Draft

Theory of Mind:

Conscious Attribution and Spontaneous Trait Inferences

Angeline Lillard

Lori Skibbe

University of Virginia

To appear in: Hassin, R., Uleman, J., and Bargh, J., The New Unconscious

Cambridge University Press.

The authors wish to thank Cara Johnson, Katie Hamm, Elizabeth Malakie, and Kimberly Patterson for their assistance with the experiment described in this chapter.

"Theory of mind" refers to the tendency to construe people in terms of their mental states and traits (Premack & Woodruff, 1978). If we see someone grimace, we might infer that they are disappointed; if we see a man running to a bus, we probably infer that he is trying to catch it. The word "theory" is applied to such mentalistic inferences for two reasons. First, mental states are unobservable, so their existence is merely theoretical. Second, our body of knowledge about the mind resembles a theory in several ways (Wellman, 1990). One resemblance is that knowledge about minds is causal-explanatory in nature, as are scientific theories. We explain the man running towards the bus as wanting to be on the bus, and thinking it is going to pull away soon. We predict that he will be relieved if he makes it and disappointed if he does not. We also can predict that certain actions might ensue if he does not: that he will try to notify others of his late arrival by telephone, that he will try to hail a taxi, or that he will consult a schedule regarding the next bus. What mediates these further actions are his mental states: wanting others to know his whereabouts, thinking that calling will accomplish that, and so on. Hence we can describe our understanding of others' mental states as a theory of mind.

The first signs of appreciation of mental states appear very early in young children. For example, Baldwin (1991) found that 1-1/2 year-olds are sensitive to an adult's focus of attention when learning words. In these experiments, a child was playing with a new toy, and an adult was nearby, looking into an opaque bucket at another new toy that the child could not see. The adult said, "It's a toma! Look at the toma!" several times. Later, the child was asked to get the toma, and was given a choice of a few toys. Children usually chose what was in the bucket (even though it had not been visible at time of naming) rather than what they themselves had been playing with. Novel words were thus mapped on to the object that an adult labeler was focused on, even when the child was playing with a different object. By 18 months, then, children are sensitive to adults' focus of attention when learning new words.

Even younger children seem cognizant about accidents, goals, and intentions. Tomasello and his colleagues found that 14- to 18-month olds imitated adults' acts when they were followed by a confident, "There!" but not when they were followed by, "Whoops!" (Carpenter, Akhtar, & Tomasello, 1998). Gergeley (1995) and Woodward (1998) have shown sensitivity to various aspects of goals even in the first year. In Woodward's experiments, infants were habituated to an adult repeatedly grasping one of two objects, one placed on the right and the other on the left side of a stage. For the test trial, the placement of the two objects was switched, and a new grasp was directed either at the old object in its new location or at the new object in the old location (where the hand

went previously). In spatial terms, the latter situation was more like what had been habituated to: the hand was moving to the same place. In intentional terms, the former was more similar: the hand was grasping the same object. Infants of 9 months and older dishabituated to the hand grasping the new object in the old location, suggesting they saw it as a different event. This implies that they understood something related to the actor's goals in getting the object. When the same movements were made by a pole rather than a human arm and hand, infants dishabituated in the opposite manner, apparently "interpreting" the situation in a physical, nonintentional way. Rudimentary mentalizing abilities thus begin during infancy.

Theory of mind capabilities evolve throughout childhood, with simple understanding of perception and its links to knowledge emerging in the toddler years, understanding of false belief and mental representation emerging around age 4, and understanding of complex emotions like surprise and pride emerging somewhat later. (For reviews, see Flavell & Miller, 1998; Wellman, in press).

The very early onset and predictable developmental course of theory of mind abilities has led some to suggest that they are supported by innate processes (Baron-Cohen, 1995; Bruner, 1990; Fodor, 1992; Leslie, 1994). These theorists argue that infants are too young to infer matters as complex as mental states, and they are not explicitly taught about them. Indeed, since mental states are invisible, their existence cannot be highlighted in the ostensive manner that many elements of the world are. An adult cannot point and say, "That is a

desire. See the desire?" as they do for so many early conceptual acquisitions (Scholl & Leslie, 1999). A second source of support for theory of mind stemming from an innate process is the ease with which normal adults make theory of mind attributions. When we are told someone tripped while learning a new dance step, we automatically assume he is clumsy (Ross & Nisbett, 1991; Uleman, Newman, & Moskowitz, 1996). Although lacking empirical evidence on this point, when we see a person running to the bus, it seems we cannot help but see them as wanting to get on it. Indeed, we even apply folk psychology to animals and inanimate entities like triangles (Abell, Happe, & Frith, 2000; Heider & Simmel, 1944). Triangles certainly do not have beliefs and desires. In sum, the early and automatic deployment of a theory of mind has led to speculation of stemming from an innate process. The type of process that has been proposed by several theorists is a module (Baron-Cohen, 1995; Fodor, 1992; Leslie, 1994).

Modularity

"A module is a specialized, encapsulated mental organ that has evolved to handle specific information types of particular relevance to the species" (Elman, Bates, Johnson, Karmiloff-Smith, & et al., 1996, p. 36). As is apparent in this definition, which corresponds to the common usage of "module" in cognitive development today, modularity and innateness are often bedfellows, because modules are thought to have evolved in the species. Innateness is not a requirement of modules (Karmiloff-Smith, 1992)—a module could quite plausibly develop over ontogeny without being prespecified—but discussion of

modules in the cognitive development literature often assumes that the module is specified a priori by the DNA.

The view that minds are structured a priori in a way that corresponds to the structure of knowledge can be traced to Plato's Meno. When the slave boy stated various geometric principles, it was agreed that he had not been taught such ideas and that the ideas must therefore be innate. Descartes (1641/1993) later conceded to innate concepts of self and of God, from which other concepts (substance, duration, number, and so on) must be derived. Gall (1835) went further, and perhaps is the first to have postulated specific brain regions being associated with specific functions. By measuring the thickness of neural tissue in specific parts or the brain (as indicated by bumps on the skull, which was soft when the tissue formed) one might determine the strength of specific personality traits in the individual.

Fodor's (1983) influential monograph, the Modularity of Mind, spawned the current wave of interest in modularity. Fodor describes modules as:

1. associated with specific brain regions;
2. showing characteristic patterns of breakdown;
3. having a specific ontogenetic course;
4. processing information very quickly;
5. concerned with specific domains of knowledge; and
6. encapsulated, producing mandatory outputs from given inputs

Fodor's (1983) initial examples of modules are from perception and language, but Leslie (1992) and (more recently) Fodor (1992) have presented arguments that theory of mind also functions as a module.

> The currency of our mental lives consists largely of propositional attitudes, even when we are interpreting the behaviours of others. […] It has been suggested that this capacity—termed a "theory of mind" (ToM)—arises from an innate, encapsulated, and domain-specific part of the cognitive architecture, in short a module.
>
> (Leslie & Scholl, 1999, p. 131)

Some of the six features of modules that Fodor described seem apt, with reference to theory of mind. Theory of mind does appear to be associated with specific brain regions, namely the amygdala, basal ganglia, and parts of the temporal cortex and frontal cortex (Frith & Frith, 1999; Schulkin, 2000). These brain regions typically show elevated rates of activity when participants are asked to consider others' mental states.

There are also specific breakdowns in theory of mind, most notably in people with autism (Baron-Cohen, 1995). Although autism has many features, one persistent and notable one is an inability to decipher others' mental states in the automatic way that most of us do. Oliver Sacks describes how Professor Temple Grandin, a high-functioning autistic woman, was puzzled as a child on the school playground.

> Something was going on between the other kids, something swift, subtle, constantly changing—an exchange of meanings, a negotiation, a swiftness of understanding so remarkable that sometimes she wondered if they were all telepathic. She is now aware of the existence of these social signals. She can infer them, she says, but she herself cannot perceive them, cannot participate in this magical communication directly, or conceive the many-leveled kaleidoscopic states of mind behind it. Knowing this intellectually, she does her best to compensate, bringing immense intellectual effort and computational power to bear on matters that others understand with unthinking ease.
>
> (Sacks, 1995, p. 272).

