On the Persistence of the ‘Problem of Other Minds’ in Psychology
Chomsky, Grice and Theory of Mind

Ivan Leudar
University of Manchester

Alan Costall
University of Portsmouth

Abstract. In this paper we discuss the historical origins and conceptual debts of the Theory of Mind framework (ToM). We investigate its affinities to Chomsky’s psychology, and Paul Grice’s work on meaning. We find that the ToM framework is resourced by the ideas found in Chomsky and Grice, adding very little new to them, and suffering from the same problems of dualism. ToM inherits the traditional dualistic problem of other minds, tries to solve it, and ends up profoundly intellectualizing social interactions.

Key Words: Chomsky, Grice, problem of other minds, Theory of Mind

The Problem of Other Minds

‘How do we know that another person is angry? . . . Do we ever know?’ (Austin, 1979, p. 76). A sceptical answer is that we can never know, but we may believe it, with varying degrees of certainty and on the basis of inference. This is an answer John Austin found not just wrong, but misconception. Many psychologists, of course, ask more generally ‘how can we know that other people have minds at all?’ and in doing so, in Austin’s words, step even further away ‘from the garden of the world we live in’ (p. 90). It is customary to blame Descartes for setting up this troublesome ‘problem of other minds’, and here is Gilbert Ryle’s distillation of what he termed the ‘Cartesian myth’:

Every human being . . . has both a body and a mind . . .

Human bodies are in space and are subject to the mechanical laws which govern all other bodies in space. Bodily processes and states can be inspected by external observers. So a man’s bodily life is as much a public affair as are the lives of animals and reptiles and even as the careers of trees, crystals and planets.
But minds are not in space, nor are their operations subject to mechanical laws. The workings of one mind are not witnessable by other observers; its career is private. Only I can take direct cognisance of the states and processes of my own mind. A person therefore lives through two collateral histories, one consisting of what happens in and to his body, the other consisting of what happens in and to his mind. The first is public, the second private. (Ryle, 1949/1973, p. 13)

Contemporary psychology claims to set aside this dualism. Yet, even though it explains mind as a mechanism and claims to be materialist rather than dualist, it inherits the problem of other minds in distinguishing mind from behaviour and accepting that mental states are private. In fact, both the distinction and the privacy are formulated very much in the way Ryle found in Descartes’ work. Behaviourists in their determination to be scientific failed to notice that they adopted a definition of behaviour that is basically Cartesian, and were themselves largely responsible for the modern dualistic meaning of ‘behaviour’. The cognitive revolution maintained the problem of other minds by reinstating the legitimacy of scientific inference to the unobservable, thereby retaining this view of behaviour as logically separate from mind, if now accepting it as indirect evidence of an essentially unobservable mind. Behaviourism and cognitivism have both operated in what has remained a fundamentally Cartesian framework for thinking about people. In this respect, the behaviourist and the cognitive revolutions were hardly revolutions (Leahey, 1992; Still & Costall, 1991).

What, then, is to be done about the problem of other minds? One can analyse the problem and its historical roots, and, as Austin did, shrewdly set it aside. The problem may, after all, be a pseudo-problem.1 Or one can try to solve it, and the Theory of Mind is the most recent and influential attempt in psychology to do this. If behaviour is merely an arbitrary ‘clue’ to a hidden mind, and not constitutive of experience, then clearly something is needed to bridge the gap: something like the theoretical leap that scientists make when drawing inferences from observable data to hidden structures has come to seem the obvious way of bridging the gap.2 On this dualist view, not just psychologists, but all of us, engage and must engage in theorizing (or something functionally equivalent to it) to understand other people as intentional beings. Since such understanding is central to our social lives, by implication, it is theorizing about mind that holds social life together. Following a long tradition going back to Plato, the Theory of Mind framework intellectualizes mundane social life—life depends upon theories to be effective and enlightened (cf. Nehamas, 1998; Ryle, 1999).

In practice, of course, we mostly manage pretty well in our dealings with other people without resorting to generalizations and theory-based inferences. Everyday language use is not Cartesian in the manner of scientific psychology but is instead a practical matter—intentionality is ‘noticeable’ without mediating inferences (cf. Merleau-Ponty, 1973, p. 9). Indeed,
Descartes himself insisted that we will never comprehend the mind–body relation just by ‘meditating’ and ‘studying things that exercise the imagination’. Although he certainly regarded mind and body as distinct substances, he took it for granted they had to be thoroughly ‘intermingled’. The mind is not in the body like a pilot in a ship. Rather, the union of mind and body must be lived to be understood: ‘it is just by means of ordinary life and conversation . . . that one learns to conceive the union of soul and body’ (Descartes, 1954, p. 280, emphasis added; see Tibbetts, 1973).

