Scaling of Theory-of-Mind Understandings in Chinese Children

Henry M. Wellman, Fuxi Fang, David Liu, Liqi Zhu and Guoxiong Liu
Psychological Science 2006 17: 1075
DOI: 10.1111/j.1467-9280.2006.01830.x

Research Article

Scaling of Theory-of-Mind Understandings in Chinese Children

Henry M. Wellman,[1] Fuxi Fang,[2] David Liu,[3] Liqi Zhu,[2] and Guoxiong Liu[2]

[1]University of Michigan; [2]Institute of Psychology, Chinese Academy of Sciences, Beijing, China; and [3]University of Washington

ABSTRACT—*Prior research demonstrates that understanding of theory of mind develops at different paces in children raised in different cultures. Are these differences simply differences in timing, or do they represent different patterns of cultural learning? That is, to what extent are sequences of theory-of-mind understanding universal, and to what extent are they culture-specific? We addressed these questions by using a theory-of-mind scale to examine performance of 140 Chinese children living in Beijing and to compare their performance with that of 135 English-speaking children living in the United States and Australia. Results reveal a common sequence of understanding, as well as sociocultural differences in children's developing theories of mind.*

People interpret each other's actions in terms of underlying mental states: the actor's beliefs, desires, emotions. What is the character and origin of this foundational theory of mind? This question has been addressed in investigations of the mental understandings of typical and atypical children. Most such investigations have focused on some single mental-state concept, especially false belief (Peterson & Siegal, 2000; Wellman, Cross, & Watson, 2001). However, achieving a theory of mind includes understanding multiple concepts potentially acquired in developmental sequences (see Harris, Rosnay, & Pons, 2005; Wellman & Liu, 2004).

How might one research such sequences? Adopting one approach, we (Wellman & Liu, 2004) devised a developmental scale composed of tasks assessing understanding of (a) diverse desires (people can have different desires for the same thing), (b) diverse beliefs (people can have different beliefs about the same situation), (c) knowledge-ignorance (something can be true, but someone might not know that), (d) false belief (something can be true, but someone might believe something different), and (e) hidden emotion (someone can feel one way but display a different emotion). The tasks were similar in procedures, language, and format, yet U.S. preschoolers evidenced a clear order of difficulty, with understanding diverse desires being easiest and understanding hidden emotion being hardest, a result confirmed by Guttman and Rasch scale analyses. A subsequent study in Australia (Peterson, Wellman, & Liu, 2005) examined the same tasks with typical preschoolers, deaf children, and children with autism. The same developmental order emerged for the first two groups, but for the children with autism, the latter part of the sequence differed.

This prior research provides a method for examining the progression of theory-of-mind understandings, and reveals a robust sequence of such understandings. However, what accounts for the consistency of sequence demonstrated so far (and for the difference in autism)? A consistent sequence could result from innately programmed modular maturations, from domain-general cognitive gains (say, increases in executive functioning as reflected across an increasingly demanding sequence of tasks), or from processes of conceptual learning in which initial conceptions lead to later concepts as shaped by relevant information and experience. These alternatives encompass contrasting accounts within current debates on theory-of-mind development.

Examining progressions of understanding for children living in different cultures and growing up in different sociolinguistic environments can help researchers to evaluate such alternatives. If sequences were universal, that would support accounts based on innate modular maturation (children in all cultures

Address correspondence to Henry Wellman, Center for Human Growth and Development, 300 N. Ingalls, 10th Floor, Ann Arbor, MI 48109-0406, e-mail: hmw@umich.edu.

mature similarly) or on domain-general cognitive gains (sequences reflect general cognitive demands of the tasks, rather than domain-specific conceptual content). In contrast, accounts based on conceptual learning from culturally variable inputs predict that sequences should differ if cultural inputs are different enough. Do children growing up in quite different sociocultural circumstances evidence the same or different sequences in development of theory of mind?

To address this question, we assessed sequences of theory-of-mind understanding in Chinese children, using the scale methods of our previous study (Wellman & Liu, 2004). Chinese children grow up in non-Western cultures, acquiring non-Indo-European languages and participating in childhood experiences different from those of their Western, English-speaking peers. For example, Chinese languages differ from English in ways that could influence early adult-child conversations, which are known to influence theory-of-mind understandings in English-speaking children (Dunn, 1995; Ruffman, Slade, & Crowe, 2002). Specifically, Chinese languages use a different collection of mental verbs than English does (Tardif & Wellman, 2000). Moreover, the English terms *think* and *believe* refer equally, and ambiguously, to beliefs that may be true or false. But Cantonese and Mandarin have specific verbs for "think falsely," and these are used by preschoolers (Lee, Olson, & Torrance, 1999; Tardif, Wellman, & Cheung, 2004). Diverging language or cultural experiences might contribute to important differences in young children's developing theories of mind, or certain developments might unfold universally for children despite such differences.