Many experimental studies of people with autism support this description: people with autism tend to fail tasks involving understanding mental states, at mental and verbal ages at which non-autistic individuals—even those with other psychological impairments—easily pass (Baron-Cohen, 2000).

Theory of mind also appears to follow a specific developmental course, both within and across cultures. For example, American and Chinese children alike talk first about desires, and about six months later, they begin to talk about beliefs, suggesting a common developmental pattern of acquisition (Bartsch & Wellman, 1995; Tardif & Wellman, 2000). Children in both developed and

nondeveloped countries appear to learn around age 4 that people can have false beliefs (Avis & Harris, 1991; Wellman, Cross, & Watson, 2001).

Regarding fast processing, the bus example given earlier suggests that people process information about others' mental states rapidly and effortlessly. Experimental support for the contention that propositional attitude information is processed quickly and automatically is lacking, however there is experimental support for fast processing of another aspect of theory of mind: trait attribution. Although the theory of mind module is normally discussed with particular reference to propositional attitudes, like beliefs, desires, and pretense, theory of mind is more generally construed to include traits (Wellman, 1990). Uleman and others (reviewed in Uleman et al., 1996) have shown that when people consider another's behavior, they appear to immediately attach to the other person a personality trait that could engender that behavior. Experimental support for the fast processing of propositional attitudes is needed.

For some of Fodor's criteria, then, theory of mind abilities do seem to be module-like: Theory of mind processing appears to be associated with specific brain regions, a particular pathology is associated with its specific breakdown, it develops in regular sequence in normal children, and, at least as regards traits, there is evidence for quick and rapid processing.

For other of Fodor's criteria, however, "module" does not seem to be a fitting descriptor for theory of mind. Somewhat problematic is the claim of domain-specificity. Theory of mind attributions are sometimes made outside of

the social domain, directed at inanimate entities (Abell et al., 2000; Heider & Simmel, 1944). One might even say of the sky, "It wants to rain," or of the machine on one's desk, "This computer is stupid." Either one has to place triangles, weather and computers within the social realm (which one might do), or the criterion of domain-specificity is problematic.

The most major concern is the claim of encapsulation: that mentalistic reasoning is encapsulated from other knowledge. As Leslie and Scholl (1999) write, "The essence of architectural modularity is a set of restrictions on information flow […] the modularized processes have no access to any external processing or resources" (p. 133). According to Leslie's (1994) descriptions, human behavior is input to the module and mental state interpretations ("He wants x, he believes x, he pretends x") are output. "The theory of mind mechanism is essentially a module which spontaneously and post-perceptually attends to behaviors and infers (i.e., computes) the mental states which contributed to them" (Leslie & Scholl, 1999, p. 147). We see someone running to a bus, and we automatically infer that he believes it is a departing bus, and that he wants to get on it. Whether we should see such inferences as modular (genetically dictated) or automatic (the result of repeated use, but formed by experience) is at issue.

The parallels between modular processing and automaticity are notable. Bargh (1994) described the "four horsemen of automaticity": efficiency, lack of awareness, lack of control, and lack of intention. Modules share these same

features, but the existence of the theory of mind module, according to Leslie (1994) and Baron-Cohen (1999), is due to its genetic predesignation. In contrast, automatic processes are considered habits of mind, produced by repeatedly processing certain types of information in the same manner (Bargh & Chartrand, 2000).

Attribution research would traditionally seem to support the view that theory of mind is modular. For example, the tendency to attribute personality traits as explanations for behavior (Jones & Davis, 1965; Ross & Nisbett, 1991) supports the notion that a theory of mind module mandatorily processes others' behaviors in internal terms. Likewise, the tendency to spontaneously infer traits when considering others' behaviors (Uleman et al., 1996) supports the idea of mandatory modular output.

However, there is a great deal of other evidence against the assertion that mentalistic explanation is truly mandatory and encapsulated, as would be expected of innate modules. Anthropologists and psychologists working in different cultures around the globe have described a wealth of variation in how minds, mental states, and actions are conceived (Lillard, 1997; 1998). These differences do not necessarily indicate how the mind is spontaneously conceived, but they do beg the question of why, if our spontaneous construals all resemble European-American views of the mind, there is such widespread variation in more considered (conscious) views the world over.

For example, the European-American concept of mind, as a self-contained, thinking entity (Geertz, 1984), contrasts with the Illongot rinawa, which shares some features with mind but leaves the body during sleep, is possessed by plants (but leaves when plants are processed), and is much more an organ of social context (Rosaldo, 1981). Cultures vary greatly also in the attention paid to minds, with Europeans and Americans seemingly the outliers in the world's cast. We pay enormous need to minds, as evidenced by the number of words in our languages that specify mental constructs, the existence of a field of Psychology, the psychological nature of parenting guides, and so on. The Chewong of Malaysia are reported to have only five terms for mental processes, translated as want, want very much, know, forget, and miss or remember (Howell, 1981; 1984). Anthropologists in many cultures have commented on the people of the cultures they study claiming one cannot know others' minds, refusing to speculate about others minds, preferring not to discuss others' minds, and simply attaching comparatively little import to minds (1984; LeVine, 1979; Mayer, 1982; Ochs & Schieffelin, 1984; Paul, 1995; Poole, 1985). Even in terms of legal responsibility, in other courts of the world intention is often not what matters when determining retribution for a crime; what matters instead is the degree of harm caused by one's actions (Hamilton & Sanders, 1992; Paul, 1995). If a module is automatically, mandatorily outputting mental state concepts, these different degrees of attention seem odd. Whole cultures do not choose to ignore

the output of vision modules or a language acquisition device; it would be impossible.

There is also wide variation in what some behaviors are attributed to as well, getting more squarely at the content of module output. If all people (except those with autism) are endowed with a brain module that perceives behavior and automatically interprets it in propositional and trait terms, we would expect to see reliance on such terms to be fairly prevalent everywhere. But again, this does not seem to be the case. In one early study demonstrating this, Miller (1984) asked Hindu Indian and Chicago children and adults to think of good and bad behaviors performed by people they knew well, and for each, to explain why the people did them. Interestingly, although by adulthood Chicagoans tended to come up with personality-trait reasons for behaviors, Indians tended to come up with situational causes. One might be concerned that this was because people in the different cultures thought of different behaviors to begin with, but in a later experiment, even when provided with behaviors (for example, a case in which an attorney left the scene of a motorcycle accident he had caused), Americans tended to give trait reasons (citing that the lawyer was a villain, for example) and Indians tended to give situational ones (the lawyer had a duty to be in court). This naïve dispositionalism on the part of Americans tends to be wrong—people's actions are strongly influenced by the situation—yet it nevertheless is how Americans tend to view behaviors (Ross & Nisbett, 1991).

We recently revisted this result with urban and rural American and Taiwanese children (Lillard, Skibbe, Zeljo, & Harlan, 2001), reasoning that rural American children might be different from the urban ones that are most often experimental participants. Using Miller's basic procedure, we found that elementary school children from rural Pennsylvania and rural Virginia tended to use situational explanations even more than did Taiwanese children. For example, a rural 7-year-old said the reason someone had shared a bicycle with her was because she had not brought her own bike. In contrast, 7-year-olds from a more urban area tended to explain behaviors with reference to mental states and traits of the actor, like "Because she wanted to help me' or "Because she was nice." This pattern has held up in two different experiments drawing on different rural geographic regions. Across all groups, mental state explanations were about four times as common as trait ones, but for rural American children, situational explanations were dominant.

Several other demonstrations of variations from the standard American attribution pattern have been accrued. Morris and Peng (1994) presented Chinese and American high school and graduate students with nonsocial (dots) and social (fish) cartoons. For both types of cartoon, a group of the entities approached a single entity, stopping at the point of contact. The single entity then moved forward. Participants were asked to rate the extent to which the movement of the single entity was caused by internal factors, and the extent to which it was caused by external factors. All of the American students and the

Chinese graduate students (from Taiwan) tended to claim the movement was caused more by internal ones, and less by external forces, than were the Chinese high school students (from mainland China). This effect was then also noted in an everyday realm: newspaper articles. Although describing the exact same Chinese and American murderers, Chinese-language papers in the United States tended to portray the murderers as a victim of their situations ("He as a victim of the China Top Students Education policy", "He had a bad relationship with his advisor") whereas English-language papers tended to portray the murderers in trait terms ("There was always a dark side to his character.") Lee, Hallanan, and Herzog (1996) have recently shown this same effect for sports articles in papers in Hong Kong and the United States. In line with their hypothesis, the effect was not seen in editorials, which they reasoned reflect deeper consideration of alternative viewpoints, but was seen in newspaper articles, thought to be more spontaneous (see Uleman, 1999).

