Imitation or the Internalization of Norms: Is 20th Century Social Theory Based on the Wrong Choice?

The dispute between simulationists and theory-theorists follows a basic pattern in philosophical discussions of cognitive science: various theoretical principles pertinent to mind, that commend themselves to us for some reason, conflict and must be reconciled, both with one another and with some partially understood or stable domains of fact (such as the empirical phenomenon of language learning), that are taken to be especially significant or problematic. For the most part this has been psychological and linguistic material. The theory-theory/simulation dispute, as well as that over connectionism and rule/computational models of the mind, have handled conflicts over such “empirical” linguistic and psychological findings in the following way. One side shows that their account is capable of “explaining,” or at least of being consistent with or of excluding the opposite of some experimental or linguistically observed finding, and argues that the alternative explanation either could not account for or was inconsistent with the finding. Sometimes the argument takes the more limited form of showing that a given approach would face particular novel explanatory burdens in attempting to explain something; the response is often to show how these burdens might be lightened.

To be sure, there are findings, such as the psychological experimental findings that make up the core of the discussion, that need to be accommodated by a successful theory. What this means in practice is that the focus of debate is on, so to speak, each side's efforts to explain the same things. This tends to narrow the discussion. This chapter, like some other contributions to the debate, will attempt to broaden the range of relevant phenomenon, and in this way to alter the balance of explanatory burdens. There is nothing sacred about the list of domains taken to be relevant. In this chapter I will extend the discussion to some different incompatibilities between principles and some different domains of “stable, partially understood facts.”

My perspective is social theoretical, and my aim in this chapter is to bring some of the topics of social theory into the discussion. I proceed from a specific thought that not all social theorists share: that certain standard accounts of culture, namely those which treat culture as a shared tacit theory, are problematic pseudo-explanations. The theory-theory has a peculiar

* This chapter was completed while the author was Fellow at the Swedish Collegium for Advanced Study in the Social Sciences, whose support is gratefully acknowledged.

relationship to these standard views of culture, which I will explain in the chapter. It appears to be compatible, but on examination, and in relation to some more or less stable facts, it turns out that it is not. The incompatibility is so difficult to resolve that it suggests that an alternative approach to mind, such as simulationism or some variant of it, is more plausible, and also that the standard account of culture is not viable. These issues become particularly apparent when we examine a set of stable, partially understood facts about the differences between psychological language in different cultures.

Enculturation

Appeals to the notion of enculturation frequently arise in the theory of mind literature, though they are rarely pursued. One very interesting suggestion, made by Peter K. Smith, indicates the reason these issues might be much more relevant than hitherto imagined. Smith notes that the best evidence for mind-reading by chimpanzees “comes from those exposed to a symbolic language code,” and observes that “if the labeling of IVs (i.e. the intervening variables in behavior that correspond to what were traditionally called “mental states”) is a crucial step, this might suggest that the presence of words (or symbolic codes) for mental states, such as desires and beliefs, in the language of parents and care-givers may be crucial.” He goes on to argue that “the availability of such codes is crucial for the developing individual to develop mind-reading abilities. . . Various aspects of the care-giving environment will facilitate their use.” Having a theory of mind, in this view, is dependent on being part of a “theory of mind”-possessing community.

This formulation quite directly poses several problems that are common to social theory and the theory-theory/simulation debate. What is “enculturation,” and what is “culture?” Is a theory of mind a cultural artifact, and is acquiring one thus socialization into a community and its culture? Suppose that it is. Smith himself thinks “that the explicit use of such codes by care-givers is likely to be important.” But the specific language of mental states—belief, intention, desire, feeling, and so forth—varies quite extraordinarily between cultural communities. And the language of “facilitate,” “crucial steps,” and so forth is open to interpretation: is the culture merely an aid to the more rapid acquisition of a capacity that would in the course of time be reached independently? Are the various languages for mental states just different expressions of the universal human theory of mind that would be arrived at, or is in some sense already possessed by the normal human? Do different cultures possess different theories of mind, or merely different modes of expression of a common underlying tacit “theory”? Are the variant “codes” constitutive or descriptive, and if descriptive, what are they descriptions of? If they are descriptions, could one culture’s theory, or mode of expressing a universal underlying theory, be said to be better than another’s? These are not idle questions, for there is a serious problem about the universality of the folk psychology and theory of mind that the theory-theory is concerned with: the terms of the folk psychology do not have translational equivalents in all the languages and cultures of the world, an issue whose significance and relevance will become evident in later sections of this chapter.

