

Does the autistic child have a theory of society?

Lawrence Hirschfeld¹, Elizabeth Bartmess², Sarah White³, & Uta Frith³

Social stereotypes provide a cognitively “inexpensive” if often inaccurate way to predict the behavior of others. We found that in spite of autistic children’s profound impairment in the ability to predict behavior on the basis of an individual’s mental state, they were just as likely as young normal children to use stereotypes to predict outcomes of novel situations. This finding is surprising only if one assumes that the ability to explain the behavior of others relies on a single mechanism. Our findings suggest that there are two distinct cognitive capacities, one that makes sense of others’ behavior in terms of psychological states (Theory of Mind) and another in terms of social group membership (Naive Sociology). Theory of Mind but not Naive Sociology is impaired in autism. This is a hitherto unsuspected islet of social ability in autism.

¹ New School for Social Research, Departments of Psychology & Anthropology, 65 Fifth Ave., New York, NY 10023, USA

² University of Michigan, Department of Psychology, 530 Church Street, Ann Arbor, MI 48109-1043, USA

³ UCL Institute of Cognitive Neuroscience, 17 Queen Square, London WC1 3AR, UK

Word count: 2,682

The ability to attribute other people's actions to unseen desires, beliefs and intentions. is often referred to as Theory of Mind (ToM) (Premack & Woodruff, 1978; Wimmer & Perner, 1983). Theory of mind, or mentalizing, is assumed to rely on a neural circuit that has been revealed in neuroimaging studies (for reviews see Frith & Frith, 2003; Gallagher & Frith, 2003; Saxe, Carey, & Kanwisher, 2004). Mentalizing ability is impaired in individuals with autism (Baron-Cohen, Leslie, & Frith, 1985) and in neurological patients where components of the underlying circuit are damaged (Siegal M & Varley, 2002).

As fundamental as ToM is to understanding the social world, and as satisfying as research on it has been, ToM meets the challenge of understanding others only in part. Group-based reasoning – reasoning through which an individual's membership in social groups is the basis for interpreting action – also makes a powerful contribution to social understanding. Despite its importance, we know surprisingly little about the cognitive specificity of “Theory of Society” (or, as we prefer, “Naive Sociology”). Given what we know about other commonsense ways of interpreting the world, one vital question is whether group-based reasoning, such as the use of social stereotypes, is based on the same cognitive capacity as that using the attributions of mental states to others. Another possibility is that group-based reasoning is based on lower-level, domain-general processes that underlie the categorization of like things together.

In the literature on group-based reasoning both these views enjoy considerable support. There is broad agreement that stereotypes are best understood from the perspective of the “cognitive miser” model that focuses attention on those processes that serve to reduce cognitive

demand by simplifying situations. In this vein, several domain-general models for stereotypes have been proposed. All share the premise that stereotypes and other forms of social perception recruit the same processes that reduce cognitive demand in nonsocial object perception (e.g., prototype and exemplar effects, associative network and schema formation, base rate biases (Hilton & von Hippel, 1996)). Another hypothesis, although not addressing whether social reasoning is grounded in general category processes, proposes that stereotypes and other group-based reasoning rely on the same cognitive mechanisms and processes as person perception, viz., the attribution of traits and dispositions (Crawford, Sherman, & Hamilton, 2002; Hamilton & Sherman, 1996)). Wellman (1990) integrates ToM in this linkage between group-based reasoning and person perception, arguing that traits and dispositions are simply stabilized representations of the propensities to believe and to desire. .

Against these views, we propose that group-based reasoning is based on a separate cognitive mechanism, triggered by a distinct range of inputs, with its own developmental pattern and possibility its own neurophysiological basis (Berreby, 2005; Furth, 1996; Hirschfeld, 1988; 1996; Jackendoff, 1992). Even though evidence for the latter is as yet lacking, it is clear that many crucial cognitive capacities are governed, particularly in their developmental trajectory, by special-purpose mechanisms. We propose that the capacity of group-based reasoning, which must be highly important in the evolution of social animals, would be similarly organized.