We propose that children's understandings reflect a mix of universally emerging insights and culturally specific ones. In support of possible universal progressions, several theorists have argued that understanding cognitive, or representational, mental states (such as beliefs and knowledge) is conceptually harder than understanding noncognitive states such as desires (see Flavell & Miller, 1998). Indeed, children in China talk about desires before beliefs (Tardif & Wellman, 2000), just as do English-speaking children (Bartsch & Wellman, 1995) and deaf children learning sign (Anderson & Reilly, 2002). This suggests that understanding desires before beliefs may represent a fundamental and universal progression in an unfolding theory of mind. If so, in terms of the scale tasks used in the present study, Chinese children would understand diverse desires before diverse beliefs and false beliefs (see Table 1 for a brief description of the tasks). Assessing Chinese children with such tasks provides an important test of such potentially universal progressions.

There are also reasons to predict intriguing differences across cultures, assuming theory-of-mind understandings are the products of social and conversational experiences that vary from one community to another. In particular, Western and Chinese adults seem to manifest very different everyday epistemologies.

TABLE 1

Summary of the Tasks and Chinese Children's Performance

Task	Proportion correct	Brief description
Diverse-desires	.89	The child judges that two persons (the child and someone else) have different desires about the same object: Given two possible snacks (ice cream, an egg), the child states his or her preference, but then must predict the choice of the other person (who has the opposite preference).
Knowledge-ignorance	.79	The child judges another person's ignorance about the contents of a container when the child knows what is in the container: The child sees a toy dog in a nondescript drawer; the drawer is closed, and the child judges (yes/no) if the other person (who has never seen inside) knows what is in the drawer (and also judges if that person saw inside).
Diverse-beliefs	.71	The child judges that two persons (the child and someone else) have different beliefs about the same object, when the child does not know which belief is true or false: The child states his or her belief that an object is under the bed and hears the other person's belief that it is in the cupboard; the child never sees where the item is, but must predict whether the other person will search under the bed or in the cupboard.
Contents false-belief	.54	The child judges another person's false belief about what is in a distinctive container when the child knows what is in the container: The child sees a familiar potato-chip tube, discovers it has pencils inside, and then must judge the belief of someone else who has never seen inside (and judges if that person saw inside).
Explicit false-belief	.49	The child judges how someone will search, given that person's mistaken belief: The child sees a book in a backpack, then hears that the character thinks the book is in the drawer, and judges whether the character will search in the backpack or in the drawer (and judges where the book really is).
Hidden-emotion	.37	The child judges that a person can feel one thing but display a different emotion: The character wants his uncle to bring him a toy gun, but the uncle brings a book; the child judges how the character will feel (sad) and what he will show on his face (happy).

Note. Detailed descriptions of all tasks are available from the authors.

Nisbett (2003) and J. Li (2002) have argued that Western epistemology is focused more on truth and belief, whereas Chinese epistemology is focused more on pragmatic knowledge acquisition. Developmentally, both English- and Chinese-speaking preschoolers acquire a word for "know" as one of their earliest mental-state verbs, before a word for "think" (Tardif & Wellman, 2000). But there also appear to be cultural differences such that Chinese children may receive more emphasis on "knowing" relative to "thinking." In conversation with young children, Chinese parents comment predominantly on "knowing" (Tardif & Wellman 2000), whereas U.S. parents comment more on "thinking" (Bartsch & Wellman, 1995). In Chinese preschools and homes, there is great emphasis on acquiring practical knowledge, such as how to fold one's blanket properly after a nap, tie one's shoes, write Chinese characters, and recite both songs and poems precisely (H. Li & Rao, 2000; Tobin, Wu, & Davidson, 1989). English-speaking preschoolers must also master new knowledge, but the Chinese emphasis on knowledge acquisition at an early age is remarkable (Kessen, 1975; Tobin et al., 1989).