We take it for granted that the problem of other minds does not reflect something timeless and transcultural about the human condition, but is instead a problem which makes sense in specific historical conditions and in its current form is partly of psychologists’ own doing. In psychology, it does not usually arise as a formally stated philosophical problem but is instead inscribed in theories, methods of investigation and in the language of psychology itself—in this respect it is a historical problem and should be treated as such.3

Kurt Danziger in his book Naming the Mind (1997) demonstrated how modern psychology has taken over terms from everyday psychological discourse, such as personality, memory and intelligence, subjecting them to radical transformation, and giving them a unity and abstraction they never previously possessed. The problem of other minds is inscribed in the language of psychology, which sets apart behaviour and mind and then joins them together again along Cartesian lines—with the behaviour providing the data for psychological inference. Austin and Wittgenstein argued cogently that behaviour is not an arbitrary sign of ‘mental states’, but is instead constitutive of intentionality. Our knowledge of others is not usually a matter of inference—one ‘notices and grasps’ directly that they are in pain, for instance, without inference being needed.4 Goffman (1979) pertinently noted that intentionality is ‘given off’ by participants, who make their behaviour ‘monitorable’ by others. This way of thinking is of course not possible if one takes ‘behaviour’ to be just bodily movement and so strips it of intentionality, relocating all that is alive and intelligent in the hidden mind.5 As has been made clear in ethomethodological and conversation analytic studies, this ‘psychologized’ way of talking about pain, beliefs and desires—as referring to hidden ‘mental states’—is inconsistent with how words about intentionality are used in everyday life (see Bilmes, 1992; Coulter, 1989, 1992, for a discussion of this issue). One of us has argued elsewhere that solipsist conception of ‘mental language’ leads to construing incorrectly whole cultures as being literally mindless (Leudar & Thomas, 2000, Ch. 2).

Why is the hidden Cartesianism of psychology not noticed more often, and why is the brain-Cartesianism so attractive to psychologists who deem their work scientific and materialist? In what follows we shall examine some of the historical sources of ToM to understand better how this frankly
dualistic approach has come to seem in psychology so plausible as the solution not just to the traditional ‘problem of other minds’ but also the problem of how in fact we all, as ‘just plain folks’, typically make sense of one another. Our argument is that ToM needs to be understood in the historical context of studies of communication and language in psychology. These, we shall see, are Cartesian in the sense just outlined; ToM follows in their tradition, without resolving the issue, in fact without ever noticing the problem. We find that ToM is above all historically contingent on both the cognitive psychology of Noam Chomsky and the pragmatics of Paul Grice. We argue that both prepared the ground for ToM conceptually even though the influence is not often acknowledged. (Costall and Leudar, this issue [2004], investigate the comparable role of Jean Piaget.)

**ToM and Studies of Communication**

Psychology tends to borrow from other disciplines—formal logic has been used as a framework for studying reasoning, linguistics as a framework for studying language, and formal theories of computation for human cognition in general. The study of communication in psychology obtained a significant impetus after the Second World War, from the developments in communication technology and in information theory (Wiener, 1948). These provided psychology with concepts for research on communication in particular and on cognition in general such as, for instance, ‘information’, ‘code’, ‘signal’, ‘channel’, ‘capacity’ (e.g. Broadbent, 1957; see Vroon, 1987). The mathematical theory of communication distinguishes between ‘signal’ and ‘meaning’, and, as Cherry (1957, p. 9) noted, the theory is concerned exclusively with signals, their transmission and information content, but not with the meaning of signals for the users of communication technologies. One can apply information theory radically, and acknowledge only those mundane meanings expressible in the language of information theory. In adapting the theory for psychology, Cherry, however, retained the distinction between ‘information content’ and ‘meaning’—meanings were psychological states, such as ‘notions’, ‘ideas’ and ‘desires’, which were, in communication, ‘framed into’ speech. By such ‘framing’ Cherry did not mean an *onomatopoeia* (p. 10): in other words, the meaning is not directly given in the ‘signal’. Instead, meanings have to be either decoded⁶ or ‘guessed inductively’. In Cherry’s account, then, meaning was not ‘visible’ but had to be inferred in one way or another—understanding a message consisted in going beyond what is observable.⁷

Cherry’s formulation of the problem of communication reproduces the problem of other minds, as did Reddy’s influential critique of mundane representations of communication (Reddy, 1979). According to Reddy, the commonsense metalanguage of English consists of a set of ‘conduit meta-
phors’, such as ‘putting (thoughts) into words’, ‘getting (one’s feelings) across’. Reddy criticized this metalanguage from a standpoint of ‘radical subjectivism’. According to him, ‘no one receives anyone else’s thoughts directly’, ‘nor can anyone literally “give you an idea”’—these are ‘locked within the skull and life process of each of us’ (pp. 286–287). Language helps ‘one person to construct out of his own stock of mental stuff something like a replica, or copy, of someone else’s thoughts, a replica which can be more or less accurate, depending on many factors’ (p. 287). On Reddy’s account of communication, everything is hidden, nothing is plainly in the open, everything has to be inferred. (But then, we wonder, how do we ever know that there is something to be inferred?) The account of communication and social interaction is telementational—individual social behaviours are caused by their doers’ mental states, and properly understood in terms of those states. Organisms do not react directly to the behaviour of others but instead to the occluded ‘mental states’ which these indicate. On this view, communicative behaviours mediate interactions that are effectively mental. You might say, ‘Well this is surely what communication is actually like?’ But in saying this you reveal how ingrained this way of thinking about communication is and how difficult to step back from.

This mentalized understanding of communication is current in social psychology and ToM inherits it, without taking it much further—besides arguing that one needs a theory to cross the gap between observable behavioural evidence and hidden meaning. As Astington (1994) has put it: ‘Social interaction is really an interaction of minds, of mental states’ (p. 43). The influence is historically mediated through the work of Chomsky and Grice, and modulated by it.