Some support for this idea comes from several lines of developmental research. Group-based reasoning appears to be as early-emerging as, if not earlier than, parallel mentalistic capacities. Infants discriminate several distinctions that are relevant to later-emerging patterns in social categorization, reliably discriminating, for example, persons in terms of gender (Miller, 1983), language spoken (Mehler et al., 1988), and relative age (Bahrick, Netto, & Hernandez-

Reif, 1998; Brooks & Lewis, 1976). Early in the second year of life, the capacity for “we intentionality” emerges, a cognitive skill that is fundamental to forming collaborative groups (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Among slightly older children, the earliest words used by the early-language-learner are kinship terms, such as *mama*, which the young child appears to understand to refer to categories of people rather than simply individuals (Hirschfeld, 1989). By 3 years of age, children display full mastery of group-based reasoning evidenced in their ready use of common gender and racial stereotypes (Aboud, 1988; Katz, 1983). These proclivities to stereotype are both robust over variation in learning environments (Aboud, 1988) and stubbornly resistant to concerted efforts to change (Aboud & Doyle, 1996)), features associated with other modular devices (Hirschfeld & Gelman, 1994).

Contrasting the hypothesis that group-based reasoning is simply an extension of ToM reasoning, or else merely a domain general categorization based on acquired social knowledge, and the alternative proposal, that group-based reasoning is part of a domain-specific cognitive device, makes a distinct set of predictions regarding the autistic children’s capacity for and willingness to use common gender and racial stereotypes. First, children with autism have core deficits in social interaction and communication. This means that the nature and scope of autistic children’s social relationships and their social knowledge base are limited, often severely so. If stereotypes are simply domain-general category effects, autistic children should have difficulty forming and manipulating broad stereotypes, displaying instead a susceptibility to narrower, and hence culturally inappropriate group-based generalizations. If, in contrast, stereotypes are an extension of Theory of Mind, ToM-impaired autistic children should have as many difficulties in using them as making mental state attributions. If, however, autistic children’s reasoning about human groupings is governed by a special purpose cognitive device contingent neither on ToM

nor domain-general category effects, we could expect them readily to use stereotypes in making sense of others' behavior. Indeed, such a finding would be powerful evidence supporting the argument that stereotyping is based on a distinct and special-purpose cognitive mechanism for social reasoning.

1. Method

1.1. Participants

Forty-five children in three groups participated in the study; an autistic group (N=21, all boys, *Mean Chronological Age* (MCA) = 8y 8m, range 5y 5m to 11y3m; *Mean Verbal Mental Age* (MVMA)= 7y 4m, range 4y 7m to 11y 10m), and two groups of normally-developing controls, 11 matched for VMA with the autistic subjects (6 boys and 5 girls; *MCA* = 6y 8m, range 5y 4m to 7y 8m; *MVMA* = 7y3m, range 5y1 1m to 8y4m) and 13 younger normally-developing 3-year-olds (6 boys and 7 girls; *M* = 3y 7m, range 3y 4m to 3y 11m). Autistic children were recruited from schools and units specializing in autism, and all had been diagnosed by DSM IV clinical criteria (American Psychiatric Association. & American Psychiatric Association. Task Force on DSM-IV., 2000). The normally-developing children attended either a primary school (older controls) or a preschool (younger controls) serving the general population.

1.2. Tasks

1.2.1 *VMA and ToM tasks*: To assess the verbal mental age of the older controls and autistic children, we used the British Picture Vocabulary Scale-II (Dunn, Dunn, Whetton, & Burley, 1997), a nationally normed verbal IQ test. To assess Theory of mind we used two well-established measures of first-order ToM: the Sally-Ann (Baron-Cohen et al., 1985) and Smarties tasks (Perner, Frith, Leslie, & Leekam, 1989). Normally-developing 6-year-olds are known to readily pass standard False Belief tasks while normally-developing 3-year-olds are known to fail them (Flavell, 1999), indicating either a gradual (Mitchell, 2003) or more radical conceptual change (Wellman, Cross, & Watson, 2001).