Moreover, various authors have described an Asian focus on persons as sharing group commonalities and interdependence and a contrasting Western focus on persons as distinctively individual and independent (e.g., Nisbett, 2003). Logically, these differences include differing emphases on common beliefs and perspectives versus diversity of individual beliefs and perspectives. Of course, children everywhere encounter persons with different mental states, but the American emphasis on individual beliefs is remarkable (Markus & Kitayama, 1991).

One way in which such differences might manifest themselves on the scale tasks used here would be that the sequence of understanding belief and understanding knowledge could differ. For Western, English-speaking children, an early understanding of beliefs (revealed in understanding of diverse beliefs) precedes understanding of knowledge-ignorance, which precedes understanding of false belief (Wellman & Liu, 2004); for Chinese children (hypothetically), an early understanding of knowledge-ignorance could precede any understanding of belief. Regardless, similar or different sequences of understandings would be informative.

A focus on scalable sequences can also help circumvent a problem endemic to cross-cultural comparisons, namely, how to validly compare children across different countries and communities. A common approach to such comparisons is to examine task performance of two same-age groups. Yet children from the United States and China, even if carefully sampled for comparable ages and socioeconomic status, differ widely (e.g., in the languages they acquire, milestones of language acquisition, family experiences, the nature of schools and preschools they attend, timing of entry into schools). Mean group differences on some task or achievement may always be due to those uncontrollable confounding factors (rather than focal conceptions or skills). Fortunately, when sequences, not ages of attainment or some absolute mean score, are compared, exact (impossible-to-achieve) matching of ages, of socioeconomic status, and so on becomes less relevant. Indeed, if a scale is rigorous and consistent, then testing a range of ages, abilities, and backgrounds contributes to a demonstration that children nonetheless exhibit a consistent sequence of acquisition. Examining trajectories of development seems especially informative for comparing contrasting theoretical accounts, and data regarding sequences provide less problematic comparisons across cultural groups as well.

METHOD

Participants

Ninety-two Chinese preschoolers from two preschools in Beijing, China, were the primary participants. The preschools largely served the children of university staff and faculty in urban Beijing. The participants included thirty 5-year-olds (M = 5.5, range: 5.0–6.1), twenty-nine 4-year-olds (M = 4.6, range: 4.0–5.0), and thirty-three 3-year-olds (M = 3.6, range: 2.9–4.0); 49 were male and 43 female.

Materials

Table 1 outlines the six tasks, parallel Mandarin versions of the tasks used in our previous study (Wellman & Liu, 2004). To construct equivalent but culturally familiar and linguistically appropriate Mandarin tasks, we had the English tasks and instructions adapted by bilingual speakers who work with children in Beijing and then back-translated. Task protocols were then reviewed and improved until experts in China and the United States, including an English-Mandarin bilingual as well as a Mandarin-English bilingual, could agree on the equivalence of the tasks and translations.

Small toy figures with Chinese visages and dark hair served as the target protagonists for the tasks. Although formats and essential wordings were not changed from the English versions, task materials were modified in several places so that they would be familiar and appropriate: The diverse-desires task compared preferences for ice cream versus a hard-boiled egg (rather than cookies vs. carrots), the contents false-belief task used a potato-chip tube familiar in Beijing (rather than a Band-Aid box), and the hidden-emotion scenario was about a boy whose uncle was returning from overseas with an unwanted gift (rather than a boy hiding negative emotion from peers). For Chinese children, false-belief tasks can be presented with either think-falsely verbs or neutral verbs (for the target question "What does X think?"); think-falsely verbs enhance children's performance (Lee et al., 1999). One of our hypotheses focused on Chinese children's possibly later understanding of belief, relative to knowledge, so our contents false-belief task used a think-falsely verb, *yi3wei2* (to provide Chinese children every opportunity to

TABLE 2

Guttman Scalogram for the Five Core Tasks

	Pattern of success (+) and failure (−)					
Task	1	2	3	4	5	6
Diverse-desires	−	+	+	+	+	+
Knowledge-ignorance	−	−	+	+	+	+
Diverse-beliefs	−	−	−	+	+	+
Contents false-belief	−	−	−	−	+	+
Hidden-emotion	−	−	−	−	−	+
Number of children	0	2	4	14	22	21

Note. Twenty-nine children exhibited patterns not shown in the table.

demonstrate understanding of beliefs). The explicit false-belief task asked where the protagonist would search for an item, avoiding use of any "think" verb.