Anthropologists' reports concur with these experimental findings. Lillard (1998, p. 15) reviews ethnographers' reports of people designating social causes for others' behaviors:

> If an Ifaluk person goes into a jealous rage, the person who left her or his valued possessions in plain sight of another is viewed by Ifaluk as being the cause (Lutz, 1985). In EA [European-American] culture, it seems more likely that the person exhibiting the rage behavior would be seen as responsible

> because people are primarily in charge of their own behavior. Hamilton and Saunders (1992) provide evidence for this: In assigning responsibility for unfortunate outcomes, Americans do not consider the effect of other people's influence as much as the Japanese do. For the American Cheyenne, behavior is seen as motivated by relationships more than by individual wills (Straus, 1977) . One's actions are generally explained by reference to someone else's actions or to one's relationship with some other ("I hit him because he hit her…I drank with him because he is my cousin"; p. 33). Straus described a social worker's frustration that the Cheyenne do not take responsibility for their actions but instead make excuses. However, Straus emphasized that these are not excuses to the Cheyenne: They truly are causes. Likewise, Briggs (1970) reported that for the Utku (Northern Territories), actions are explained in terms of other people's desires, not their own. Harre (1981) also wrote that 'many travelers have reported the extraordinary degree to which Eskimos seem to be influenced by their fellows. When one weeps, they all weep.'

Other cultures turn more often to ethereal causes of behavior than do mainstream Americans, claiming that their actions were caused by gods or spirits. Evans-Pritchard (1976), for example, tells of a man attributing his own

tripping on a stump to witchcraft. When asked whether he shouldn't take some of the blame, since of course he had not seen the stump and was clumsy, he insisted that had witchcraft not been operative, he would have seen the stump. Americans are much more prone to assume personal responsibility. (There are certainly exceptions to this, however, with the recent propensity of Americans to sue others for damages that in another time would be attributed to the suer or bad luck, and with a subset of Americans invoking the hand of God to explain some if not all events, as described in Weeks & Lupfer, 2000.)

To summarize, in their conscious explanations for behaviors, most Americans appear to differ from many other cultures in the world by locating explanations inside the person, in terms of theory of mind constructs like beliefs, desires, and traits. These findings are not perfectly consistent across studies, and there may be some cultures—Taiwan and Korea in particular—in which the difference is not as frequently observed or as strong (Choi & Nisbett, 1998; Fiske, Kitayama, Markus, & Nisbett, 1998; Krull et al., 1999; Lillard et al., 2001; Morris & Peng, 1994). Yet certainly it has been widely noted as a cultural difference in both experimental and observational work (Markus & Kitayama, 1991). If theory of mind were mandatory and encapsulated, we think we would see much less variation in these conscious interpretations. Importantly, cultural variation in behavior interpretation extends to unconscious interpretation as well, which mandatory, encapsulated modular processing does not allow for. Automatic

trait processing is particularly a feature of individualistic cultures (Duff & Newman, 1997; Newman, 1993; Triandis, 1994; Zarate, Uleman, & Voils, 2001).

Drawing on the well-replicated finding that people from more collectivist cultures use fewer trait constructs in person descriptions than do people from more individualistic ones, and the finding that trait thinking is particularly strong in middle childhood (Livesley & Bromley, 1973; Shantz, 1983), Newman (1991) examined spontneous trait inferences in Anglo-American suburban versus Hispanic urban 5th graders. He found that only the Anglo children appeared to have spontaneously made trait inferences in a word recognition test. In later work, Duff and Newman (1997) gave adult participants the idiocentrism scale developed by Triandis and his collegues (Triandis, Bontempo, Villareal, Asai, & et al., 1988) along with an inference task designed to reveal the extent to which participants spontaneously infer either situation or trait causes of behaviors. They found that idiocentrism was positively correlated with trait, but not with situation, cued recall (see also Newman, 1993).

In a further example of this cultural difference, Zarate, Uleman, and Voils (2001) examined spontaneous trait inference with adult Hispanic and Anglo-American subjects. In a first experiment, they found that Euro-American college students had significantly faster reaction times to trait words in lexical decision-making tasks when trait words were primed by trait-implying sentences. No such difference was seen for Hispanic participants. A second experiment repeated this effect with savings in relearning tasks. In sum, there is strong

evidence for cultural variation in the unconscious process of spontaneous trait inference.

Findings concerning cultural variation in both behavior explanations and spontaneous trait inference are at odds with a modular theory of how we interpret the behaviors of others (Lillard, 1999). If modular processing is mandatory, and if all normally-functioning people have a theory of mind module that interprets behaviors, then all people should automatically infer mental states and traits in response to behaviors. Just as we cannot help, when viewing Muller-Lyer stimuli, but see one line as longer than the other, we should mandatorily see behaviors as rooted in theory of mind causes (mental states and personality traits).

Scholl and Leslie (1999) have taken exception to the view that variation in conscious behavior explanations is problematic for modularity theory. They point out that how one explains behavior after the fact, on an attribution task, and how one perceives behavior on line, can be quite different: "you can't decide not to interpret lots of situations as involving intentional agents, although you can ignore the interpretation" (p. 135). (They have not responded to the STI literature, but might disregard it due to its concern with traits rather than mental states, as discussed later.) According to their claim, regardless of how rural children subsequently explained the behavior, when they originally perceived the behavior, they perceived it in theory of mind terms. Everyone who sees someone running for a train might perceive them as wanting to get the train,

although when asked why the person was running, some might be more apt to say, "Because he wanted to catch the train" and others might be more apt to say, "Because the train was leaving." This is possibile, and it raises an interesting issue: how does a person's style of explaining behaviors after the fact relate to a person's on-line encoding of those behaviors? In other words, how does spontaneous trait inference relate to deliberate attribution?

The literature suggests some divergence (Miller, Smith, & Uleman, 1981); indeed when spontaneous trait inference works less well, it appears to be because behaviors were too considered (Zarate et al., 2001; Zelli, Cervone, & Huesmann, 1996). "Spontaneous impressions are guided by chronically accessible constructs, whereas intentional impressions are guided more by temporarily activated goal-relevant constructs and procedures, and by implicit theories (about the meanings of actions, relationships of traits to each other, etc.)" (Uleman, 1999, p. 146).

One way to address this issue, and Scholl and Leslie's critique, is to examine the relationship between the strength of a person's tendency to make spontaneous trait inferences (STIs) (Uleman et al., 1996) and the strength of a person's tendency to explain behaviors in theory of mind terms. If automatic, unconscious theory of mind processing is modular, and after the fact explanations of behavior are not, there should be no consistent relationship between spontaneous trait inference and how behavior is explained. On the other hand, if theory of mind processing is an automatic process, instilled

through years of exposure to and practice of certain forms of inference, then the degree to which a person makes spontaneous trait inferences and uses theory of mind constructs in conscious behavior explanation should be related.

To test for the relationship between STIs and behavior explanation styles, we gave 45 undergraduate participants both types of task. Participants were recruited from the University of Virginia psychology classes, and were tested in groups of up to 10 persons.

The STI task used a cued recall procedure, based on Tulving's encoding specificity principle (Tulving & Thomson, 1973). The stimuli were previously used by Duff and Newman (1997, Exp. 2). Participants were shown 10 sentences, projected one at a time for 6 seconds, on the wall by an overhead projector. Prior to the first sentence being projected, participants were told that their memory for these sentences would be important later in the experiment. The first and last sentences were filler sentences used to reduce primacy and recency effects in memory (Anderson, 1990). The 8 remaining sentences were presented in a random order, followed by two 60-second filler tasks (list as many of the U.S. states as possible, and write any thoughts you had as you were viewing the sentences.). The filler tasks were intended to clear the participants' short-term memory of the sentence content, and, in the case of the second filler task, to check whether participants consciously attempted to use internal or external cues to memorize sentences.