We subdivided the participants with autism into ToM failers and ToM passers. Autistic children were considered ToM failers if they failed one of the two test questions. On that criterion, 10 autistic children were identified as ToM failers. Of these, 100% failed the Sally-Ann task and 80% failed the Smarties task. Verbal mental ages for these groups revealed that the ToM failers had lower VMAs (mean = 6y8m) than the ToM passers (mean = 7y10m), a difference that failed to reach significance ($t(19) = 1.640, p=0.118$). (see Table 1)

1.2.2 *A Desire screening task*, similar to tasks used by Wooley & Wellman (1990) and Repacholi & Gopnik (1997) assessed whether the children understood that if someone (habitually) wants something they will (habitually) take action to fulfill that desire. Wellman & Liu (2004) found that desire is virtually the first ToM-relevant mental state that children understand, and we anticipated that even autistic children would grasp the notion by 5 years of age. Our materials consisted of three triads, each made up of a target drawing depicting a representation of desire and two test drawings, each depicting one of two choices. In one trial, desire was represented by a color drawing of a face looking at one of two choices; in another, by

a finger pointing at one choice; in the third, desire was depicted as represented by a face with a speech bubble declaring a desire. For example, in the third trial, the child was shown a target picture of a face with a speech bubble saying “I like the white one” above a brown and a white cookie. The child was told, “Sam says, ‘I **like** the white one.’” The child was then shown two comparison pictures, in one of which Sam was in profile holding a brown cookie, in the other a white cookie. The experimenter then asked “Which does he choose?” The order of trials was counterbalanced.

1.2.3 *Preschool Racial Attitude Measure (PRAM II) & Conflict task.* The PRAM II assesses children’s racial and gender stereotypes, using a forced-choice format (Williams, Best, & Boswell, 1975). On each of 24 racial and 12 gender trials, children were presented with a brief vignette describing a positive or negative quality (in the racial trials) or a common sex-role stereotype (in the gender trials). The task was introduced in the following way: “What I have here are some pictures I’d like to show you, and stories to go with each one. I want you to help me by pointing to the person in each picture that the story is about.” Children are then shown color drawings that depict either one white and one black person (in the racial trials) or a male and a female (in the gender trials). To illustrate: one negative racial trial reads, “Here are two men. One of them is a bad man. He took money out of his children’s piggy bank and never put it back. Which is the bad man?” A positive racial trial reads, “Here are two boys. One of them is a friendly boy. He has lots of friends. Which one is the friendly boy?” A gender trial reads, “Here are two people. One of these people has baked two delicious apple pies. Which person baked the pies?”

In the *Conflict* task the child was presented with a novel situation in which one prediction could be made from an individual's mental state, while a different prediction could be made from his or her social category membership. The Conflict task was introduced in the same way as the PRAM II. Each vignette presented the child with a situation in which an individual could act either in accord with his or her mental state (in this case a habitual preference) or in accord with his or her social position (i.e., in a group stereotype- or group attitude-consistent manner). Five trials pitted mental state against social expectations about race and five trials pitted mental state against common sex-role stereotypes. For example, in a sex-role vs. habitual preference trial the child was shown a picture of a man and a woman and told, "Here are two people. This is James and this is Grace. Grace doesn't like to cook for people. One of these people has baked biscuits. Which person baked biscuits?" In one of the racial attitude versus habitual preference trials, the child was shown a picture of a black woman and a white woman and told, "Here are two women. This is Georgina [indicates drawing of black woman] and this is Jasmine [indicates drawing of white woman]. Georgina likes to help people. One of these women takes care of her husband and children. Which woman takes care of her husband and children?" (NB. names were chosen from frequency lists of British children's names.)