The six tasks had roughly two formats, although all used toy figurines, pictures, and props to show objects and situations, and asked two parallel focal questions (see Table 1). The diverse-desires, diverse-beliefs, and explicit-false belief tasks formed one subset of tasks in which children saw a toy figure and a drawing with two pictured choices (e.g., egg and ice cream or bed and cupboard). To answer the questions, the children verbally chose one of the pictured options. The format was similar for the hidden-emotion task, but in this case children chose one of three pictured options. For the knowledge-ignorance and contents false-belief tasks, children saw a container with a hidden item or items inside (e.g., a potato-chip tube with pencils inside), and answers were verbal choices (e.g., "Does he think there are chips or pencils?"). Each subset of tasks included a false-belief task (explicit false-belief vs. contents false-belief). Each child received one trial on each of the six tasks; thus, all children were given both false-belief tasks. We did not expect the two false-belief tasks to scale differently, but rather expected them to be similar in difficulty. Including both tasks allowed us to assess whether responding would be similar with the two differing task formats when conceptual content was meant to be the same.

Procedure

The children were tested in a quiet room in their preschool by one of three trained testers. Each child received all six tasks, presented in one of two orders. In both orders, the diverse-desires task was presented first (to help children warm up to the procedures with a task hypothesized to be easy to respond to), and the hidden-emotion task was presented last. Otherwise, the orders scrambled the tasks in two different sequences; 47 children received Order 1; 45 received Order 2. The two orders were nearly identical to the two orders most frequently used in our previous study (Wellman & Liu, 2004) to allow direct comparison of the results.

RESULTS

Table 1 shows the proportion of children who performed each task correctly. As expected, the two false-belief tasks were similar in difficulty; 54% of the children were correct on the contents false-belief task, and 49% were correct on the explicit false-belief task; a McNemar's chi-square test was not significant. For subsequent analyses, we excluded the explicit false-belief task and looked at the same five core tasks closely examined in our previous study (Wellman & Liu 2004).

A preliminary Age (3) × Order (2) × Gender (2) analysis of variance on each child's total correct (0–5 out of 5 tasks) yielded a main effect of age, $F(7, 80) = 15.34, p < .001$, but no effects of gender or order. Three-, 4-, and 5-year-olds averaged 2.91, 3.69, and 4.23 correct, respectively.

Guttman scales capture sequences in which if a respondent is correct on one item, he or she is correct on all easier items. Table 2 shows the Guttman scaling that best fits the Chinese data; 68% of the children fit this pattern exactly. As calculated via Green's (1956) procedure, the coefficient of reproducibility was .93 (.90 or higher is considered significant). Green's coefficient of consistency, a more conservative measure, was .25 (n.s.; .50 or higher is considered significant). Guttman scale analyses are quite stringent and determinate—a pattern of responses either fits the scale exactly or fails. Contemporary approaches have been developed to allow consideration of less strict, probabilistic progressions. In particular, Rasch analyses (Rasch, 1960; Wright & Masters, 1982) compute probabilistic models for Guttman-like progressions. Rasch analyses derive best-fit scale sequences, indices of fit for derived sequences, and estimates of the "distance" between items for a derived scale sequence.

We used the WINSTEPS/BIGSTEPS computer program (Linacre, 2003; see Wellman & Liu, 2004, for details) to conduct a Rasch analysis of the five items. The results are summarized in Table 3. This analysis showed the same scale sequence as obtained with Guttman scaling (Table 2). The infit and outfit indices assess the fit of the derived scale to the data. When infit and outfit indices are less than 2.0 (with means and standard deviations close to expected values of 0 and 1.0, respectively), as in the present case, the sequence is considered stable and scalable.

In Table 3, for simplicity, the item-difficulty scores on the linear logits scale derived by the analysis were rescaled so that the contents false-belief task (arbitrarily considered the anchor of the five tasks) had a scale score of 5, and the other scores show scale distances relative to that anchor. For comparison, the scale scores for the Rasch analysis of U.S. data (Wellman & Liu, 2004), again using the contents false-belief task as the anchor, are shown. The distance from the contents false-belief task to the diverse-desires task is very similar for the U.S. and Chinese data (approximately 4.5 scale units), and in both cases, the diverse-beliefs and knowledge-ignorance tasks fall in between. How-