Chomsky is fabled for having delivered modern psychology from behaviourism. His influence on psychology was indeed considerable, even though, from our point of view, somewhat Mephistophelean. In marginalizing behaviour, Chomsky reinforced the Cartesianism implicit in scientific psychology, and even though he did not actually write about it, he established a niche for the Theory of Mind and has indeed been praised for doing so (e.g. Smith, 1999, pp. 25–26). Moreover, as we see it, his programmatic formulation of the pragmatics of language prefigures the theory of mind.

Chomsky and the Status of Theory in ToM

Nobody needs to be reminded of Chomsky’s distrust of the notion of learning (when he writes the word, it is more often than not in inverted commas). Indeed his striking entry into psychology involved showing unequivocally that the behaviourists’ notion of learning was insufficient to account for language acquisition (Chomsky, 1959). Human language could not be acquired by an equivalent of a ‘finite state automaton’ through piecemeal associations, sentence-by-sentence. What were acquired were
rules of language, organized in a recursive grammar, and capable of producing a bounded infinity of sentences. In 1958, in the preface to his Words and Things, Roger Brown wrote that acquiring a language is learning a system, and knowing a language is knowing its rules (p. viii). Brown located the beginnings of psycholinguistics to the early 1950s, singling out an interdisciplinary meeting of psychologists and linguists at Cornell University in 1951. Notably, the participating linguists were ‘descriptive linguists’ not ‘generativists’—that school of linguistics did not yet exist (Brown, 1958, preface). Brown’s comments on the origins of psycholinguistics are cogent. He was perhaps less persuasive, however, when he wrote the following:

Following Chomsky’s talk there was an exchange that went something like this:

Brown: ‘It sounds to me as if a transformational grammar might be what children learn when they learn their first language.’

Chomsky: ‘Oh, do you think so?’

and thereby apparently took credit for turning Chomsky towards psychology (Brown, 1970, p. 17). Chomsky regarded himself as engaging in psychology from the outset (see Chomsky, 1986). His critique of Skinner’s Verbal Behavior (Chomsky, 1959) was addressed to a psychological audience already prepared for the ideas it conveyed—but it amplified and focused them.

Chomsky contributed significantly to the acceptance within psychology of the notion that language is a system of grammatical rules (e.g. Greene, 1972), and to the extension of this way of thinking to other psychological ‘faculties’ (e.g. the ability to think and reason, cf. Johnson-Laird & Wason, 1977). Premack and Chomsky engaged in an ardent debate on whether the acquisition of such language was restricted to humans, or potentially in common with primates. An interesting passage in the seminal article on ToM by Premack and Woodruff (1978), however, reveals an affinity: they accepted that the Theory of Mind consisted of rules (albeit not as tightly structured as a grammar of, say, English), and that these would not be learnable through associative mechanisms. In other words, like generative grammar, ToM is explicitly opposed to behaviourism, and specified as a system of rules internal to the individual. The influence of Chomsky is clear, if unacknowledged. The affinity goes further. Premack and Woodruff define theory of mind thus:

In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and others. . . . A system of inferences of this kind is properly viewed as a theory, first because such states are not directly observable, and second because the system can be used to make predictions, specifically about the behavior of other organisms. (p. 515)
The terms Premack and Woodruff use here to describe how individuals accomplish understanding severely intellectualize the matter. Even so, is a ‘system of inferences’ used to predict behaviour or a ‘rough kind of system with which we all work’ necessarily a theory (cf. Grice, 1989, p. 285)? It is arguable that Chomsky’s work prepared the ground for the ‘theory’ in ‘Theory of Mind’. He wrote that the term ‘theory’ ‘covers general common-sense and belief’ (Chomsky, 1969, p. 53) and this seems to be a relatively fixed element of his psychology. In an apocryphal story in his *Reflections on Language* (1976), he considered that even face recognition might reflect an implicit theory:

S might study, for example, the ability of his subjects to recognize and identify complex physical objects and predict their behavior under various circumstances. He might find that there are qualitative differences in their ability to recognize human faces and other objects of comparable complexity. This investigation might lead S to attribute to his subjects, as an element of common sense, an abstract theory of possible faces and a system of projection which (abstracting away from memory restrictions and the like) enables the subject to predict how a face will appear under a range of empirical conditions. (pp. 139–140)

Like ‘common sense’, linguistic competence might be based on a tacit theory too. A grammar is an aspect of a theory of language in two senses. It is formulated by linguists to describe and explain language behaviour (and this sense we take to be relatively uncontroversial). For Chomsky (1980a), however, grammar is a psychological object:

To know a language, I am assuming, is to be in a certain mental state, which persists as a relatively steady component of transitory mental states. What kind of mental state? I assume further that to be in such a mental state is to have a certain mental structure consisting of a system of rules and principles that generate and relate mental representations of various types. (p. 48)

In this sense, the grammar of a language is ‘known’ by lay individuals. But is it to be thought of as a theory? Chomsky seemed to think so when characterizing language acquisition as hypothesis testing and as did some developmental psycholinguists who at the time wrote of children learning to speak as ‘little linguists’ (cf. Thomas, 2002). Indeed, Jerome Bruner (1978), although highly critical of Chomsky’s claim that knowledge about language learning could be regarded as independent of any wider knowledge of other people or the world, nevertheless had no objection at all to what he took to be Chomsky’s fundamental insight:

I would take the view that the child’s knowledge of pre-linguistic communication, related as it is to world of action and interaction, provides him with tell-tale cues for constructing and testing hypotheses about the meaning and structure of the discourse into which he quickly enters. He does, as LAD [the Language Acquisition Device] would have us believe,
have a stunning capacity to infer and to generate rules, indeed to over-
generalize them. (p. 83)

Nevertheless, the relationship between linguistic competence and linguistic
theory is by no means clarified in Chomsky’s work (cf. Searle, 1972), and in
practice children and linguists do very different things. We suspect that it is
when explaining language in very general and formal terms (as, e.g., ‘a
grammar’, or a ‘projection problem’) that important differences between
what children do as children learning language and what linguists do as
scientists become obscured. Even at the time, the way Chomsky spoke of
grammar as known by lay speakers of language was controversial (see, e.g.,
Stich, 1980), and Chomsky did not of course argue that ordinary folks know
language in exactly the same way as generative linguists. The lay ‘theoretici-
ans’ merely, as Chomsky put it, ‘cognize’ the grammar of a language—
their knowledge is not reflexive but provided as a biological structure (a
language module, perhaps). This knowledge, if it should be regarded as
knowledge at all, is then inevitably tacit and ‘inaccessible to consciousness’.
It is of course debatable whether the unconscious rules of grammatical
competence postulated by Chomsky count as rules in the everyday sense of
the word (see Chomsky, 1969, 1976, pp. 179–195; Hacker, 1988; Quine,
1969; Searle, 1980). Why should we take ‘cognizing’ to be a kind of
knowing? Nevertheless, leaving this debate aside, it is evident that Chom-
sky’s psychology proposes a partial symmetry between expert and lay
‘knowledge’ of language, as there is in the ToM framework (see Costall &
Leudar, 2004). The affinity between Chomsky’s ideas and ToM is again
clear and it is occasionally acknowledged explicitly in ToM writings:
‘Infants have an implicit theory of mind . . . in the same way that two-year-
olds have an implicit theory of grammar’ (Astington, 1994, p. 47).

Chomsky and the Status of Behaviour in ToM

The behaviour of individuals in the Theory of Mind framework has the
status of observable ‘evidence’ for inferring their occluded mental states. A
crucial part of Chomsky’s psychology is what is now referred to as ‘the
poverty of the stimulus’ argument. Usually, and especially in developmental
psychology, this is taken to mean that the ‘language input’ (i.e. linguistic
behaviour of others) does not contain information sufficient to determine a
grammar of the language for a child (cf. Pullum & Scholz, 1992). The
poverty of the stimulus is, however, conceived in somewhat broader terms
by Chomsky, presumably following from his rejection of behaviourism
(Chomsky, 1959). Chomsky did not just argue that children cannot induce a
grammar from ‘language input’ without an evolutionary endowment. He
also disdained the use of a language corpus as an adequate source on which
to base linguistics (e.g. Chomsky, 1964; cf. Sampson, 2002). He argued that
the object of linguistics, language, is an abstraction and hence not a proper
object for a science (e.g. Chomsky, 1976, Ch. 2; 1984, Ch. 2), and declared that he was doing cognitive psychology rather than linguistics, claiming that linguistics was a branch of cognitive psychology (e.g. Chomsky, 1976, p. 36; 1980a, p. 4). With respect to ToM, it is crucial to note that Chomsky viewed behaviour as evidence of knowledge, but not as a criterion of knowledge. The following texts, written by Chomsky in the 1980s, presumably relate to his exchange with Quine, but could have been in response to Wittgenstein’s comments on the same issue (Chomsky, 1969; Quine, 1969; Wittgenstein, 1953):

... one might attempt to characterize the knowledge of language—perhaps knowledge more generally—as a capacity or ability to do something, as a system of dispositions of some kind, in which case it is perhaps not unreasonable to think of behavior as providing a criterion for the possession of knowledge. In contrast, if such knowledge is characterised in terms of mental states and structure, then behavior simply provides evidence for possession of knowledge, as might facts of an entirely different order—electrical activity of the brain, for example. (Chomsky, 1980b, p. 5)

... if such knowledge is characterized in terms of mental states and structure, behavior simply provides evidence for possession of knowledge, as might facts of an entirely different order—electrical activity of the brain, for example. (Chomsky, 1980a, p. 48)

Chomsky thus acknowledges that behaviour may be thought of as either evidence or a criterion of knowledge, this, according to him, depending on one’s ontology. Somewhat paradoxically, however, the ‘fact of the matter’, as far as Chomsky is concerned, is that behaviour merely counts as evidence:

But I see no reason to deny that there is a fact of the matter, however difficult it may be to establish, on that behavior is only one kind of evidence—sometimes not the best, and surely no criterion of knowledge. (Chomsky, 1980b, p. 6, our emphasis)

One can of course retort that meaningful behaviour provides the criterion for telling whether electrical activity in the brain has anything to do with a particular psychological function. Staying with history, however, the poverty of stimulus argument seems contingent on a neo-Cartesian ontology and it sets up behaviour as evidence, and poor evidence at that, for psychological knowledge. Theory of Mind inherits this conception of behaviour and the problems of other minds that it carries.

