have yet become normal members of a language-using population. Perhaps language is first acquired and afterward becomes conventional. … I am not concerned with the way in which language is acquired, only with the condition of a normal member of a language-using population when he is done acquiring language. (Lewis [1975], 23)

Thus, Lewis appeals to notions governing rational activities to suggest that his account of linguistic activity as governed by conventions cannot be disconfirmed via the appeal to evidence involving the merely habit-based linguistic abilities of young children.

Unfortunately, however, it is not clear how Lewis’s argument is supposed to function as a defense of his convention-based account against the evidence presented in the preceding section—and against the alternative account represented by the principles of veracity and credulity. The evidence presented in the previous section suggested that young children are incapable of forming beliefs about the beliefs of their interlocutors. However, the evidence from Ninio [1995] to the effect that 22-month olds are capable of disagreeing with propositions and correcting utterances would suggest that young children, prior to possessing the concept of belief, are sensitive to the goals of assertional practice. If this is correct, it is not clear what the cash value of Lewis’s requirement that a rational activity be one involving “habits which would … be overridden if they ceased to serve the agent’s desires according to his beliefs” is in the case under consideration. If Lewis’s characterization is intended to capture the extent to which linguistic agents can, for example, withhold belief in cases in which they have countervailing evidence, then the evidence suggesting that 22-month olds can correct utterances and reject propositions would suggest that they are engaged in a rational activity. That is, 22-month olds are capable of correcting or rejecting those utterances that the child knows to violate the principle of veracity. It is crucial to note, however, that children can satisfy this requirement for rational activity without being able to attribute beliefs
to others, and thus without satisfying Lewis’s convention-based account of linguistic activity.

Indeed, it is difficult to see how one could characterize the notion of linguistic competence as a rational activity in such a way as to essentially involve the ability of agents to form beliefs about beliefs. The option that suggests itself would be one requiring that—at least in the case of rational activities involving the production of beliefs—such activities could be overridden if they ceased to serve an agent’s desires according to his beliefs regarding beliefs.

Note, however, that in order to make the assessment one that essentially involves second-order beliefs about beliefs, it would seem necessary to think of the assessment as one involving universal generalization or some sort of modality operator. That is, such a condition on rational doxastic activity would seem to involve the possibility of generalizing over the outputs of activities so as to assess their propensities—as types—for producing true beliefs, and of opting out of such activity types that fail to meet appropriate standards. If, however, this is the characterization that Lewis intends—i.e., one such as to involve the possibility of opting out of the practice of assertion, tout court—then it is not clear that even most adults’ linguistic practices would involve “rational” activity, as understood under this interpretation of Lewis’s requirement. Thus, consider as a parallel the doxastic activities of believing on the basis of perceptual or memory experience. On this account of rational activity, neither activity would count as rational for the vast majority of adults.

Thus, we see that the most plausible reading of the notion of “rational activity” is not one to which Lewis can appeal in order to rule out pre-four-year old children as participants in linguistic practices qua rational activities. If this is so, however, then the challenge of the empirical evidence presented in the previous section would seem as pressing as ever for those who wish to maintain that the ability to form beliefs about beliefs is a necessary condition for participation—even full, rational, participation—in an assertional practice.
Another attempt, along the lines sketched by Lewis, to link participation in assertional practices to a robust notion of rationality in such a way as to rule out cases that pose problems to intention-based conceptions of linguistic activity may be found in Bennett [1976].^14 Bennett allows that there may be “careless” or “slow-witted” hearers that do not form beliefs concerning the intentions of an utterer $U$ to communicate a proposition $P$ by uttering a sentence $S$, but who nevertheless come to belief $P$ on hearing $S$ by acting according to the rule, “whenever $S$ is uttered, $P$ is true.” But, Bennett argues,

[g]iven time to reflect [such a hearer] will no doubt realize that he ought not to infer $P$’s truth unless he thinks $U$ intended to communicate $P$; but in the heat of the communicative moment [he] might blunder, and fall back on (i) [whenever $S$ is uttered $P$ is true] as though it were basic, thus inferring $P$’s truth in a non-Gricean manner. (Bennett [1976], 194)

Bennett here is suggesting that it would be a “blunder” to move from hearing a sentence $S$ to believing $P$, without first considering positive evidence to support the idea that $U$ in fact intends to communicate $P$ by uttering $S$. That is, Bennett takes it to be a dictate of rationality that one believe what one hears only after having gathered positive evidence with respect to the beliefs and intentions of one’s interlocutors. As Millikan has emphasized, (Millikan [1984], 66) such

---
^14 In characterizing Bennett’s dialectic in this way, I am glossing over subtleties in his presentation that are not pertinent to the discussion here. In fact, Bennett accepts that a theory of linguistic behavior cannot require that linguistic agents be capable of attributing beliefs and intentions to others, and argues merely for the weaker conclusion that fully rational adult human linguistic behavior ought to be characterized as involving the attribution of beliefs and intentions to others. It is this point that will serve as the focus of our discussion here. For Bennett’s extremely thoughtful account of these issues, cf. his [1976], 192-210.
a dictate amounts to the claim, “[t]he rational man … should take no shortcuts.” Thus, it would be appropriate to submit this general claim to greater scrutiny.\textsuperscript{15}

Consider a hearer $H$ who possesses the following beliefs:

- **[RELIABILITY]** Usually when $S$ is uttered then $P$.