TABLE 3

Item Measure Summary and Fit Statistics for the Five-Item Rasch Model

Task	Scale score	Comparison score (Wellman & Liu, 2004)	Error	Standardized infit	Standardized outfit
Hidden-emotion	6.36	7.73	0.34	1.5	1.1
Contents false-belief	5.00	5.00	0.27	−3.2	−2.2
Diverse-beliefs	3.65	2.43	0.34	1.5	1.6
Knowledge-ignorance	2.80	3.61	0.36	−0.5	−1.1
Diverse-desires	0.75	0.48	0.63	−0.1	0.1
Mean	3.71		0.39	−0.2	−0.1
SD	1.91		0.13	1.7	1.4

Note. Standardized infit and standardized outfit have expected means of 0 and expected standard deviations of 1.0; fit statistics > 2.0 are indicative of misfit. Rasch analyses also provided data for individuals' fit to the scale. Summarizing for individuals, infit was −0.1 (*SD* = 0.5) and outfit was 1.1 (*SD* = 0.6); again, fit statistics > 2.0 are indicative of misfit.

ever, for the U.S. data, the diverse-beliefs task is easier than the knowledge-ignorance task (with a distance of approximately 1.0 between them), whereas for the Chinese data, the order is reversed (with distance of approximately 1.0, but in the opposite direction).

These results suggest an overall similarity between the sequences in China and the United States, but also a crucial difference regarding the scale order of understanding knowledge-ignorance and diverse beliefs. The similarity is that both U.S. and Chinese children's understanding appears to progress systematically from diverse desires to knowledge-ignorance to contents false belief to hidden emotion. To confirm this progression, we tested a four-item scale of just these tasks, removing the diverse-beliefs task. In a Guttman analysis of that scale, 86% of Chinese children fit the scale exactly (diverse desires < knowledge-ignorance < false belief < hidden emotion); both reproducibility (.96) and consistency (.50) were significant. Analysis of this same four-item Guttman scale for the U.S. data (Wellman & Liu, 2004) produced very similar results; 88% of the U.S. children fit the same four-item scale exactly, and both reproducibility (.96) and consistency (.63) were significant. (Rasch analyses showed similar equivalence of the four-item scales for the two groups, as expected when responses so closely fit the stricter Guttman model.) Comparisons of pairs of tasks confirmed these similarities. For the Chinese children, the diverse-desires task was easier than the knowledge-ignorance task, McNemar's $\chi^2(1, N = 14) = 11.13$, $p < .001$; the knowledge-ignorance task was easier than the contents false-belief task, $\chi^2(1, N = 27)) = 40.20$, $p < .001$; and the contents false-belief task was easier than the hidden-emotion task, $\chi^2(1, N = 39) = 5.49$, $p < .02$. The same comparisons were significant for the U.S. children (all *p*s < .02; see Wellman & Liu, 2004).

Consider next the key difference, which concerns which sort of cognitive mental state is easier for children to judge. For Chinese children, understanding knowledge-ignorance is first and easier than understanding diverse beliefs, but for English-speaking children, the order is the reverse, with understanding diverse beliefs first and easier than understanding knowledge-ignorance. This difference was tested directly by comparing children in the United States and China who failed one but not both of these tasks. Sixty-four percent (18 of 28) of the Chinese preschoolers who failed one of these tasks failed the diverse-beliefs task and passed the knowledge-ignorance task. Sixty-eight percent (15 of 22) of U.S. children who failed one of these tasks failed the knowledge-ignorance task and passed the diverse-beliefs task (Wellman & Liu, 2004), $\chi^2(1, N = 50) = 3.98$, $p < .05$.

We confirmed this key comparison between Chinese- and English-speaking children by collecting further data in China and comparing them with data from 60 English-speaking Australian preschoolers (Peterson et al., 2005). In this extension of our study, there were 48 preschoolers, eighteen 3-year-olds, fourteen 4-year-olds, and sixteen 5-year-olds, from a third Beijing preschool with a makeup similar to that of the first two. (In Peterson et al., the diverse-beliefs task was always presented immediately before the knowledge-ignorance task. For the additional Beijing data, preschoolers were also given the diverse-beliefs task immediately before the knowledge-ignorance task.) In these two groups, 89% of the 28 Chinese children who passed one but not both tasks passed the knowledge-ignorance task and failed the diverse-beliefs task, whereas 56% of the 16 Australian preschoolers who passed one but not both tasks showed the opposite pattern, passing the diverse-beliefs task and failing the knowledge-ignorance task, $\chi^2(1, N = 44) = 8.47$, $p < .01$.