**Chomsky’s Pragmatics, Grice and ToM**

Chomsky (1966) explicitly declared his linguistics to be in Cartesian mode. He did not, of course, sign up to all aspects of Descartes’ philosophy. He endorsed Descartes’ reflections on human linguistic creativity (pp. 3–6). Where the creative use of language indicated immaterial soul to Descartes,
to Chomsky it implied a recursive grammar which, like the soul, was unique to humans. Since grammar is a psychological faculty realized in the human brain (an aspect of ‘mind/brain’ in his later writings), Chomsky is a materialist, whereas Descartes was an ontological dualist. Chomsky did, however, import into psychology other aspects of Cartesianism, namely the categorial distinction between behaviour and mind. Even though he wrote that he did not believe ‘that bodies are Cartesian automata’ and accepted that the idea of ‘the body’ is evolving (Chomsky, 1980a, p. 6), he also argued, as we have seen, that behaviour is not constitutive of the mind but just poor evidence for mental states.

Do individuals have a privileged access to their own mind in Chomsky’s psychology? Shanker (2002) has pointed out that Chomsky could hardly argue this was so, at least with respect to the organization of the language faculty, since such knowledge is tacit for individuals (‘cognized’ in Chomsky’s terms) but only accessed through scientific investigation of generative grammarians. Nevertheless, we shall see that Chomsky endorsed an essentially ‘telementational’ notion of communication, and hence the notion that one human mind is occluded from another and only revealed, indirectly and ambiguously, in language and other behaviour. As a consequence, he accepted that in everyday social interactions one’s mind is occluded, and so he became committed to the problem of other minds with a vengeance. The important link between Chomsky’s work and Theory of Mind is, then, through his conception of language use: pragmatics. Chomsky conceived of the human cognitive system as consisting of autonomous and biologically fixed modules, or, in another context, ‘competences’ (Chomsky, 1980a, 1980b; see Smith, 1999). The core modules in his psychology were ‘linguistic competence’, ‘conceptual competence’ and ‘pragmatic competence’.

Grice on Meaning and ToM

Grice’s work turned around two ideas—‘non-natural meaning’ and ‘implicature’ (Grice, 1957, 1968, 1969, 1975). The former is the idea that
effects in linguistic communication are accomplished by means of intention recognition, and communication is successful to the extent that communicative intentions are recognized.\textsuperscript{15} The latter idea was that hearers infer implicatures so as to preserve the validity of maxims of conversation.\textsuperscript{16} Initially it was just about possible to understand Grice’s work as an exercise in philosophical logic and mundane reasoning taking place in talk, and not as cognitive psychology. After all, Grice was a philosopher, concerned with defining linguistic meaning, and was doing this partly by comparing different types of meaning. As a philosopher, he would, quite properly, have nothing much to say on how intentions are ‘put into’ utterances or inferred from them. According to this interpretation, Grice should be regarded as writing about the logic of language-games and was certainly not concerned with psychological questions, such as how children move from communicating in pre-linguistic (‘perlocutionary’ in the terminology of developmental pragmatics) mode to linguistic (‘illocutionary’) mode.\textsuperscript{17} Grice’s latter work, however, makes this non-psychological reading impossible to maintain (Grice, 1982, 1989). Grice (1989) writes of communication as the production of psychological states by one person in another. For instance:

\[
\ldots \text{a certain psychological state } \Psi^1 \text{ in certain circumstances is followed by a certain utterance } U, \text{ made in certain circumstances, which in turn, if the circumstances are right, is followed by a particular instance of a further psychological state } \Psi^2, \text{ a state not now in the communicating creature but in the creature who is communicated to. (p. 287)}
\]

The operation of such creatures as I have been talking about is at least in certain circumstances going to be helped and furthered if there is what one might think of as shared experience. In particular, if psychological states which initially attach to one creature can be transmitted or transferred or reproduced in another creature (a process which might be called F-transmission), that would be advantageous. (pp. 286–287)

Grice in fact considered, albeit reluctantly, that this psychological exchange is underwritten by ‘a theory’:

The laws I have mentioned are vulgar laws. The kind of theory in which I think of them appearing would not be a specialist or formalised psychological theory, if indeed there are such things; I am perhaps not very comfortable with the word ‘theory’ being applied to it. It would be the rough kind of system with which we all work, and the laws in it are to be thought of as corrigible, modifiable and \textit{ceteris paribus} in character. (p. 285)

Grice’s ‘non-natural meaning’ became a part of John Searle’s definition of speech acts (Searle, 1969)—their essential condition—and influential especially in cognitive science and in developmental psychology (e.g. McShane, 1980). The reworking of Grice in speech act theory is not quite immaterial—at least not as far as understanding the acceptance of ToM is concerned.
Approaching pragmatics from the direction of speech acts, one is inferring not what a speaker’s communicative intention happens to be, but what speech act was issued. But speech acts are organized in a system and theorized (Searle & Kiefer, 1980). This sort of pragmatics thus implicitly postulates a theory of interpretation—something explicitly declared in ToM. Thus, although Searle took issue with Chomsky’s conception of rules and innateness (see Searle, 1972, 1980) and did not see himself as describing psychological competence, his conception of speech acts did nevertheless provide a link to ToM. In fact, much of the work in developmental pragmatics in the 1970s and 1980s used speech act theory (see, e.g., the work brought together in Ochs & Schiefflin, 1979).