2. Results

Table 1 shows the results of the desire task. A Kruskal-Wallis test indicated a group difference, $\chi^2(3) = 26.73, p < 0.001$. Post-hoc comparisons revealed that the younger controls had greater difficulty with the task than all other groups who did not differ from each other (all $ps > 0.13$). A 2 (task) x 4 (group) ANOVA showed that all groups were more likely to give stereotype consistent responses in the PRAM (approx 70%) than in the conflict task

($F(1,41)=155.73$, $p<0.001$), whereas 3-year-olds and autistic ToM failers produced more stereotypes than controls and autistic ToM passers in the conflict task but not in the PRAM (significant interaction: $F(3,41)=3.36$, $p=0.028$). Thus, when there was a conflict, desire was less important for the former than the latter groups, and they responded about 40% of the time in line with group stereotype. Only the 6-year olds and the autistic ToM passers showed a clear avoidance of stereotypes with less than 20% stereotype choices.

3. Discussion

These findings are striking on several counts. First, regardless of their success with standard ToM false belief tasks, the autistic children participating in our study *know* a great deal about familiar stereotypes. Moreover, they can readily recruit this knowledge in predicting and explaining others' behavior. Stereotypes are preternatural units of cultural knowledge, and vary considerably by place and historical epoch. Virtually all accounts of the acquisition of cultural knowledge are premised on the conviction that such knowledge is the product of persisting, and often consciously guided, engagement with the flow of normal social life. Autistic children whose experiences are assumed to be outside this flow and whose social engagement is limited, should, on this account, be disadvantaged in developing normal cultural competence (Rogers, Hepburn, Stackhouse, & Wehner, 2003; Tomasello et al., 2005). Yet our findings show that they have an intact capacity for group-based social reasoning. Our study also shows that relevant *conceptual* impediments, here an impairment in mentalistic reasoning, do not significantly affect children's use of common social stereotypes

Indeed, the findings we report here are surprising *only* when we assume that group-based reasoning is acquired as a result of non-specific processing of social information. They fall into

place when we acknowledge that stereotypic knowledge relies on special-purpose cognitive processes with their own developmental trajectory. These processes must be remarkably robust if they work in early childhood, when social experience is limited, and in autism, when social experience is abnormal, one hallmark of an innate cognitive competence. Interestingly, given the pattern of reasoning evidenced by normally-developing 3-year-olds, our results can be taken as tentative evidence that group-based reasoning is highly elaborated even before crucial aspects of mental state attributions emerge.

Our findings thus support the hypothesis that the capacity for Naïve Sociology is unimpaired in autism, and distinct from the capacity for Naïve Psychology or Theory of Mind, which is impaired. When we combine these findings with what we know of the sophisticated skills in social reasoning among many primate species, it seems plausible that group-based explanations of behavior are a fundamental dimension of the human mind. We would like to suggest that Naïve Sociology, just as Naïve Psychology, relies on a neurocognitive mechanism whose neurophysiological basis may yet be revealed.

References

- About, F. E. (1988). Children and prejudice. York, NY: B. Blackwell.
- About, F. E., & Doyle, A.-B. (1996). Parental and peer influences on children's racial attitudes. International Journal of Intercultural Relations: Special Issue: Prejudice, discrimination and conflict., 20(3-4), 371-383.
- American Psychiatric Association., & American Psychiatric Association. Task Force on DSM-IV. (2000). Diagnostic and statistical manual of mental disorders : DSM-IV-TR. (4th , text revision. ed.). Washington, DC: American Psychiatric Association.
- Bahrnick, L. E., Netto, D., & Hernandez-Reif, M. (1998). Intermodal perception of adult and child faces and voices by infants. Child Development, 69(5), 1263-1275.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? Cognition, 21(1), 37-46.
- Berreby, D. (2005). Us and them: Understanding your tribal mind. New York: Little, Brown.
- Brooks, J., & Lewis, M. (1976). Infants' responses to strangers: Midget, adult, and child. Child Development, 47(2), 323-332.
- Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entatativity, stereotype formation, and the interchangeability of group members. Journal of Personality and Social Psychology, 83(5), 1076-1094.
- Dunn, L., Dunn, L. M., Whetton, C., & Burley, J. (1997). British Picture Vocabulary Scale. (2nd (BPVS-II) ed.). Windsor, Berks: NFER-Nelson.
- Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. Annual Review of Psychology, 50, 21-45.