Across all these studies, there are 140 Chinese preschoolers (from Beijing) and 135 English-speaking preschoolers (from the United States and Australia). Of those children who failed one but not both tasks (56 Chinese children and 38 English-speaking children), 77% of Chinese children failed the diverse-beliefs task (while passing the knowledge-ignorance task), whereas 63% of English-speaking children failed the knowledge-ignorance task (while passing the diverse-beliefs task), $\chi^2(1, N = 94) = 13.51$, $p < .001$.

DISCUSSION

Data assessing theory of mind from decidedly non-Western communities and cultures have typically focused on single task paradigms, essentially the false-belief paradigm (e.g. Callaghan et al., 2005; Vinden, 1999). Worldwide, it seems that children come to understand false belief in childhood, but the timing of that development is variable—as early as 3 1/2 years on average in some countries, but 6 or 7 years in others (Wellman et al., 2001). Do differences across countries represent simply delays and accelerations, or do they reflect different paths of conceptual development? Data about extended progressions of mental-state understandings can address such questions, but until now, data on sequences of understanding have come only from children growing up in Western cultures.

Chinese children, just like typical children elsewhere, show a consistent sequence of theory-of-mind understanding in the preschool years, an important finding in itself. Not only is their sequence consistent, but it is also similar in many ways to that evident for children in the United States and Australia. The essential similarities were clear in our analyses of a four-item scale, which showed closely parallel sequences and levels of significance. Both English-speaking children and Chinese children demonstrated an early understanding of desires that preceded understanding of a sequence of cognitive mental states (knowledge, false belief) that preceded understanding of hidden emotions. This sequence was apparent even considering pairs of tasks with almost identical task formats, for example, the diverse-desires task versus the diverse-beliefs task or the knowledge-ignorance task versus the contents false-belief task. Certain progressions in theory-of-mind development, such as the ones demonstrated here, may be universal.

At the same time, there was a salient and intriguing difference. For Chinese children, earliest understanding of cognitive mental states was evident on the knowledge-ignorance task, whereas for English-speaking children, earliest understanding was evident on the diverse-beliefs task. This finding seems robust in that it was initially demonstrated in a Beijing-U.S. comparison but then confirmed with additional data from China and Australia. In our final analysis, this reversal of sequence was confirmed for 135 English-speaking and 140 Chinese children. Moreover, the tasks are rigorous and revealing in that they fit within a larger, scaled sequence of related understandings. Indeed, the current comparison across cultures seems especially informative because it rests on scale sequences rather than exact ages of attainment or absolute mean differences. Furthermore, it seems unlikely that this key difference in sequences is simply due to peculiar tasks or task formats; the fact that the order of difficulty reversed across groups indicates that one task was not intrinsically more difficult than the other because of complexity of language or format.

These data shed light on several alternative hypotheses. First, it seems unlikely that these sequences represent mere maturational unfolding of modular theory-of-mind conceptions, because the sequences differed importantly in the two cases. These sequences are also unlikely to simply represent childhood increments in executive function, or cognitive complexity, manifest on tasks that increase such demands step by step (Frye, Zelazo, & Palfai, 1995). Previously (Wellman & Liu, 2004), we argued that all our scale tasks place very similar demands on executive function and require similar levels of cognitive complexity. For example, all deal with two alternatives, and one alternative must be inhibited to correctly choose the other (e.g., diverse-desires task: inhibit my preference, answer on the basis of the other person's preference; diverse-beliefs task: inhibit my belief, answer on the basis of the other person's belief; knowledge-ignorance task: inhibit my knowledge or reality, answer on the basis of the other person's knowledge). More important, however, suppose one were to devise more subtle task analyses to claim that there are step-by-step increments in executive-functioning or complexity demands across these tasks. Any such analyses proposed to account for the detailed U.S. data (e.g., that the knowledge-ignorance task requires more inhibition or more complex reasoning than the diverse-beliefs task) would be challenged by the Chinese data. Our sequence differences for understanding diverse beliefs and knowledge-ignorance seem to require a different sort of account.