Even in the absence of explicit references to the ToM literature in Grice’s later writings, one is tempted to speculate whether his work did not just prepare the conceptual ground for ToM but was also subsequently affected by it. Some followers of Grice certainly were—Sperber and Wilson (2002), for instance, now refer to the inference of communicative intentions as ‘mindreading’, and argue that an autonomous part of ToM ‘brain module’ is dedicated to doing relevance-theoretic pragmatics. In their work, Gricean pragmatics became assimilated into the Theory of Mind framework, even to the extent of offering speculations about the evolutionary origins of the ‘relevance module’ (cf. Sperber, 2000; Wilson, 2000). Grice’s ‘reluctance’, we noted above, indicates that he was unlikely to have accepted that there is a formal theory with which one can bridge the Cartesian divide between persons. His approach, however, assumes such a divide and, endorsed by Chomsky, Grice provided psychology with a system which mentalizes communication.

Unlike the link to Chomsky, the historical link of ToM to Grice is frequently acknowledged by those working within the ToM framework. Astington (1994), who has adopted a pragmatics based on Gricean principles, wrote:

Social interaction is really an interaction of minds, of mental states, but we have to communicate those states to others. We have to let the other person know we want something, or that we want them to believe something, and so on. Human beings are not mind readers, not in any telepathic sense anyway, and in order to know what is in another person’s mind we have to give that information to one another. . . . You have your thoughts, your beliefs, desires and so on, and I have mine. We share them in language, in the talk that passes between us. (pp. 45–47)

Baron-Cohen (1992) wrote:

Speech Act Theorists such as Grice (1967/1975), Austin (1962), and Searle (1965) had argued that a theory of mind is also essential for normal communication, both verbal and non-verbal. Put simply, the argument is that all communication requires both participants to take into account the background knowledge and presuppositions of the other person in the
dialogue, as well as their intentions in communicating. Such mental state attribution is necessary, it is argued, if a dialogue is to respect the conversational rules of pragmatics—if it is to be appropriate and relevant to the social context (Sperber & Wilson, 1986). (p. 10)

We have seen that Grice was ambivalent about Theory of Mind and, to our knowledge, neither Searle and certainly not Austin ever argued for a Theory of Mind. Baron-Cohen (1988) unfortunately does not provide a clear outline of the relationship between pragmatics and Theory of Mind. He implies that he does not equate pragmatics with Theory of Mind, but regards ToM instead as a precondition of pragmatics. What, then, according to Baron-Cohen, is a Theory of Mind (i.e. in addition to knowing what are beliefs and intentions, etc.)? For him it seems to entail instrumental means–end reasoning about actions in general (cf. Bilmes, 1986). This, however, is to misunderstand what pragmatics is about.

Pragmatics aims to account for the varieties of ways in which people reason in situ about what they themselves and other people do, what they intend and believe, and the circumstances. It usually does this by analysing everyday reasoning practices, such as those Grice described for deriving implicatures. There are of course exceptions in pragmatics—Sperber and Wilson's (1986) relevance theory, for instance, ignores the varieties of mundane relevance, individualizing, mentalizing and technicalizing the notion. Sperber and Wilson (2002) are, however, more clear than Baron-Cohen about the relationship between pragmatics and Theory of Mind. Pragmatics and Theory of Mind are subserved by independent brain modules, with their functional organizations described by relevance theory and by means–end logic, respectively. Consequently ToM is not a precondition for the capacity for doing pragmatics. Sperber and Wilson's position is a notably Chomskyan one, even though in their writing they credit the notion of modularity to Fodor. On the whole, the relationship between ToM and relevance-based pragmatics seems fluid in the ToM framework. Historically, it is clear that the relationship between Gricean pragmatics and the ToM framework is far from unilateral. Grice provided the conceptual grounding for ToM, but his pragmatics were transformed (or at least some varieties of it were) in the process, and subsumed in a speculative modular and evolutionary psychology.

Conclusion

Our historical analysis demonstrates that there are affinities between Chomsky's psychology, Grice's pragmatics and the Theory of Mind framework. Chomsky and Grice passed on the Cartesian problem of other minds to the ToM framework inscribed in their theories of language and meaning. Chomsky further bequeathed it the ideas that behaviour is an impoverished
source of evidence about mind, and, furthermore, that the Cartesian divide is to be spanned, not just in science but also in everyday life, by a theory. This formulation presents mundane understanding as a quasi-scientific endeavour, a form of scientific investigation, albeit one obviously lacking in rigour (cf. Taylor, 1992, Ch. 1).

The common theme is the acceptance of the Cartesian divide, represented in the distinction between ‘mental states’ and ‘behaviour’, with the former being treated as unobservable hypotheses inferred on the basis of behavioural evidence. The Cartesian divide is conceived of as a universal fact about human beings rather than something which is historically contingent and perhaps unnecessary.