---

Understanding and Norms: the Conventional View

Questions about the nature of understanding are not entirely the product of the debate over theories of mind: social theorists and philosophical writers on interpretation have long written on these topics. The discussion of the problem of understanding in social theory has developed in two traditions: Verstehen or empathy, the German tradition of Dilthey and Weber, and in taking the role (or attitude) of the other originating in the thought of G. H. Mead. Each regards understanding as both an activity of theorists (or historians and other analysts) and of human beings themselves in the course of their dealings with one another. The problem of culture or norms takes two basic forms as well. The dominant one, rooted in Durkheim and Parsons, treats norms, or culture, as a shared body of rule-like cognitive objects which are understood to be tacit.\(^3\)

The connection between the problem of norms and the problem of understanding, though close, is rarely discussed. Stated as a thesis, the conventional view is this: understanding, properly speaking, presupposes the sharing of norms. Usually this is stated even more strongly as the claim that communication (or some other feature of human interaction) would be impossible without the sharing of norms. There is also a close and rarely discussed affinity between the problem of norms and cognitive science models of mind that treat the mind as a rule governed machine. To the extent that norms are thought of as akin to linguistic rules, and rule governed models of mind as the best approach to explaining the way in which individuals are able to master and follow rules of grammar, tacit social norms may be thought of in the same way. The connection is implicit in the commonplace analogy between tacit societal norms and tacit linguistic norms.

Some social theorists have been skeptical of the “culture as rules” model, and a few have provided alternatives to it. One skeptic was Durkheim’s main French rival, Gabriel Tarde, who proposed a theory of “imitation” to explain most of the same phenomena. The basic idea behind this alternative is present in Nietzsche as well. Children, Nietzsche observed, were natural apes, and primitive human beings were, he argued, the same. A herd instinct, rooted in, or taking the form of, a horror of difference, was the psychological basis of morality, whose specific contents were acquired by childish imitation. Tarde’s notion of imitation was more complex, and included various kinds of imitation, such as “rational imitation.” Weber, who was unenthusiastic about the norms model, accepted “imitation” as a genuine phenomenon and also suggested that morality was, if not rooted in, buttressed by, a biological urge to conform— a model that was not so different from Nietzsche’s.

These alternative approaches are very interesting in connection with the questions raised above about the nature of a “theory of mind.” They deal with the phenomenon of social norms

---

\(^3\) The model figures prominently in the philosophical literature as well: perhaps its simplest statement is to be found in John Searle’s *Speech Acts*, in which it is claimed of language “that the speakers of a language are engaging in a rule governed form of intentional behavior” and the "rules account for the regularities in exactly the same way as the rules of football account for the regularities in a game of football.” Searle 1969, 53. Cf. Stephen Turner “Searle’s Social Reality,” in press.
without appealing to any notion of an underlying set of rules or rule like shared mental contents, such as a shared tacit theory. Imitation is wholly external: One can imitate only what one can see or hear; that is to say the externals of an act, thus the content of imitation, is limited by our ability to identify something to copy. We may imitate unconsciously, but this does not mean that we have special powers of unconscious discernment that allow us to discern anything other than the external aspects of what we imitate.

Fundamental Psychological Mechanisms: The Alternative Approach

“Imitation” is a basic psychological mechanism, so approaches to culture based on imitation proceed by first taking a backward step to the following question: what are the basic psychological mechanisms for such things as understanding, that is to ask what is the basic mechanism with which children build an understanding of minds, and what internal givens and external data do they build with? Mead gave a distinctive answer to this “basic mechanisms” question that conflicts with the theory-theory.⁴

Mead’s approach is similar enough to simulation and an older “imitation” theory that Mead began his work by criticizing the idea that the three might be assimilated to one another. I will ignore the differences between the three approaches in what follows, though obviously they were crucial to Mead, and call this family of theories the “emulationist” approach. The common idea of the three theories is that social interaction and “understanding” is the result of the workings of basic psychological mechanisms which enable an individual to emulate another person in thought, for example to guess what they are going to do, and to refine their expectations and responses to others on the basis of the external data given by the responses and conduct of others.