Participants were asked to recall the sentences, using one of two recall sheets. Each recall sheet had a list of 8 cues on it, 4 of which were situational and cued half the sentences, and the other 4 of which were internal and cued the remaining sentences. An example of the sentences and associated cues is "The engineer/ picks up /the papers/ from the floor." For this, the internal cue was "neat" and the external cue was "dropped them." Another example is, "The accountant /gets the day off work /with some fellow employees/ and takes the orphans to the circus," which had an internal cue of "caring" and an external cue of "job obligation." The slashes in the sentences correspond to sentence parts for which recall was scored.

The cues were derived through extensive pretesting at other universities. When asked to explain these sentences, in pilot work, about 50% of undergraduates had tended to spontaneously supply a situational reason , and about half had tended to spontaneously supply a trait reason. The cues used in this study were the reasons most commonly given by undergraduates in that pretesting (Duff & Newman, 1997). Participants were given 6 minutes to record any parts of the sentences that they could remember, and were told that if one of the cues on their sheet helped them to remember any part of a sentence, then they should record that sentence part next to the relevant cue. Participants were also told that if a sentence part was recalled with no help from the cues then they should write it on a blank line at the bottom of the recall sheet.

Following the cued recall task, three types of behavior explanation tasks were employed: Forced choice, rating scale, and open-ended. Participants first answered 16 forced-choice questions concerning other people's behaviors. For each question, participants read a description of someone engaging in an action and were asked to choose which of two explanations was more likely. One of the two was judged to be internal, and the other external, according to a coding system developed for a prior study (Lillard et al., 2001). Internal explanations involved reasons internal to the actor: personality traits, beliefs, desires, emotions, and so on. External reasons were those that lay outside the actor: the situation the actor was in, another person, a relationship, a role the person had to play, and so on. Kruglanski (1975) and Ross (1978) pointed out the difficulty in making internal/external splits, in that internal reasons are often embedded in external ones and vice-versa. We take the point, yet believe that when a respondent chooses to emphasize the actor's internal qualities versus circumstances in giving explanations, important distinctions are being preferred (for further discussion, see Lillard et al., 2001). Opting to explain a behavior with reference to the actor, versus with reference to the situation, is a non-arbitrary choice which can reflect the person's schemas and world view. Evidence for this position would be accrued if internality of behavior explanation were found to be related to STI.

The explanation choices were selected from responses given by a different group of undergraduates who participated in a pilot study. In the pilot study, 85

undergraduate participants had been given the same actions to explain, but in an open-ended format. Their most frequent responses served as choices in the present study. Examples of items on the forced-choice tasks are: "The girl gave cookies to her neighbor," with the choices being, "The neighbor just moved in" (external) or "The girl wants to be kind," (internal) and "He hit another person in the mouth," with the choices being, "He disliked that person." (internal) and "They had an argument" (external).

Next, for the rating-scale, participants read 6 sentences describing actions. Each sentence was followed by four reasonable explanations for each behavior. The sentences and the explanations were selected from responses given in the same pilot study just mentioned. Of each set of 4 explanations, 2 were internal in nature and 2 were external. For each explanation, participants were asked to rate how likely each explanation was on a four-point Likert scale, with 1 being not likely, and 4 being very likely. An example of an item from this task is "Sue helped Mary with her schoolwork." Participants were asked to rate how likely it was that "They are friends," and how likely it was that "Sue likes to help others."

Finally, for the open-ended task, participants answered two open-ended questions like those used in the prior study: to think of and explain a good behavior and a bad behavior. Specifically, they were asked to, "Think of a good [bad] behavior a friend of yours engaged in during the past two weeks. What did your friend do and why did he/she do that?"

A further goal of this study was to shed light on what kinds of person variables are associated with arriving at more external or internal reasons for behaviors. In other work we considered several possible contributors (Lillard et al., 2001). These speculations were based on the fact that American children from less densely populated areas tended to use internal explanations for behaviors less so than did children from more urban areas. However, the rural/urban factor was confounded in these experiments with income and education levels of the parents. The present experiment asked participants about the rural/urban nature of their childhood communities and about income and education levels in order to examine how these factors might interplay with how behaviors are explained.

Regarding more rural versus urban communities, in more rural communities, there are fewer people among whom to hide, and perhaps a greater sense of responsibility to the group (Paul, 1995). This could lead to people in more rural communities privileging external causes of behavior—doing things because of rules, or because they are the right thing to do. Yuill (1992) explained the use of fewer personality trait attributions in rural communities as being due to residents being better intuitive psychologists, because they interact more closely and frequently with the same small group of others. On the other hand, Hollos (1987) found that children in more rural communities performed much more poorly on role-taking tasks than children from urban areas, whereas performance on logical tasks, like conservation, was

similar. Role-taking has been linked to theory of mind in many ways, from pretending to be other people to simulating the mental states of others to the perspective-taking studies that partially instigated theory of mind research (Astington, Harris, & Olson, 1988; Flavell, 1992; Harris, 2000; Lillard, in press). To help us determine whether they were from more rural or urban environments, most participants filled out a brief demographic survey asking them to rate, on 5-point scales, the rural vs. urban nature of their childhood community.

Another factor that might be at issue with regard to behavior explanation is income level. We reasoned that the predominant income level of their community could influence the degree to which people view themselves as victims versus controllers of circumstances. People with more money tend to have more choices in America, and people with less are more restricted (see Lillard et al., 2001). Those who are more at the mercy of their external circumstances would tend to focus more on external causes of behavior, we reasoned, whereas those with more means would tend to see behaviors more as emanating from beliefs and desires.

Parents' education level (1 being "some high school" and 5 being "graduate degree") was also thought to be an important factor for behavior explanation, for the same reason as income was: those with more education would see more possibilities, and would see internal factors as leading to their choosing among those possibilities. In another vein, Hollos (1987) suggests that

the lack of talking that she found rural children were engaged in could be a factor in their poor role-taking skills. Harris (1996) has suggested that conversation could give rise to the idea that others have distinct mental perspectives that lead to their actions, and more educated parents are known to talk to children more than less educated parents. Hence lack of conversation among rural families might also explain decreased use of theory of mind explanations.

Finally, strength of religious conviction was also examined. Weeks and Lupfer (2000) found that more religious individuals show a stronger tendency to endorse God as a cause of events. In keeping with this, we reasoned that stronger religious convictions could be associated with a tendency to place power outside the self, in the hands of God.

Commensurate with these variables was another variable that more directly assesses what several of the demographic variables were intended to get at: locus of control. Participants completed the Rotter (1966) I-E Scale and the Levenson (1981) Internality, Powerful Others and Chance Scale. Locus of control refers to whether people believe that most outcomes are the result of factors that they themselves control (hence an internal locus of control) or are the result of factors over which they have no control (hence an external locus of control) (Lefcourt, 1991). The Rotter I-E Scale (1966) is the classic measure of this construct, and was used because of its solid foundation in the literature. Levenson's (1981) scale builds on Rotter's (1966), and is particularly well-suited

to the specific issues addressed in this experiment. In particular, it addresses separately the respondent's beliefs about his or her own power over events, beliefs about other people's power over outcomes, and beliefs in chance or luck.

It was expected that the income and education level variables would actually be commensurate with locus of control, and that locus of control might also subserve behavior explanation choices. In particular, people who see behaviors as stemming more from inside the actor also were hypothesized to have a more internal locus of control, and those who see behaviors as stemming from situations were hypothesized to have a more external locus of control. Other research has shown that locus of control varies reliably with rural/urban background as well as with other dimensions that were confounded in our sample, namely family education background and income level (Gurin, Gurin, & Morrison, 1978) (but see Witt, 1989; Zimbelman, 1987).

Half the participants received the Rotter scale first, and half received the Levenson scale first. The Rotter (1966) I-E Scale is a forced choice questionnaire. Participants read a series of pairs of statements, with each pair expressing divergent views about an issue, and then chose the situation that most closely corresponded to their own view. For example, one pair of statements reads, "Many of the unhappy things in people's lives are partly due to bad luck" (control is placed outside the person) and "People's misfortunes result from the mistakes they make" (control is seen as being inside the person). Participants circled the statement that they agreed with more. The Levenson (1981)

Internality, Powerful Others, and Chance Scale also presents statements about the causes of events. However, in this case the statements systematically address internality, powerful others, and chance, and rather than asking for agreement as a forced choice, participants are asked about their level of agreement with each statement. For example, one statement reads, "Whether or not I get to be a leader depends mostly on my ability." Participants were asked whether they strongly disagree, disagree, slightly disagree, slightly agree, agree or strongly agree with each statement. The 24 statements constitute the three different subscales, and are provided in a predetermined and intermingled order on the single scale.