- **[INSINCERITY]** It is not the case that usually when $S$ is uttered without the intention of communicating $P$, then $P$.

Thus, the following beliefs are also rational for $H$:

- **[DEPENDABILITY]** Usually when $S$ is uttered then it is with the intention of communicating $P$.

- **[TRUTHFULNESS]** Usually when $S$ is uttered with the intention of communicating $P$, then $P$.

The question we are considering is whether Bennett is right to suggest that whenever $H$ believes $P$ on the basis of an utterance $S$, the rational course of action for $H$ is to ground her belief in an inference of the following form:

\begin{align*}
(A) & \quad U \text{ uttered } S. \\
(B) & \quad U \text{ had the intention of communicating } P. \\
(C) & \quad \text{[TRUTHFULNESS]} \\
\text{Thus, } & \quad (D) \quad P.
\end{align*}

Alternately, of course, in those cases in which $H$ has discovered not (B) but (B'), viz., $U$ did not intend to communicate $P$, the rational course of action will be to inhibit the inference from (A) to (D) via [RELIABILITY] by considering the countervailing evidence presented by (B') in conjunction with (A) and [INSINCERITY].

\textsuperscript{15} For the argument in the following paragraph I am indebted to the presentation in Millikan [1984], 66-7.
crucial point is that Bennett must hold that in every case the rational course of action is to consider the beliefs and intentions of one's interlocutor in forming beliefs on the basis of her utterances.

There are three cases that may be considered when evaluating the appropriateness of such a claim:

1. The only sort of evidence that some hearer $H$ ever gets regarding whether $U$ intends to communicate $P$ relies on [DEPENDABILITY]; i.e., it always involves $H$'s experiences of $U$'s uttering $S$.

2. $H$ has some way of independently determining whether $U$ in fact intends to communicate $P$. In a particular situation $s$, $H$ has used this method to discover that $U$ does not intend to communicate $P$.

3. $H$ has some way of independently determining whether $U$ in fact intends to communicate $P$. In a particular situation $s^*$, $H$ has acquired no information one way or another regarding whether $U$ intends to communicate $P$.

It seems clear that in situations like (1), there would be no point in $H$'s engaging in an inference of the type Bennett envisages. For in such cases the only evidence that $H$ has for believing that $U$ intends to communicate $P$ is the same as her evidence for believing $P$ itself. Thus, the rational course of action in (1) would be for $H$ to infer $P$ directly upon hearing $U$ utter $S$.

Situations like (2) seem equally clear. In such cases, the rational course of action would be to block the inference to $P$ via the employment of the principle [INSINCERITY].

The true test cases involve situations like (3). Bennett must argue that, in such situations, it is irrational for $H$ to believe $P$ on the basis of $U$'s uttering $S$. Rather, Bennett must claim that the onus is on $H$ first to find some evidence as to $U$'s intentions if $H$ is to act rationally in believing $P$. Why stop there, however? Certainly, there will be conditions under which the evidence upon which $H$ standardly relies in determining $U$'s intentions will itself be unreliable. If $H$ is to be rational, would she not have to rule out such cases.
before relying upon her evidence regarding U’s intentions? In this
case too, though, there will be further conditions under which the
evidence upon which H relies in ruling out those earlier conditions
will itself be unreliable. So again H will have to rule out *those*
conditions before being able rationally to rule out the earlier
conditions. And off we are on an infinite regress of reasons.

In this respect, understanding linguistic communication—as
Reid noted in his debate with the Humeans\(^\text{16}\)—is exactly parallel to
understanding the evidence of the senses. As with linguistic
communication, perception is also subject to conditions under which
the inferences that are normally reliable break down, e.g., in the outré
thought-experiments involving barn facades, painted mules, or
envatted brains that are so beloved by epistemologists in the analytic
philosophical tradition. In the case of perception, however, we have
long ago given up the idea that it is a mark of rationality that one rule
out deviant cases before accepting the belief that immediately
presents itself—namely that all is as our senses tell us that it is. If our
discussion here is correct, then, as with perception, so too with
testimony. The rational method for understanding assertion is the
one licensed by the Reidian disposition-based account: absent
evidence to the contrary, one may “confide in the veracity of others
and … believe what they tell us.”

\(^{16}\) For a detailed examination of the debate between Reidians and
Humeans with respect to the epistemological issues involved in
linguistic communication, see my [1999].
References