Mead’s variant on this basic idea arose out of a critique of the notion of imitation with which he attempted to account for the problems with the notion of imitation and, more interesting for the purposes of this chapter, to replace it as the fundamental mechanism needed for an explanation of certain key facts about individual social psychological development.⁵ The topic of imitation became important in American psychology for reasons quite separate from its role in social theory. James Mark Baldwin wrote an influential work based on observations of his two young daughters.⁶ He observed that young children imitate, something that was not controversial then or now: we may take it as a “stable domain of fact” that young children imitate. We may also stipulate the following: Children role-play, and do so endlessly and without apparent motives. They also do various other things, such as repeat, endlessly ask “why” questions that adults find non-sensical, repeat the answers to them, initially without

---

⁴ A. F. Goldman has already noted that Mead is a precursor to simulation, in “Empathy, Mind, and Morals.” Goldman 1995, 196.

⁵ In what follows, I will rely on two useful guides to these issues: Gary A. Cook. 1993. George Herbert Mead: the Making of a Social Pragmatist; Hubert Kessler. 1941. Basic Factors in the Growth of Mind and Self: Analysis and Reconstruction of G. H. Mead’s Theory. Dissertation, University of Illinois. This dissertation was directed by Mead’s student and follower A. E. Murphy.

⁶ Baldwin, 1895.
“understanding” them, and apply and alter the answers in new settings, and so forth. Moreover, children vary a great deal in what they do.

The issues between Mead and Baldwin were conceptual, not empirical. Mead argued that one problem with the notion of imitation as a fundamental mechanism is that it presupposes what it needed to explain, namely conscious agency, since imitating is already an intentional act. This of course is an arguable point. Psychological research on imitation establishes that unconscious imitation does in fact occur. But the issue points to a more fundamental problem for theories of mind about the causal or genetic priority of particular psychological mechanisms or phenomena. If we attribute a given capacity to the mind prior to its development through external inputs and attempt to explain other capacities as derivative (e.g. by showing how they are produced by the initial capacity operating on inputs, such as experiences, of various particular kinds), we must be careful not to sneak into our account of the fundamental mechanism features of the phenomena to be explained. In the case of consciousness, or alternatively of intentionality, the temptation is always present, though usually in a concealed form.

Typically this problem arises through, so to speak, reading motives appropriate to developed minds into undeveloped minds. For Mead, imitation was an paradigm case of this problem. People, like the proverbial monkeys, do “copy” the external behavior of others. But the language of copying and imitating is already intentional. We would not say that the accidental repetition of one action by another person or monkey under similar conditions and in close contact was imitation, but we would want something more, some additional mental element to distinguish the accidental from the genuine case. When we begin to describe this additional element, for example in terms of awareness of what the other is doing, we already have imported a minimal notion of intentionality, namely the capacity of the imitator to recognize something as a doing or an action. If we add to this something like self-awareness of copying we have imported a notion of conscious behavior as well. We no longer have a fundamental mechanism, but a mechanism that presupposes a developed mind.

Simulation, imitation, and taking the attitude of the other are similar in a key respect. They are presented as fundamental mechanisms, that is to say, as mechanisms that can explain various mental capacities, but which do not presuppose or require a great deal of prior mental equipment. Simulation differs from imitation in the directionality of the relationship. Imitation, as I have noted, operates by an external pattern being reproduced by an individual. In simulation, the individual has a capacity to think hypothetically or off-line about the individual's own behavior or experiences and to project the results of this thinking to the outside world, for example in the form of expectations about the behavior and responses of others, and to get feedback about the success of these projections. To understand the anger of another person, for example, it is not necessary for the person doing the understanding to have a theory, but only for them to be able to place themselves imaginatively in the other person's shoes and to monitor their own reactions to the imagined situation. The capacity enables learning, in that one can act on one’s understanding, and “test” it by acting on it, and modify one’s simulative capacities accordingly.

In each of these accounts, the fundamental mechanisms, plus learning by experience through the employment of the fundamental mechanisms, do the explanatory work. Baldwin argued that infants were endowed with a capacity and drive to imitate, and that they developed mentally through attempts at imitation. Mead, as I have noted, objected to this on grounds that at
the time must have seemed to persuasive, namely that imitation presupposed what it aimed to explain, namely capacities for recognizing people and consciously copying. The issue is important for our purposes because what Mead is talking about here is precisely the issue raised by simulation theory, namely whether interacting with other people requires those who interact to have a so-called theory of mind already or whether one’s conceptions of other people and capacities for reasoning about their mental processes derived from some prior act, such as simulating the attitudes of others from which one generates, through a process like hypothesis testing, a set of expectations and capacities that are the functional equivalent of a theory of mind.