We argue that culturally shaped differences in input are at work. More precisely, working within a common (and possibly universal) set of theory-of-mind insights, children in these different cultural communities receive different information and experiences pertaining to mental states. These differences result in different sequences of understanding that are apparent quite early in development. This hypothesis garners important support from the exact nature of the differences we found. That is, our findings are consistent with a Chinese cultural emphasis on knowing, and on children acquiring practical knowledge (of the sort we outlined in our introduction), and with a contrasting Western emphasis that is more slanted toward belief, truth, falsity, and differences in belief. As a result of these different emphases, knowing may be more salient and important in the conceptual lives of young Chinese children, and thinking and believing may be more salient and important for Western children. Thus, for Western children in the United States and Australia, the earliest understanding of cognitive mental states (built on an earlier understanding of desires) may be some sense that two people can have differing thoughts about the same situation. For Chinese children in mainland China, that key step beyond an understanding of desires may more likely come in the appreciation that persons can be knowledgeable versus ignorant.

These data underwrite descriptions of theory of mind that emphasize sociocultural influences on children's developing theory of mind (Lillard, 1998; Wellman & Miller, 2006), as well as descriptions that emphasize universal acquisitions during childhood (Callaghan et al., 2005; Wellman et al., 2001). A full account must include both factors. Of course, our data cannot fully address universalities or sociocultural differences; they encompass a limited set of tasks, only three communities (the United States, China, and Australia), and three languages (English, Mandarin, and Australian Sign). Nonetheless, the current data contribute to a fuller account of theory of mind by showing differences as well as similarities in developmental sequences of understanding in different cultural contexts.

Acknowledgments—Support for this research came from the National Institute of Child Health and Human Development (Grant HD-22149) and from the Chinese Academy of Sciences.

REFERENCES

Anderson, D., & Reilly, J. (2002). The McArthur communicative developmental inventory: Normative data for American sign language. *Journal of Deaf Studies and Deaf Education*, *7*, 83–104.

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

Callaghan, T., Rochat, P., Lillard, A., Claux, M.L., Odden, H., Itakura, S., Tapanya, S., & Singh, S. (2005). Synchrony in the onset of mental-state reasoning. *Psychological Science*, *16*, 378–384.

Dunn, J. (1995). Children as psychologists: The correlates of individual differences in understanding of emotions and minds. *Cognition & Emotion*, *9*, 187–201.

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

Frye, D., Zelazo, P.D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483–527.

Green, B.F. (1956). A method of scalogram analysis using summary statistics. *Psychometrica*, *21*, 79–88.

Harris, P.L., Rosnay, M., & Pons, F. (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science*, *14*, 69–73.

Kessen, W. (1975). *Childhood in China*. New Haven, CT: Yale University Press.

Lee, K., Olson, D.R., & Torrance, N. (1999). Chinese children's understanding of false beliefs: The role of language. *Journal of Child Language*, *26*, 1–21.

Li, H., & Rao, N. (2000). Parental influences on Chinese literacy development: A comparison of preschoolers in Beijing, Hong Kong and Singapore. *International Journal of Behavioral Development*, *24*, 82–90.

Li, J. (2002). A cultural model of learning: Chinese "heart and mind for wanting to learn." *Journal of Cross-Cultural Psychology*, *33*, 248–269.

Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin*, *123*, 3–32.

Linacre, J.M. (2003). *User's guide and program manual to WINSTEPS: Rasch model computer programs*. Chicago: MESA Press.

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

Nisbett, R.E. (2003). *The geography of thought: How Asians and Westerners think differently—and why*. New York: Free Press.

Peterson, C.C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind & Language*, *15*, 123–145.

Peterson, C.C., Wellman, H.M., & Liu, D. (2005). Steps in theory of mind development for children with autism and deafness. *Child Development*, *76*, 502–517.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*, 734–751.

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

Tardif, T., Wellman, H.M., & Cheung, K.-M. (2004). False belief understanding in Cantonese-speaking children. *Journal of Child Language*, *31*, 779–800.

Tobin, J.J., Wu, D.Y.H., & Davidson, D.H. (1989). *Preschool in three cultures*. New Haven, CT: Yale University Press.

Vinden, P.G. (1999). Children's understanding of mind and emotion: A multi-cultural study. *Cognition & Emotion*, *13*, 19–48.

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.

Wellman, H.M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development*, *75*, 523–541.

Wellman, H.M., & Miller, J. (2006). Developing conceptions of responsive intentional agents. *Journal of Cognition and Culture*, *6*, 27–55.

Wright, B.D., & Masters, G.N. (1982). *Rating scale analysis: Rasch measurement*. Chicago: MESA Press.

(Received 11/11/05; Revision accepted 3/14/06;
Final materials received 3/20/06)