These affinities are only occasionally acknowledged by ToM researchers. One objection to our conclusions should be that the similarities which we uncovered reflect the fact that the approaches share the same scientific paradigm. This is indeed so, and the next step in our investigation will be to establish specific episodic and biographic connections between Chomsky and Grice, on the one hand, and those who pioneered Theory of Mind, on the other. For instance, is it a mere coincidence that the seminal article in the history of ToM by Baron-Cohen, Leslie and Frith (1985), on ToM and autism, was produced only a few doors from the University College London Department of Phonetics and Linguistics, which is renowned for its work on Chomsky and Gricean pragmatics (see, e.g., Smith, 1982, 1999; Smith & Wilson, 1979; Sperber & Wilson, 1986)? Did Chomsky, Grice and Searle discuss pragmatics, and, if so, how did they part?

We have tried to avoid critique and focus on the internal logic of the approaches and the conceptual links between them. It must be obvious, though, that we consider the intellectualization of everyday communication inappropriate. Everyday communicative understanding simply does not consist in inferring unobservables using a theory. The issue is how people make sense of what other people do, and this is not just hard but impossible to resolve if one starts by assuming that people’s minds are concealed and only knowable indirectly through ‘behavioural cues’. Such cues are said to be limited and unreliable since there is allegedly no systematic relation between what we observe about other people and their ‘hidden’ thoughts and feelings. A host of experimental studies supposedly ‘confirm’ that people are inconsistent in ‘encoding’ their mental states and inaccurate in making psychological attributions (e.g. Fiske & Taylor, 1991). Conversation analytic studies, however, show that conversations are intricately organized so that individuals’ contributions to social interactions incarnate intentionality, more often than not in an unambiguous way (e.g. Drew, 1995; Mandelbaum & Pomerantz, 1991; Pomerantz, 1990). Moreover, if there really were no systematic relation between how people act and what they feel and think, it is not easy to see how anyone, even a psychologist, could determine the accuracy or inaccuracy of psychological attributions. An explicit connection...
should be made between how real people make sense of other people in everyday interactions, and the problem of how psychologists go about doing their kind of science. The interactions between investigators and participants in psychological experiments are rarely analysed as dialogical engagements, but there are exceptions which show incongruity between the precepts and practices of experimental research. Participants’ behaviours are demonstrably interactively contingent on those of experimenters (cf. Leudar & Antaki, 1996). We note a curious inversion: psychologists’ accounts of mundane communication are modelled on practices of experimental psychological investigations rather than experiments analysed dialogically. Yet studies of the dialogical engagements of experimenter and participants would inevitably throw many of these experiments in doubt, as incoherent and as investigating possibilities that the experiments themselves presuppose. What psychology needs is a good account of communication which is applicable to its own investigative practices—what it gets is one designed to resolve the spurious problem of other minds.19

The Theory of Mind framework is caught in a quandary. On the one hand it draws a sharp distinction between the mind and behaviour, so that the former cannot be reduced to the latter. On the other hand, however, behaviours are said to be the evidence from which mental states are inferred. The problem then is, if ‘behaviours’ are a distinct order of phenomena altogether, how can they signify ‘mental states’? How can the observed behaviours possibly suggest that an inner mental state is a possible inference to make, and how can they justify ‘mental states’? A mentalist could riposte that the problem is a general one (and so nothing specific to the Theory of Mind framework)—how can anything be evidence for something else? Peirce was clearly aware of this problem in arguing that the unity of a sign was provided by an ‘interpretant’. Plainly, the connection between evidence and the conclusion it supports needs to be understandable and ‘warrantable’. But then one can also ask (and sometimes we do ask), ‘How do I know that the warrant justifies the connection between the evidence and the conclusion?’ To avoid a potentially infinite regress, Stephen Toulmin (1958) argued some time ago that valid arguments are grounded. The problem for ToM behaviour–mental state inferences, with the two terms placed in separate and autonomous domains, is that such ‘grounding’ is unattainable. Why should this be? What ‘grounding’ is may vary, but one kind is practical action. Looking at a computer monitor, for instance, I know that it has a far side which cannot be seen just now. The grounding connecting our experience now and the knowledge of the occluded part of the monitor is practical action—we could always turn the monitor round or walk round the desk. But what could be the grounding for the behaviour to mental state inferences? We certainly can’t stroll from someone’s behaviour to someone’s mind (and, as Sharrock and Coulter [2004] argue cogently in this issue, the theory to connect behaviours to mind epistemically is not learnable).
Speaking about the relationship between ‘language’ and ‘reality’, and dealing with a like problem, Hilary Putnam (1995) writes: ‘The real worry is that sentences cannot be true or false of an external reality if there are no justificatory connections between things we say in language and any aspect of that reality whatsoever’ (p. 65). Putnam’s solution is to argue that language and reality are not autonomous, but ‘interpenetrate’ each other. What would this solution require from the Theory of Mind framework? Basically ToM would have to abandon epistemic dualism and accept that intentionality and behaviour are not different in kind but aspects of the same phenomenon—activity. But then there would be no need for a Theory of Mind.