Coding was done as follows. For the Cued Recall task, each sentence was divided into four parts in accordance with criteria provided by Duff and Newman (1997) and elaborated on by Newman (personal communication, May, 2000). Internal cued recall was calculated as the number of internally-cued sentence parts perfectly recalled, divided by the total number of sentence parts perfectly recalled. Sentence parts were counted as internally cued when the sentence for which they were a part was cued internally on that sheet. They were counted as such even if not written on the line with the cue, since recall can be cued unconsciously. Truly uncued recall could be expected to appear equally on both sheets and hence would be equivalent across the sample. Partially or imperfectly recalled sentence parts were not scored, making the scoring criteria stringent, as in Duff and Newman (1997).

For the forced-choice behavior explanation task, the total number of internal explanations selected was divided by the total number of answers provided by each participant, giving a percent-internal score. For the scaled behavior explanation task, responses for each category of explanation were averaged, producing a mean likelihood score for both the external and the internal explanations. For the open-ended behavior explanation questions, the total number of internal explanations was divided by the total number of explanations given, giving a percent internal score.

The locus of control scales were coded in their standard manners. The subject's score on the Rotter (1966) I-E Scale is the total number of external statements that the subject endorsed, out of a total of 29 statements. On the Levenson (1981) Internality, Powerful Others and Chance scales, subjects receive three separate scores, one for each subscale, ranging from 0 to 48 for each.

The demographic survey revealed that the sample was fairly homogenous. This was not expected, given that the university where testing took place draws a diverse population, from rural mountainous regions to the urban areas around Washington, D.C. However, in our sample, only one participant was from a very rural community (rated as 1 on the 5-point scale); most were from suburban areas (n=17, 3 on the scale) or more urban ones (n = 10 marking 4 or 5 on the scale). Education levels of their parents were also high. Every parent had finished high school, all but 3 mothers and 1 fathers had some college education, and 20 fathers and 14 mothers also had post-graduate

education. The mean education level was 4.2 on the 5-point scale. The income levels of their communities when growing up were rated as upper-middle income, with a mean of 3.8 on the 5-point scale. None were rated as lower- or middle-lower in income (1 and 2 on the scale). Religiosity was normally distributed with a majority (n = 15) declaring themselves moderately religious, 11 deeply or fairly committed, and 10 barely or not at all religious. Correlational analyses revealed that these factors were not related to any of our other measures, as might be expected from the homogeneity of the sample. The homogeneity of the sample was unexpected, and made the test even more stringent, since such a sample would be expected to use a preponderance of internal explanations (based on Lillard et al., 2001).

On the cued recall task, participants recalled an average of 21% of the sentence parts (6.4 of 32), comparable to results found in other research of this type (Uleman et al., 1996). The mean percentage of recalled sentence parts that were from internally-cued sentences was 53 (range = 0 to 100; SD = 27), indicating that on average participants recalled sentence parts corresponding to internal and external cues on their recall sheets about equally. This was to be expected given that the stimuli were preselected to be ones that undergraduates spontaneously infer trait and situation explanations for about half the time. The interesting issue is how recall cued by external or internal cues corresponds with how each participant consciously explained behaviors.

For the forced choice behavior explanations questions, 51% of choices were internal (range = 25 - 88%, SD = 15), indicating that on average participants chose internal and external explanations equally often when those explanations were provided, but showing sufficient range for possible correlation with the spontaneous trait inference task.

For the rating scale portion of the behavior explanation task, the mean rating for internal explanations was 2.99 (range = 2.50 to 3.58; SD =0.27), and the mean rating for external explanations was 3.02 (range = 2.25 to 3.58; SD =0.27), suggesting internal and external explanations were rated about equally plausible overall. Participants found all explanations fairly likely on average when both were listed for them and they did not have to choose between them (as they had for forced-choice). Indeed, ratings on the internal and external scales were significantly correlated, $r = .38$, $p < .05$, suggesting that the ratings reflect participants' individual tendencies to use more extreme versus mid-points on the scale rather than propensity to construe others' actions in internal or external terms. As Solomon (1978) suggested, internal and external ratings are not necessarily inverse, and this correlation makes that point clearly. This aligns with prior work (Zarate et al., 2001; Zelli et al., 1996), and makes it less likely that one might find correlations between spontaneous trait inference and this task.

For the open-ended behavior explanations, the typical adult American finding obtained: the mean percentage of internal responses was 78 (Range = 0 to 100; SD = 26), demonstrating that, when the participants were asked to come up

with their own explanations for behaviors, they usually postulated internal ones. The range was sufficient to allow for correlations with spontaneous trait inference. One participant used no internal explanations, one used 33% internal explanations, and 23 used 100% internal explanations, with the remaining participants between 50 and 100% internal.

Cronbach's alphas were calculated for all three behavior explanation measures, and were .59, .47, and .42, for the ratings scale, forced-choice, and open-ended measures, respectively. Note that the open-ended task concerns only two items. For the other tasks, items sometimes pull for internal or external responses, hence one might not expect particularly high alphas for tasks of these sorts.

Correlations between the behavior explanation and spontaneous trait inference tasks were examined next. First, there was no significant correlation between STI and the rating scale task, as expected from the results with the rating scale: participants tended to highly endorse both the internal and external explanation for each behavior.

A significant correlation was obtained between STI and the percentage of internal explanations for behavior for the forced choice explanation task ($\underline{r} = .34$, $p < .05$). When forced to choose between an external and an internal explanation for each behavior, participants who tended to choose the internal options also tended to have spontaneously inferred traits from the brief descriptions of behaviors they read earlier in the study.

There was also a significant correlation between STI and the percentage of internal explanations participants provided on the open-ended behavior explanation task ($\underline{r}$ = .37, p = .01). Those who tended to provide more internal explanations also tended to infer traits more, as evidenced by their higher degree of trait-cued recall.

These two correlations were obtained despite the fact that the sample was not a very diverse one. Although the range of internal explanations provided on the open-ended task was as large as possible (0 to 100%) the majority of respondents provided mostly internal responses, averaging close to 80% internal. Had rural populations been better represented in the sample, perhaps even stronger correlations would have been obtained.

Interestingly, although the behavior explanation scores were both independently related to internally-cued recall, they were not correlated with each other. Whereas many (n= 23) participants were 100% internal on the open-ended behavior explanation task, none were 100% internal on the forced choice task. Indeed, only 6 participants chose internal choices on 75% or more of the forced-choice items. This makes the point that when participants are shown external options, they are apt to think them plausible, although they might not have come up with them spontaneously. Performance on the rating scale task supports this, since most participants endorsed both internal and external explanations as likely. Performance on the rating scale task was not correlated with performance on the other theory of mind tasks.

The correlation pattern suggests that some participants' internal orientations were revealed by their answers to the open-ended questions, and others' were revealed by their choices on the forced-choice questions. The reasons for this are a topic for further research. One possibility is that participants scoring high on STI come from two camps, which might be dubbed Libertarians and Psychologists. The Libertarian camp emphasizes free choice, and explained behaviors internally on the forced-choice measure. Supporting this, Miller et al. (1981, Exp. 2) obtained evidence suggesting that when American respondents choose trait options on forced-choice tasks, they are not so much claiming that the trait caused the behavior, as that the behavior was freely chosen. Despite their presumed idiocentrism (suggested by their choice of trait options), Libertarians might not spontaneously use internal reasons to describe those choices. The other camp, the Psychologists, use many internal explanations on the open-ended task, score high on STI, and yet do not necessarily opt for the internal options in the forced-choice task. Such people habitually consider all manner of internal constructs, including traits and mental states, in considering what causes behaviors, but when faced with the external option in the forced choice task, their knowledge that situations often drive behavior leads them to often choose the external option. .