Explaining Our Capacity to Understand

One reason the question of what is prior is especially important relates to understanding: If the way in which individuals acquire the capacity for understanding others resembles the acquisition of a theory, such as a theory of mind, and if, in the cultural case, it requires the acquisition of some set of normative rules that are interiorized or internalized, the things acquired are assumed to be essentially the same for everyone, and the condition for interaction and mutual understanding is the fact that each of us share the same theory. Empathy, understanding within a culture, normative action, and so forth are, on this view, like the situation in which two scientific observers in possession of the same instruments and theory are able to make the same calculations and communicate it to one another precisely because they have the same theory. But this describes the case in which people already possess the same theory. The developmental question is “how do they acquire it (or its functional equivalent)?”

One problem with the “shared theory” account, from a developmental point of view, is that individuals’ experiences in the course of growing up are extremely diverse. In the case of language, the massive quantity of data overwhelms and obliterates individual differences, at least with respect to such things as basic grammar. But for the sorts of things that people have in mind when they talk about empathy, individual differences are bound to be of greater significance. G.H. Mead’s account of development holds that children take the attitude of others, and grants that children imitate. What distinguishes children from one another is content: the roles that they take and the things that they imitate. And here experiences are in some respects common and in other respects very diverse. Little boys in the United States a half-century ago took the role of cowboys, and little girls took the role of mommies and role-played with dolls, though obviously not all little girls and little boys did these things, and some did both. But in the course of their

7 The conceptual point is important, but in this case the empirical evidence supports the idea that some capacity for imitation is very basic. As one psychologist puts it, “by the second day of life, babies can reliably recognize their mothers; and also imitate facial gestures such as mouth opening and tongue poking.” Ellis 1998, 23 This is imitation preceding any sort of feedback which would enable the infant to learn how to “copy,” and thus before having an “intention to copy.” For this reason I would be skeptical about any attempt to salvage a pure form of Mead’s account.

8 A simple account of the way such a process would work is available in Elgin 1996, 205-211, esp. 210).
development children take quite different roles primarily because they are in different social environments and are compelled to master interaction through this role-playing with people who are themselves taking different kinds of roles. Roles are themselves stereotypical, a point to which I will return shortly.

Role-taking understood as a mode of acquisition of competencies (as well as imitation understood in the same way) falls into a category that for convenience I will label diversifying: the same mechanisms, applied in the actual situations in which individuals find themselves, lead to diverse capacities of understanding. The differences bear very directly on the problem of intercultural comparison and understanding, and provide a simple explanation of the main problems in this domain. The reason we don’t have empathy for, say, historical figures in ancient Rome is that we simply do not understand the roles that they are enacting because we never enacted these roles ourselves. Thus we are inclined to reinterpret empathically the activities of Caesar or Cleopatra in the stereotypic forms of the roles that we have in excess, so that Cleopatra presented on the big screen consists of historical figures behaving in stereotypic ways associated with lovers and statesmen of our own era.

These movies are, so to speak, hypotheses about these historical figures in which we, living in the present, make the historical record conform to expectations about attitudes and the like which we have acquired in the course of our own role taking and role playing activities. Not surprisingly, it is usually only possible to make these stories psychologically compelling if we fill in the historical record in a way that better fits our expectations-- to make Cleopatra behave in familiar ways, with familiar emotions and attitudes. But in real history and archeology matters are much more difficult: the conduct of the ancient Romans, especially with respect to sexuality, is a constant challenge to our powers of empathy. It is simply difficult for us to make heads or tails out of their attitudes and motivations because we do not share them, and we do not share them, not because we are some sort of fundamentally different human type, but because they acquired their capacities from a pool of stereotypic courses of action that we no longer possess, and which are essentially irrecoverable for us because the social world is in which they acquired them has disappeared.

The Meadian view of understanding is that understanding another person consists in taking the attitude or the role of the other. Indeed social interaction and its continued enactment and operation depended very heavily, in the view of Mead’s followers, on the fact that people could rely on various stereotypic expectations. If I go into a doctor’s office I know how to behave as a patient and the physician knows how to behave as a doctor. By virtue of these roles we understand one another by imaginatively projecting ourselves into the role of the other. If the doctor asks me a personal question, I respond to it by understanding that it reflects and is meant in terms of the role of the physician and not, say the role of a new acquaintance or a clerk.