Frith, U., & Frith, C. (2003). Development and neurphysiology of mentalizing. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, 358(March 29), 459-473.

Furth, H. (1996). Desire for society: Children's knowledge as social imagination. New York: Plenum Press.

Gallagher, H. L., & Frith, C. D. (2003). Functional imagine of 'theory of mind'. Trends in Cognitive Sciences, 7(2), 77-83.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. Psychological Review, 103(2), US American Psychological Assn.; 1996, 336-355.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. Annual Review of Psychology, 47, 237-271.

Hirschfeld, L. A. (1988). On acquiring social categories: cognitive development and anthropological wisdom. Man, 23, 611-38.

Hirschfeld, L. A. (1989). Rethinking the acquisition of kinship terms. International Journal of Behavioral Development, 12, 541-568.

Hirschfeld, L. A. (1996). Race in the making : cognition, culture, and the child's construction of human kinds. Cambridge, MA: MIT Press.

Hirschfeld, L. A., & Gelman, S. A. (1994). Mapping the mind: Domain specificity in cognition and culture. New York: Cambridge University Press.

Jackendoff, R. (1992). Language of the mind: Essays on mental representation. Cambridge, MA: MIT Press.

Katz, P. (1983). Developmental foundations of gender and racial attitudes. In R. Leahy (Ed.), The child's construction of social inequality . New York: Academic Press.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. Cognition, *29*, 143-178.

Miller, C. L. (1983). Developmental changes in male/female voice classification by infants. Infant Behavior & Development, *6*(3), 313-330.

Mitchell, P. (2003). Acquiring a theory of mind. In A. Slater & G. Bremner (Eds.), An introduction to developmental psychology (pp. 237-257). Malden, MA: Blackwell Publishing.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. Child Development, *60*(3), 689-700.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, *1*(4), 515-526.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. Developmental Psychology, *33*(1), 12-21.

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. Journal of Child Psychology and Psychiatry, *44*(5), 763-781.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. Annual Review of Psychology, *55*, 87-124.

Siegal M, & Varley, R. (2002). Neural systems involved in "theory of mind". Nature Review Neuroscience, *3*, 463-471.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. Behavioral & Brain Sciences, *28*, 675-735.

Wellman, H. M. (1990). The child's theory of mind. Cambridge, Mass.: MIT Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. Child Development: Special Issue., *72*(3), 655-684.

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. Child Development, *75*(2), 523-541.

Williams, J. E., Best, D. L., & Boswell, D. A. (1975). The measurement of children's racial attitudes in the early school years. Child Development, *46*(2), 494-500.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, *13*(1), 103-128.

Woolley, J., & Wellman, H. M. (1990). Young children's understanding of realities, nonrealities, and appearances. Child Development, *61*(4), 946-961.

Acknowledgments

This work was supported by MRC programme grant 65013 to UF and a research grant to Lawrence Hirschfeld and Elizabeth Bartmess by the Culture and Cognition Program, University of Michigan USA. The study obtained ethical approval from the Joint UCL/UCLH committees on the Ethics of Human Research.

	3-year-olds N=13	6-year-olds N=11	Autistic ToM failers N=10	Autistic ToM passers N=11
Age (years;months)	3;7 (0;2)	6;8 (0;9)	7;11 (2;2)	9;5 (1;4)
VMA (years;months)	N/A	7;3 (0;9)	6;8 (1;5)	7;10 (1;10)
Desire task (out of 3)	1.77	2.90	2.60	3.00

Table 1

Participant characteristics and performance on desire screening test (means, SDs)

VMA = Verbal Mental Age

Figure 1
Percentage of stereotype consistent responses in the two experimental tasks