This brings up another important refinement to the results. Internal scores on the open-ended behavior explanation measure reflect both trait and mental state explanations. The actual breakdown of the results is that participants gave

mental state explanations about 64% of the time, and trait ones about 14% of the time. In other words, the vast majority of participants' open-ended explanations are trait ones. As is pointed out by Malle (this volume), prior experiments have tended to focus on traits to the exclusion of mental states (e.g., Miller, 1984). Although Heider's (1958) seminal monograph considered both mental state and trait interpretations, attribution research has focused mainly on traits. Yet our participants rarely volunteered trait explanations, as compared to mental state ones. Indeed, spontaneous trait inference related strongly to mental state explanations on the open ended task ($r = .45$, $p < .01$) and was unrelated to trait explanations alone. The lack of relation to trait explanations could be primarily due to how few trait explanations were provided overall, yet still the relation between STI and belief-desire reasons was striking.

One goal of this study was to provide evidence for what personality or demographic variables are associated with explaining behaviors in more internal or external terms. As stated, the demographic variables yielded no information on this, possibly because of insufficient variation. The personality variable of locus of control was tested to examine whether it might undergird avoidance of internal explanations in rural populations (Lillard et al., 2001). Scores on the locus of control scales were what one would expect for the population, given the results of other studies (Lefcourt, 1991). The mean score on the Rotter was 11.49, with a range of 5 to 19 and a SD of 3.52, hence the sample tended towards an internal locus of control. The mean on the Levenson Internal subscale was 32.96,

with a range of 16 to 42., and a standard deviation of 4.98, suggesting the sample was rather internal, overall. The Powerful Others subscale yielded a mean of 17.96, with a range of 3 to 35.0, and a SD of 7.88. The Chance subscale yielded a mean of 21.44 with a range of 5.0 to 34.0 and a SD of 6.31.

Scores on the Internal subscale of the Levenson correlated negatively with scores on the Rotter, $r = -.34$, $p < .05$, suggesting validity of these measures. Contrary to expectations, the locus of control scores did not correlate with any behavior explanation measures. It may be the case that our hypothesis is simply wrong, and that other factors, like individualism/collectivism, undergird the rural-urban findings and propensities to regard internal or external factors as responsible for behaviors (Duff & Newman, 1997; Newman, 1993; Zarate et al., 2001). Alternatively, it may be that the college sample we tested did not have sufficient variation in locus of control to allow correlations to be revealed with this size of sample.

Although significant relations with locus of control were not seen, the relationship between spontaneously inferring traits in response to behaviors, and the tendency to explain behaviors in internal terms, emerged clearly. Our findings come full circle with others in the literature. Miller (1984), Morris and Peng (1994), and others have found that people from more collectivist cultures tend to explain behaviors in more external, situationist terms than do Americans. Others (Duff & Newman, 1997; Newman, 1993; Zarate et al., 2001) have found that more individualistic participants, like Americans who score high on

idiocentrism measures, are more apt to encode behaviors with traits, whereas collectivists are more likely to encode behaviors with situations. The present study demonstrates that within a single college sample, those who were more apt to explain behaviors with reference to internal factors like traits and mental states also were more apt to make STIs. This result was obtained both when participants were asked to come up with explanations for behaviors, and when they were asked to choose between two plausible explanations.

The finding that how people explain behaviors—in external or internal terms—is related to how they encode behaviors—according to situations or traits—suggests first that how one chooses to explain behavior is not an arbitrary semantic decision (Kruglanski, 1975; Ross, 1978). It relates importantly to unconscious person perception processes.

Second, it suggests that theory of mind may not arise from modular processes. Automatic person inference processes were related to more considered ones. Yet Leslie's theory of mind module, as discussed earlier, has been proposed to explain our spontaneous and quick invocation of propositional attitudes—beliefs, desires, and pretense, for example—to explain behavior, not traits. Leslie found behavior explanations irrelevant to the module because they were after the fact. STIs, likewise, might be irrelevant because they concern traits, not mental states. Ideally, a tasks similar to the STI task could be constructed, aimed at propositional attitiudes. However, clear how one would do so regarding beliefs and desires. Take the bus example given earlier. If

participants saw a sentence reading, "The man raced towards the departing bus", what would be the correct propositional-attitude cue word? "Want" for "wants to catch"? "Desires? "Think"? How to use such cue words effectively across multiple sentences is also very problematic, since "desire" would presumably apply to several situations in a way that "clumsy" does not. Many mental states are ubiquitous; specific traits are not.

The question arises then, does this data address the modularity account of mindreading? If a theory of mind module is strictly limited to spontaneous propositional attitude interpretations of behaviors, then it probably does not. There is no measure, to our knowledge, of spontaneous mental state inference of the unconscious sort Leslie's theory appears to require. The open-ended behavior explanation is the closest we have to spontaneous explanations, and this task is done consciously. If theory of mind is limited to the propositional attitudes, we believe the onus is on modular theorists to come up with a task that assesses their spontaneous, encapsulated use.

Theory of mind is generally seen as more encompassing than belief-desire reasoning, and includes traits. If modular processes are thought to provide social cognitive interpretations more generally, then they should be responsible for trait attributions as well as mental state ones. As pointed out earlier, automatic processes are akin to modular ones in several important ways: they are both fast and efficient, and their output is predictable. "The essence of architectural modularity is a set of restrictions on information flow," (Scholl & Leslie, 1999, p.

133). "TOM interpretations […] seem to be relatively […] fast (they typically occur without lengthy and effortful reasoning), and mandatory (you can't decide not to interpret lots of situations as involving intentional agents, although you can ignore the interpretation" (p. 135). Spontaneous trait attribution in this way looks like a modular process: one sees the behavior, and one infers the trait. However, it cannot be an innately established one, given as part of "our genetic endowment" as is claimed for TOM (Scholl & Leslie, 1999, p. 134) since it is not universal, but appears to vary in culturally-informed ways. The information flow is not restricted; it is open to the influence of culture. The closest test we have, then, of a modular process for social interpretation processes, is more supportive of an automatic than a modular account of that process. The correlations show that how one explains behavior after the fact (mostly with mental state interpretations, in the form of propostiional attitudes) is clearly related to that process, the tendency to infer traits spontaneously when reading about a behavior.

This said, it should be noted that proponents of modular theory do not always strongly endorse a strong version of their modules. Fodor (1983) states, "Whenever I speak of a cognitive system as modular, I […] shall always mean to some interesting extent" and Scholl and Leslie (1999) state that the restrictions on information flow are always "a matter of <u>degree"</u> (italics in original; p. 133). Yet once one begins loosening the criteria for what constitutes a module, one loses the very essence of a module. If a module lacks restrictions on information flow,

then why call it a module? Because such soft positions leave one with nothing to evaluate, in this chapter we have addressed the strong form of innate modules.

A second note is that Scholl and Leslie (1999) discuss TOM as having an innate basis; they are willing to concede that over development, cultural "extramodular" (p. 137) processes use the modular output in various ways. Yet "The essential character of ToM a person develops does not seem to depend on the character of their environment at all" (p. 136). We agree that the essential character of theory of mind—some concepts that map at least roughly to our concepts belief, desire, seeing, feeling, and so on--and an understanding of individuals as having some degree of agency (although the degree and circumstances under which that agency exists may differ) is universal (Lillard, 1998). But we think the evidence falls to favor the view that these similarities arise from similarities in the people that infants come to construe, rather than a module prespecifying that people are to be interpreted in these ways.

Why might there be a relationship between spontaneous trait inference and the use of mental constructs (beliefs and desires) in explaining behaviors? We see three possibilities. First is the possibility that idiocentric thought feeds both unconscious trait inference and conscious belief-desire reasoning. Uleman (1999) considers spontaneous inferences as part of the underground stream of unconscious thought, and intentional ones as part of an above-ground aqueduct system. Taking this analogy further, both could be fed from the same source. Cultural influences could lead both to the unconscious habits of mind that are

evident in spontaneous trait inference, and to the more considered, but again habitual, schematic tendency to consciously think of others' behaviors as stemming from internal sources (Lillard et al., 2001; Miller, 1984). Traits occur spontaneously to people who are idiocentric, who think we act freely based on our internal proclivities rather than based on outside sources. And those same people, asked to reason consciously about why people do things, arrive at belief and desire explanations for those actions.