In what I have said here I have vacillated between the “stereotypic” aspects of this process and its “diversifying” ones. The literature itself tends to vacillate in this way, and in one sense the notion of role is extremely congenial to the notion of a culture as a set of normatively defined roles, and it is in this form that it entered the thought of Parsons. Roles are simply understandable as enactments of normative expectations and norms can be understood as a notion correlative to roles. If you are sick, for example, one must conform to the sick role in order to be “understood.” Parsons, indeed, studied precisely these expectations. But everything depends on how one thinks of the stereotypes. One might think of them as common possessions,
as a kind of theory that each member of a group acquires. Or one might think of them as internal mental devices, as a private shorthand (consisting, perhaps, of expectations) which is partly structured by public codes, that is to say by language, or what Mead called significant symbols. Weber seems to have thought of what he called ideal-types in this way as well: as a private set of simplifying cognitive devices, whose employment, consciously or unconsciously, is demanded by the economy of thought in the face of the infinity of perceptual material. But he also recognized that one could be trained, as a lawyer is trained, in the use of conventional ideal-types, such as those of legal categorization, and that one communicates scholarly findings by constructing such types for others.

This “private shorthand” model preserves the personal or diversifying character of stereotypes. We may each use the same terms in describing a “sick” person, and the practical necessities of communication provide feedback, and a kind of public disciplining of our usage occurs through the various encounters in which we use the term. But the mental content of the term for us is our own, the product of our experiences, and it could hardly be otherwise: language learning is done through experience and feedback like any other learning.

Two points should be made here, to forestall any misunderstanding and to prepare the ground for the final discussion. Nothing in what has been said here to characterize the “emulationist” approach conflicts with the thought that, among the basic psychological possessions or starting points of individuals, along with a basic mechanism such as simulation or imitation, are some basic emotions, perhaps associated with basic facial expressions. Indeed, it helps this approach if the basic material with which “understanding” develops includes some common starting points, so that children do not have to learn what pain is and what expresses it, but can instead use this capacity in interpreting others. Common starting points are sometimes called basic emotions, and there is important evidence that there are some universal feelings with universally correlated facial expressions.

If there are some such common (and non-cultural or universal) starting points, it makes it possible for us to create a bridge into another language or into the emotional experiences of others, by linking our “hypotheses” about the feelings of others with a set of facts, namely the facts conveyed by facial expressions. There might also be “non-basic” emotions, that are not indicated by facial expressions. These might well arise through various transformations and training out of, or building on, basic emotions. Patriotic feeling, for example, is perhaps rooted in fellow-feeling and pride, but is the product of training or the experience of participating in

9 Weber is quite clear that these concepts cannot represent the shared theory of people: the “ideas which govern the behavior of a population of a certain epoch...empirically...exists in the minds of an indefinite and constantly changing mass of individuals and assumes in their minds the most multifarious nuances of form and content, clarity and meaning.” If, for example, the “Christianity” of the individuals of the Middle Ages “could be completely portrayed,” he adds, it would be a “chaos of infinitely differentiated and highly contradictory complexes of ideas and feelings.” Weber 1949, 95-96.

10 Ekman 1993, 384-92. The idea is also found in Darwin, who attributes them to animals, (cf. The Expression of Emotions in Man and Animals, 1988), and has a close affinity to Wittgenstein, who says “the common behavior of mankind is the system of reference by means of which we interpret an unknown language.” Wittgenstein 1958[1953] §206.
specific rituals, and is non-universal and non-basic. There is no guarantee that such emotions will be accessible to those who have not shared the experiences, or analogous ones. Feelings of sexual attraction, though certainly rooted in something “basic,” may also develop in very different and mutually incomprehensible ways, depending on the experiences of the individual.

With these various distinctions in mind, we can consider the question of what is it that people “acquire” through feedback from their experiences with others. One answer would be what might be called the “culturalist” one— that people acquire a culture, in a specific sense of this term. Clifford Geertz, himself a student of Parsons, famously defined culture as

an historically transmitted pattern of meanings embodied in symbols, a system of inherited conceptions expressed in symbolic forms by means of which men communicate, perpetuate, and develop their knowledge about and attitudes toward life.\(^{11}\)

We will shortly reconsider this definition with respect to the notion of “inheritance.” The feature that is crucial here is the notion of “system of . . . conceptions,” which is to say a theory-like intellectual object. In an earlier part of this discussion I asked the question whether the theory of mind was such a system of inherited conceptions, and delayed answering it.