Notes
1. See Bennett and Hacker (2003, pp. 316–322) for an introductory précis of philosophical arguments on why the problem of other minds is misconceived.
2. Ter Hark (2001) has noted that the theoretical solution of scepticism consists in assimilating ordinary knowledge claims about other minds to procedures of scientific inference.
   By transforming ordinary use into a form of hypothetico-deductive entailment, Fodor and Churchland have made propositions about other minds and propositions about mathematics and physics into similar instruments—the way out of the sceptical problem is to recognize that we measure with different instruments. (p. 215)
3. If this intuition is correct, then the problem of other minds will be resolved not just through critique, or by thinking up a new non-dualist ontology, but by a change in the language and practices of psychology.
4. We may of course occasionally not know what others are feeling, thinking or wanting, especially if they work on concealment, and this problem clearly crosses historical and cultural boundaries. That one has to work on concealing psychological states, however, implies that usually our experiences are ‘noticeable and graspable’ by others directly, not that there is a profound ‘problem of other minds’.
5. The point stands, whether that mind is conceived as occluded in principle or just physically (coextensive with the brain and ‘in the head’).
6. The code was defined as ‘an agreed transformation, usually one to one, and reversible, by which one set of messages may be converted from one set of signs to another’ (Cherry, 1957, p. 7). Cherry also argued, quite sensibly, that human languages are not precise like artificial languages and so understanding cannot consist in mechanically decoding the signals.
7. It is important to mind one’s language here. One line of argument is that Cartesian dualism is not (just) a philosophical thesis, but is built into the language of psychology and into its practices of investigation. Were we ourselves to accept the terms ‘mental states’ and ‘behaviour’ in their usual psychological senses, and declare that the former is visible in the latter, we would maintain and propagate the dualism. So we avoid these terms and say that thoughts, intentions, fears, pains, jealousies, etc. (i.e. experiences, intentionality
in general) are ‘noticeable and graspable’, and this from how an action is done
in situ. (The terms are borrowed from Merleau-Ponty, who used them to avoid
ty ing understanding of others to the five senses.) Our empirical research is partly
concerned with how interactions are organized to make experiences ‘noticeable
and graspable’ (or occluded!).

8. The nativist corollary, that grammar ‘grows’ rather than is learned, was not,
however, as broadly accepted at the time. Roger Brown, for instance argued
consistently that grammar of language is induced from a corpus (e.g. Brown,
1958, p. viii; see also Cromer, 1980; cf. Shanker, 2002).

9. Not all differences, though, always disappear. Stanley Peters (1972), for
instance, made clear at the time that he considered it important that the linguistic
corpus is unordered, but children experience language in an ordered manner.

10. This is a notable difference between Chomsky’s and Descartes’ conceptions
of knowledge. Chomsky allows ‘cognizing’ (i.e. unconscious knowledge),
whereas, according to him, it is crucial to Descartes that knowledge is accessible
(Chomsky, 1976, p. 23)

11. Astington was characterizing the response of the ‘theory theorists’ to evidence
of early understanding of other minds.

12. Not many psychologists, however, accepted that Chomsky was doing psychol-
ogy. Psycholinguists did accept Chomsky’s grammar as a linguistic description
of language but insisted that it be tested experimentally for ‘psychological
reality’. The criterion of the model being ‘psychological’, then, seems to have
involved the method of investigation. Doing experiments was the right method;
(1970) thought psycholinguistics should be an interdisciplinary science requir-
ing psychologists to become sophisticated linguists.

13. Chomsky would presumably endorse the modular version of the Theory of Mind
but not the ‘theory-theory’ version.

14. Grice (1982), however, did not endorse transformational grammar.

15. Grice’s definition of ‘non-natural meaning’ was, roughly: Speaker S means
something x by an utterance U if he utters it with an intention I, wants this
intention to be recognized, and the recognition to be instrumental in achieving
the intended communicative effect. Grice (1957) did not have much to say about
natural meaning, except by exclusion.

16. The maxim of quality is in fact a precondition for the operation of the
intentional mediation model of communication (Grice, 1989; cf. Leudar &
Browning, 1988).

17. In developmental pragmatics the word ‘perlocutionary’ tends to mean ‘in-
strumental’ rather than what Austin meant by the word—an actual consequence
of a speech act. The idea is that a child somehow moves from instrumental to
conventional use of language, and in this respect developmental pragmatics
antecedes ToM—conventional use of language involves attributions of in-
tentionality.

18. Our suspicion is that the ToM framework owes much to the University College
London connection.

19. One can say that in neo-Cartesian accounts of social interaction (such as the
theory of ToM) ‘one does not encounter others’ (Merleau-Ponty, 1973, p. 3.)
References


**Ivan Leudar** was born in Czechoslovakia but moved to the United Kingdom and studied psychology at Birkbeck College and University College London. He is now Reader in Psychology at the University of Manchester. His work in psychology is influenced by ethnomethodology and pragmatics. One of his main interests is how membership categories for people on the margins are produced, maintained and contested in talk and mass media. His book with T. Thomas, *Voices of Reason, Voices of Insanity*, was published in 2002 by Routledge. **Address**: Department of Psychology, University of Manchester, Oxford Road, Manchester M13 9PL, UK. [email: I.leudar@manchester.ac.uk]

**Alan Costall** is Professor of Theoretical Psychology at the University of Portsmouth. His empirical, theoretical and historical work explores the implications of a (broadly) ecological approach to the human sciences. He recently translated Georges-Henri Luquet’s *Children’s Drawings* (Free Association Press, 2001). **Address**: Department of Psychology, University of Portsmouth, PO1 2DY, UK. [email: alan.costall@port.ac.uk]