A second possibility is that trait reasons are initially elicited in the behavior explanation task, but are withheld because they are deemed less acceptable. This reasoning follows Gilbert's (1989) theory that social inference processes involve a characterization followed by a correction. By this reasoning, participants in the open-ended task have (usually) corrected an initial trait inference, and often have offered a belief-desire reason in its place. This correction might occur because of sensitivity to the fundamental attribution error among the Psychology 101 students who participated in this experiment.

A third possibility is that both belief-desire reasoning and trait reasoning are spontaneous, unconscious processes, and that belief-desire reasoning is actually primary. The spontaneous inference task only taps trait inferences, so the automaticity of belief-desire reasoning is simply untapped. But because it is primary, belief-desire reasoning is what emerges most readily when people explain behaviors. These spontaneous, unconscious processes are habits of mind, with individualistic or collectivist cultures having provided a "continual

priming effect" (Lillard, 1998; Shweder & Bourne, 1984) leading to people's use of such constructs.

The factors that lead to these habits of mind are primarily cultural, but are also rooted in our animal biology. Infants begin life with some degree of self-awareness and understanding others' core similarity to oneself, as evidenced by infants' imitation of others bodily movements in the first hours and weeks of life (Lillard, 1999; Meltzoff & Moore, 1995). Later, as infants become aware of their own intentions, they begin to read intentions into the actions of others (Woodward, 1998). Rizzolati, Gallese and their colleagues (Gallese, 2000; Rizzolatti & Arbib, 1998) have found in monkeys what may be the neural analog of this process. When monkeys observe experimenters engaging in particular movements, some of the same neurons fire in the monkeys as fire when the monkeys themselves make those movements. Humans have been shown to have similar processes. (For a superb recent review of this topic, see Dijksterhuis & Bargh, 2001.) Frith and Frith (2001) as well note that the medial frontal areas of the brain that are involved in self-monitoring also are activated in interpreting others behaviors. Although some assume that the involvement of specific brain regions suggests genetic prespecification, modules can arise through physical characteristics of different brain regions that come to take over types of inputs because they are more efficient at processing those inputs (Elman et al., 1996).

From this process of matching oneself with others, and perceiving some of one's own mental states, young children begin to understand mental processes in

others. This is the root of all cultures having some theory of mind: all people really do operate in part from mental processes, and there is some awareness of these processes from early in ontogeny. However, cultures differ in how much those processes are emphasized and in how much they are acted upon (Lillard, 1998). Middle-and upper-middle class educated Americans at the turn of the millenium seem to be at the extreme of thinking about inner life: witness large self-help sections in bookstores, and how parents urge even very young children to make their own choices, based on their own desires. In contrast, in most other cultures (often termed "collectivist") emphasis is on the society and fitting into its mores at the expense of personal desires and other internal features of the individual. In such cultures, minds are not so emphasized, although they are still understood. The habits of mind that give rise to STIs and internalistic explanations for behaviors are nurtured in individualistic cultures, and are more akin to automatic than innate-modular processes.

References

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, 15(1), 1-16.

Anderson, J. R. (1990). Cognitive psychology and its implications.

Astington, J. W., Harris, P. L., & Olson, D. R. (Eds.). (1988). Developing theories of mind. New York: Cambridge University Press.

Avis, J., & Harris, P. L. (1991). Belief-desire reasoning among Baka children: Evidence for a universal conception of mind. Child Development, 62, 460-467.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. Child Development, 62, 875-890.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. J. S. T. K. Wyer (Ed.), Handbook of social cognition (Vol. 1, pp. 1-40). Hillsdale, NJ, USA: Lawrence Erlbaum Associates, Inc.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & J. C. M. (Eds.), Handbook of research methods in social and personality psychology (pp. 253-285). New York: Cambridge University Press.

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. London: MIT Press.

Baron-Cohen, S. (2000). Theory of mind and autism: A 15-year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), Understanding other minds: Perspectives from developmental cognitive neuroscience (2nd ed., pp. 3-21). Oxford: Oxford U. Press.

Bartsch, K., & Wellman, H. M. (1995). Children talk about the mind. Oxford: Oxford University Press.

Briggs, J. (1970). Never in anger: A portrait of an Eskimo family. Cambridge, MA: Harvard University Press.

Bruner, J. (1990). Acts of meaning. Cambridge, MA: Harvard University Press.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. Infant Behavior and Development, 21, 315-330.

Choi, I., & Nisbett, R. E. (1998). Situational salience and cultural differences in the correspondence bias and actor-observer bias. Personality & Social Psychology Bulletin, 24(9), 949-960.

Descartes, R. (1641/1993). Meditations on first philosophy (Cress, D.A., Trans.). (3rd ed.). Indianapolis: Hackett.

Dijksterhuis, A. P., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), Advances in experimental social psychology, Vol. 33 (pp. 1-40). San Diego, CA, US: Academic Press, Inc.

Duff, K. J., & Newman, L. S. (1997). Individual differences in the spontaneous construal of behavior: Idiocentrism and the automatization of the trait inference process. Social Cognition, 15, 217-241.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., & et al. (1996). Rethinking innateness: A connectionist perspective on development. Cambridge, MA, USA: Mit Press.

Evans-Pritchard, E. E. (1976). Witchcraft, oracles, and magic among the Azande. Oxford: Clarendon.

Fiske, A., Kitayama, S., Markus, H., & Nisbett, R. (1998). The cultural matrix of social psychology. In D. T. Gilbert & F. S. T. (Eds.), The handbook of social psychology (Vol. 2, pp. 915-981). Boston, MA, USA: Mcgraw-Hill.

Flavell, J. H. (1992). Perspectives on perspective taking. In H. Beilin & P. Pufall (Eds.), Piaget's theory: Prospects and possibilities . Hillsdale, NJ: Erlbaum.

Flavell, J. H., & Miller, P. H. (1998). Social cognition. In D. Kuhn & R. S. Siegler (Eds.), Handbook of child psychology: Vol. 2. Cognition, perception, and language development. (5th Ed.) (pp. 851-898). New York: Wiley.

Fodor, J. (1992). A theory of the child's theory of mind. Cognition, 44, 283-296.

Fodor, J. A. (1983). The modularity of mind. Cambridge, MA: Bradford Books/MIT Press.

Frith, C. D., & Frith, U. (1999). Interacting minds--a biological basis. Science, 286(5445), 1692-1695.

Frith, U., & Frith, C. (2001). The biological basis of social interaction. Current Directions in Psychological Science, 10, 151-155.

Gall, F. G. (1835). Works: On the functions of the brain and each of its parts. (Vol. 1-6). Boston: March, Capon, & Lyon.

Gallese, V. (2000). The acting subject: Toward the neural basis of social cognition. In T. Metzinger (Ed.), Neural correlates of consciousness: Empirical and conceptual questions (pp. 325-333). Cambridge, MA, US: MIT Press.

Geertz, C. (1984). "From the native's point of view": On the nature of anthropological understanding. In R. A. Shweder & R. A. LeVine (Eds.), Culture theory: Essays on mind, self, and emotion (pp. 123-136). Cambridge, England: Cambridge Unversity Press.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. Cognition, 56, 165-193.

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), Unintended thought: Limis of awareness, intention, and control (pp. 189-211). New York: Guilford.

Gurin, P., Gurin, G., & Morrison, B. M. (1978). Personal and ideological aspects of internal and external control. Social Psychology, 41, 275-296.

Hamilton, V. L., & Sanders, J. (1992). Everyday justice. New Haven: Yale University Press.

Harre, R. (1981). Psychological variety. In P. Heelas & A. Lock (Eds.), Indigenous Psychologies (pp. 79-104). New York: Academic Press.

Harris, P. L. (1996). Desires, beliefs, and language. In P. Carruthers & P. K. Smith (Eds.), Theories of theories of mind (pp. 200-220). Cambridge, England: Cambridge University Press.

Harris, P. L. (2000). The work of the imagination. Oxford: Blackwell.

Heider, F. (1958). The psychology of interpersonal relations. New York: Wiley.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. American Journal of Psychology, 57, 243-259.

Hollos, M. (1987). Learning in rural communities: Cognitive development in Hungary and Norway. In G. D. Spindler (Ed.), Education and culural process: Anthropolgical approaches (2 ed., pp. 401-429). Prospect Heights, IL: Waveland Press.