Another answer to the question “what do people acquire” would be this: all that understanding another person means to “understand” is to be able to interact, and this means to play the relevant roles, to take the attitude of the other in the relevant stereotyped ways, and to employ the significant symbols in this activity of attitude-taking. What we “acquire” is no more than the results of our own attempts to interact on the basis of our hypotheses about the attitudes of others, and the feedback that enables us to improve our attempts to take the attitudes of others is the success and failure of our interactions. There is no more than this to understanding, and specifically nothing in the way of “a system of conceptions” that must be “inherited” in order for us to understand.

This “emulationist” account of mutuality is also an account of culture, or rather an alternative approach to the phenomena that are usually explained as “cultural.” Acquiring a culture, in this approach, is nothing more than acquiring the ability to take the attitude of others with respect to stereotyped roles and significant symbols. “Culture is thus not a system that needs to be “inherited,” but a set of learned capacities organized around objects, such as symbols, ritualized actions, and so forth. The capacities are individualized in the sense that personal history of learning is not irrelevant, but disciplining, mutual interaction, and experiences around particular symbols may tend to make people respond in ways that are outwardly uniform.

Culture vs. Psychological Universals

The linguistic psychologist Anna Wierzbicka, who works on the problem of the language of emotions, imagines the following dialog between psychologists and linguists:

P: Fear and anger are universal human emotions.

\(^{11}\) Geertz 1973, 89.
L: Fear and anger are English words, which don’t have equivalents in all other languages. Why should these English words, rather than some words from language X, for which English has no equivalents—capture correctly some emotional universals?

P: It doesn’t matter whether other languages have words for fear and anger or not. Let’s not deify words.

L: Yes, but in talking about these emotions you are using culture-specific English words, and thus you are introducing an Anglo perspective on emotions into your discussion.

P: I don’t think so. I am sure that people in these other cultures also experience fear and anger, even if they don’t have words for them.

L: Maybe they do experience fear and anger, but their categorization of emotions is different from that reflected in a standard English lexicon.

P: Let’s not exaggerate the importance of language.

Q: This conflict is important for the theory-theory, but for a reason that is difficult to explain. The conflict between the linguist and the psychologist is closely analogous to a conflict between a “culturalist” interpretation of the “theory” involved in the kind of reasoning about the minds of others that is learned by children, and a “psychological” one.

How does one explain the diversity of emotion terms in different languages? Wierzbizka considers Angst, an emotion that has no proper English equivalents. The problem with Angst is this. It is widely reported as an emotion among Germans, but not reported widely in other cultures, and indeed is not part of the language at all in most other cultures. This emotion, then, is “cultural.” The psychological theory is this: there are certain common basic emotions that people have. Angst is therefore not “basic.” So there is no direct conflict here. But there is a conflict nevertheless, and it is a profound one.

Both theories attempt to explain the same kinds of facts, emotions. The linguistic or culturalist explanation suggests that emotions are in some sense cultural, dependent on the things that language is dependent on. The psychological explanation holds that at least some emotions, and perhaps all real emotions, are universal, and explicable by features of human psychology, biology, and development that are themselves universal. Neither can be completely right. Either the “culturalist” explanation is specious for some emotions, such as anger, or the psychological account is false, a mistaken attempt to turn a cultural fact into a universal feature of psychology. There is a burden on the psychological account to explain such emotions as Angst, that apparently are not universal. And there is a burden on the culturalist account to explain why some emotions appear to be cultural and others appear to be universal or non-cultural, even though there is no apparent reason that the same cultural explanations that are invoked for Angst could not also be invoked for anger.

Of course, there are a variety of possible solutions here. But the important point about these “solutions” is this: they impose additional burdens on the explanation, burdens that are not apparent from the initial explanation. The same point holds in the case of the theory-theory, in its relationship to culturalist explanations of “folk psychology,” where there is a similar problem about language, a problem which is in some respects more serious. Folk psychological language is theoretical, and for this reason even more varied than emotion language. Four examples can be

given. The term “believe,” in which the “false belief” problem is itself defined, lacks good equivalents in some languages— a problem discovered at a practical level, as Rodney Needham has observed, by Bible translators. In many languages, it is impossible to describe sensations: the locutions for description are literally translated in forms like “object x stands in relation y with me.” In classical Chinese, it is impossible to formulate statements taking a form analogous to “X believes P,” where P is a proposition. To translate “X believes the moon is red” one might say “X uses the moon and deems it red.” In short, many of the elements of folk psychology that make up the false belief problem and the conventional theory of mind are not linguistically or culturally universal.

This kind of variation poses a problem for enterprises like the theory-theory which can be understood by analogy with the difficulties that arise with emotions. They come down to this. The theory-theory is an account of developed intentional thinking, complete with an ability to solve the false belief problem. Its answer is that people, all people (with the exception of people with genuine psychological abnormalities, like autistics), somehow acquire this theory. The theory, incidentally, is false, that is to say it is a part of “folk psychology” and cannot be developed into a scientific theory of mind. If normal people do not acquire the theory, the theory-theory is itself false. If they appear to acquire different theories of mind, as the surface differences in language suggest that they do, then these differences need to be shown to be irrelevant to the validity of the theory-theory as an explanation.

It would be convenient for the theory-theory if the language of folk psychology was the same or more or less completely and easily intertranslatable the world over. If so, it could be said that children are aided to acquire this theory (which it is agreed they do not have from birth, since there is a stage in which they cannot solve the false belief problem) by maxims, codes, and the like, i.e. the stuff of culture. If all cultures were uniform with respect to these maxims and codes, it would be possible to list them among the universals of culture, and perhaps even to theorize that the reason they are found in the same form in all cultures is that they reflect a universally acquired psychological “theory” which is in some sense not linguistic but is prior to language or in the language of the mind itself. In this case the surface differences of language could be dismissed, in the manner of “P” above.

13 Needham 1972.
14 As Chad Hansen explains, “The core Chinese concept is xin (the heart-mind). As the translation suggests, Chinese folk psychology lacked a contrast between cognitive and affective states ([representative ideas, cognition, reason, beliefs] versus [desires, motives, emotions, feelings]). The xin guides action, but not via beliefs and desires.” Hansen 1998. At least one philosopher, Herbert Fingarette, has suggested that classical Chinese has no psychological theory. Fingarette 1972, 37-56. This would be a quite literal example of a non-“theory of mind” community.
15 It seems that intentional language is more or less translatable, but that “belief” language is not. As Needham says, “belief makes an extreme contrast. It is not a necessary concept, and it is not a distinct capacity of inner state; other languages make no recognition of a mode of consciousness of the kind, and other people order their lives without reference to any such capacity.” Needham 1972, 146. “Belief,” however, is a key concept in the “folk psychology” which the theory-theory purports to explain.
The fact that the diversity of mental language is apparently greater than this is inconvenient, and inconvenient because it imposes additional explanatory burdens. It could mean that people in different cultures have different mental terms and thus genuinely different theories, each of which happens to enable them to solve the false belief problem, but differently, which is to say that the theories are cultural artifacts. This would mean that the theory-theory is gratuitous for its main purpose, explaining “folk psychology,” because the specific folk psychologies of particular groups are already explained as distinct cultural facts, as something acquired with and as a part of a culture. But this leaves the explanation entirely at the level of culture and cultural adaptation, and treats acquisition as a matter of acquiring culture, which of course does not occur in one developmental step.

So there is a new explanatory burden created by the “cultural” explanation, namely a problem of explaining the apparent universality of the solution of the false belief problem. Perhaps it is not such a difficult one. It does seem odd that all cultures should develop a false theory, and more or less the same false theory, and that this is the adaptive theory. But perhaps this is merely odd. One can deny that the commonalities amount to much more than the result: the common ability to solve the inferential problems called in English the false belief problem. If the whole business of mental language and “theory” is cultural and culturally relative, it is perhaps surprising that all cultures have the equipment to solve the false belief problem in the form that it can arise in that culture. But perhaps it is not. Perhaps it is simply analogous to the fact that all cultures can in some sense think “causally” without all having precisely verbally equivalent concepts of causation. Both ways of thinking are profoundly adaptive (as a culture of autism would not be), so there simply is no extant culture that lacks them.