Howell, S. (1981). Rules not words. In P. Heelas & A. Lock (Eds.), Indigenous Psychologies (pp. 133-144). New York: Academic Press.

Howell, S. (1984). Society and cosmos. Oxford: Oxford University Press.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. Advances in Experimental Social Psychology, 2, 219-266.

Karmiloff-Smith, A. (1992). Beyond modularity. London: Bradford/M.I.T.

Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. Psychological Review, 82, 387-406.

Krull, D. S., Loy, M. H.-M., Lin, J., Wang, C.-F., Chen, S., & Zhao, X. (1999). The fundamental fundamental attribution error: Correspondence bias in individualist and collectivist cultures. Personality & Social Psychology Bulletin, 25, 1208-1219.

Lee, F., Hallahan, M., & Herzog, T. (1996). Explaining real-life events: How culture and domain shape attributions. Personality & Social Psychology Bulletin, 22, 732-741.

Lefcourt, H. M. (1991). Locus of control. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), Measures of personality and social psychological attitudes. Vol. 1 (Vol. 1, pp. 413-499). San Diego: Academic Press.

Leslie, A. M. (1992). Pretense, autism, and the theory-of-mind module. Current Directions in Psychological Science, 1, 18-21.

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. A. Hirschfield & S. A. Gelman (Eds.), Mapping the mind: Domain specificity in cogntion and culture (pp. 119-148). Cambridge: Cambridge University Press.

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In H. M. Lefcourt (Ed.), Research with the locus of control construct. Vol 1: Assessment methods . New York: Academic Press.

LeVine, R. A. (1984). Properties of culture: An ethnographic view. In R. A. Shweder & R. A. LeVine (Eds.), Culture theory: Essays on mind, self, and emotion (pp. 67-87). Cambridge, England: Cambridge Unversity Press.

LeVine, S. (1979). Mothers and wives. Chicago: University of Chicago Press.

Lillard, A. S. (1997). Other folks' theories of mind and behavior. Psychological Science, 8, 268-274.

Lillard, A. S. (1998). Ethnopsychologies: Cultural variations in theory of mind. Psychological Bulletin, 123, 3-33.

Lillard, A. S. (1999). Developing a cultural theory of mind: The CIAO approach. Current Directions in Psychological Science, 8, 57-61.

Lillard, A. S. (in press). Pretend play and cognitive development. In U. Goswami (Ed.), Handbook of Cognitive Development . London: Blackwell.

Lillard, A. S., Skibbe, L., Zeljo, A., & Harlan, D. (2001). Developing Explanations for Behavior in Different Communities and Cultures . Unpublished manuscript, University of Virginia.

Livesley, W. J., & Bromley, D. B. (1973). Person perception in childhood and adolescence. London: Wiley.

Lutz, C. (1985). Cultural patterns and individual differences in the child's emotional meaning system. In M. Lewis & C. Saarni (Eds.), The socialization of emotions . Plenum Press: New York.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98, 224-253.

Mayer, J. (1982). Body, psyche, and society: Conceptions of illness in Ommura, Eastern Highlands, Paua New Guinea. Oceania, 52, 240-259.

Meltzoff, A. N., & Moore, M. K. (1995). A theory of the role of imitation in the emergence of self. In P. Rochat (Ed.), The self in infancy: Theory and research. Advances in psychology, 112 (pp. 73-93). Amsterdam, Netherlands: North-Holland/Elsevier Science Publishers.

Miller, F. D., Smith, E. R., & Uleman, J. (1981). Measurement and interpretation of situational and dispositional attributions. Journal of Experimental Social Psychology, 17, 80-95.

Miller, J. G. (1984). Culture and the development of everyday social explanation. Journal of Personality and Social Psychology, 46, 961-978.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. Journal of Personality and Social Psychology, 67, 949-971.

Newman, L. S. (1991). Why are traits inferred spontaneously? A developmental approach. Social Cognition, 9, 221-253.

Newman, L. S. (1993). How individualists interpret behavior: Idiocentrism and spontaneous trait inference. Social Cognition, 11, 243-269.

Ochs, E., & Schieffelin, B. (1984). Language acquisition and socialization. In R. Shweder & R. LeVine (Eds.), Culture theory: Mind, self, and emotion (pp. 276-322). Cambridge, England: Cambridge University Press.

Paul, R. A. (1995). Act and intention in Sherpa culture and society. In L. Rosen (Ed.), Other intentions: Cultural contexts and the attribution of inner states (pp. 15-45). Santa Fe NM: School of American Research Press.

Poole, F. J. P. (1985). Coming into being: Cultural images of infants in Bimin-Kuskusmin folk psychology. In G. M. White & J. Kirkpatrick (Eds.), Person, self, and experience (pp. 183-244). Berkeley: University of California Press.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1, 515-526.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. Trends in Neurosciences, 21(5), 188-194.

Ross, L. (1978). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), Cognitive theories in social psychology (pp. 337-384). New York: Academic Press.

Ross, L., & Nisbett, R. E. (1991). The person and the situation: Perspectives of social psychology. New York: McGraw-Hill.

Sacks, O. (1995). An Anthropologist on Mars. New York: Knopf.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development, and 'theory of mind'. Mind and Language, 14, 131 - 153.

Schulkin, J. (2000). Roots of social sensibility and neural function. Cambridge, MA, US: The MIT Press.

Shantz, C. U. (1983). Social cognition. In J. H. Flavell & E. M. Markman (Eds.), Handbook of child psychology: Vol. 3. Cognitive development (Vol. III, pp. 495-555). New York: Wiley.

Shweder, R. A., & Bourne, L. (1984). Does the concept of the person vary cross-culturally? In R. A. Shweder & R. A. LeVine (Eds.), Culture theory: Essays on mind, self, and emotion (pp. 158-199). Cambridge, England: Cambridge Unversity Press.

Solomon, S. (1978). Measuring dispositional and situational attributions. Personality & Social Psychology Bulletin, 4, 589-594.

Straus, A. (1977). Northern Cheyenne ethnopsychology. Ethos, 5, 326-352.

Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin- and Cantonese-speaking children. Developmental Psychology, 36(1), 25-43.

Triandis, H. C. (1994). Culture and social behavior. New York: McGraw-Hill.

Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & et al. (1988). Individualism and collectivism: Cross-cultural perspectives on selfngroup relationships. Journal of Personality & Social Psychology, 54(2) Feb 1988), US, http://www.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80(5), 359-380.

Uleman, J. S. (1999). Spontaneous versus intentional inferences in impression formation. In S. T. Y. Chaiken (Ed.), Dual-process theories in social psychology (pp. 141-160). New York, NY, US: The Guilford Press.

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreteres: Evidence and issues from spontaneous trait inference. Advances in Social Psychology, 28, 211-279.

Weeks, M., & Lupfer, M. B. (2000). Religious attributions and proximity of influence: An investigation of direct interventions and distal explanations. Journal for the Scientific Study of Religion, 39(3), 348-362.

Wellman, H. M. (1990). The child's theory of mind. Cambridge, MA: Bradford Books/MIT Press.

Wellman, H. M. (in press). Understanding the psychological world: Developing a theory of mind. In U. Goswami (Ed.), Handbook of Cognitive Development . Oxford: Blackwell.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. Child Development, 72, 655-684.

Witt, L. A. (1989). Urban-nonurban differences in social cognition: Locus of control and perceptions of a just world. The Journal of Social Psychology, 129, 715-717.

Woodward, A. L. (1998). Infant's selectively encode the goal object of an actor's reach. Cognition, 69, 1-34.

Yuill, N. (1992). Children's conception of personality traits. Human Development, 35, 265-279.

Zarate, M. A., Uleman, J. S., & Voils, C. (2001). Effects of culture and processing goals on the activation and binding of trait concepts. Social Cognition, 19, 295-323.

Zelli, A., Cervone, D., & Huesmann, L. R. (1996). Behavioral experience and social inference: Individual differences in aggressive experience and spontaneous versus deliberate trait inference. Social Cognition, 14(2) Sum 1996), 165-190.

Zimbelman, K. (1987). Locus of control and achievement orientation in rural and metropolitan youth. Journal of Rural Community Psychology, 8, 50-55.