There are alternatives that can save the theory-theory from reduction to culture, but these create some even more peculiar explanatory burdens. If the universal features of the developmental process explain acquisition of the theory-theory, and the theory is false, then one must explain the cultural and linguistic diversity of folk psychological idioms. But how does one explain cultural differences if the mechanisms of acquisition of the theory are themselves general and psychological? The problem is analogous to the problem of emotion terms, but not precisely analogous, since the terms in question are not primary descriptions, but theoretical terms about non-observable facts. Why would people hold the same “theory” tacitly and non-linguistically, but articulate it in different ways? One could say “error” and see the folk theory of different cultures as bad theory, like bad science or ethno-science. But this is overly complicated. It is more plausible in this case, in contrast to the emotions case, to drop the hypothesis of a universally shared tacit folk psychology entirely. Yet dropping that hypothesis leaves the process by which children develop these cognitive capacities, the empirical findings which have motivated much of the discussion, unexplained, and apparently inexplicable.

What this suggests is that both the culturalist and the psychological universalist approach to theories of mind bring in their train serious explanatory burdens. A combined approach would

---

16 The culturalist account is not free of profound explanatory burdens either. One is this: even a culturalist account is forced to posit some universal cognitive machinery. Recall the metaphors of “inheritance” that recur in Geertz’s definition of “culture.” What is the cognitive mechanism by which “systems of concepts” are passed from one person to another intact, i.e. as a “system?” This question is the theme of my The Social Theory of Practices.
have its own burdens. Chomsky’s linguistics is often the model that thinkers in cognitive science have in mind when they consider the nature of the mechanisms they study. But Chomsky’s approach carefully avoids this problem, by modelling the relation of the universal claims it makes to the actual diversity of languages on the relation between language and dialects: all human languages, it is claimed, are dialects of the universal language. This is not an option open to theory-theorists. Theories are the same only if the terms of the theory are the same, and the problem here is that the terms of folk psychology are not the same, and some of them do not even appear in other languages.

Congeniality and Explanatory Burdens

What sort of argument is this? What we have arrived at is a very crude result. The alternatives of the theory-theory and simulation have been extended to deal with some more or less stable facts that are usually of concern to social theory. Various lines of development of the basic idea of the theory-theory turn out to run into peculiar explanatory burdens, like explaining the actual cultural diversity of terms for beliefs, intentions and so forth. When we arrive at these issues, the difficulties mount up. The contrast between “cultural” concepts and a “universal but tacit” folk psychology turns out to produce particularly serious problems, problems of compatibility between explanations. Put simply, the problem is that two “theory-theories,” a universalistic psychological one and a relativistic culturalist one, attempt to explain the same thing, and thus conflict. Neither can be reduced to or made easily compatible with the other. Neither is implausible.

Simulation and related fundamental mechanisms, however, appear to avoid these problems. They don’t rely on “possession of a theory” as an explanation, and consequently there is no conflict between two theories, of different kinds, each of which explains some of the same things. The problem is to see if the basic mechanisms, together with some plausible starting points, such as emotions and recognition of emotions expressed on faces, can account for the things one wishes to account for. And in this case it is plausible, at least for a large range of things. Actual experience, notably the empirical experiences gained in role-playing, imitation, and simulation tested by experience, provides a great deal of the kind of psychological content that is needed to have the capacity to interact socially. And the capacities thus understood are congenial with an alternative account both of understanding and of culture itself— not necessarily an account in which a culture is a kind of a theory and in which understanding is by virtue of sharing this theory.

I use the term “congenial” for a reason. Though the relationships here are relationships of a particularly loose kind, they are nevertheless meaningful. The relations between the “stable domain of fact” I have appealed to and the very abstract models at stake in the discussion are not so well defined as to make the facts into very persuasive tests of the models. More importantly, the models can be developed in various ways, on the basis of the internal resources of the models, to account for the findings or facts that are claimed to problematic for them. But difficulties do arise, and the conflict between the culturalist notion of the “theory” that explains

\[17\] And, incidentally, to think intentionally about others, which, as we have seen, Needham suggests is universal, in contrast to other elements of “folk psychology.”
intentional language and the psychological notion of the “theory” that underlies the inferences needed to solve the “false belief” problem is a significant one. The less burdened path in the face of this conflict is to avoid the notion of “theory” in either the cultural or the psychological context, and this is congenial with the path I have labeled the “ emulationist” model, which avoids appealing to shared tacit theory, psychological or cultural. It is the path that twentieth century social theory for the most part did not take. Perhaps the balance of explanatory burdens has now changed, and it is the path that should be taken.

References