We compared 20 prelingually profoundly deaf adolescents (age: 11–16 years) and 20 matched, hearing adolescents on a picture-sequencing task and on a social judgment test. In addition, we also tested 14 younger deaf children (age: 6–10 years) and compared their data with those from 20 hearing peers as well as those from the older deaf participants on the picture-sequencing task. The results from this study did not provide evidence for the hypothesis that deaf adolescents possess significantly poorer knowledge about social reasoning than age-matched hearing peers, but it did present further additional support for Peterson and Siegal’s (1995) conversational hypothesis: a proposal that a deprivation in conversations about mental states leads to an impairment in the development of an awareness of mental states in the younger deaf children.

The term “theory of mind” (ToM) describes the awareness in an individual of human behavior as ruled by a complex system of mental states, involving feelings, beliefs and desires (Premack & Woodruff, 1978; Wellman 1990). Recent research has focused on the development of moral and social reasoning abilities in both “normal” children and children with various disabilities. For example, Leslie (1987), Butterworth, Harris, Leslie, and Wellman (1991), and Perner (1991) have studied how ToM develops in children, both with and without cognitive or sensory problems. Baron-Cohen, Leslie, and Frith (1985) conducted a study of the concept with autistic children, the results of which consistently indicated serious deficits or delays in their acquisition of ToM, a finding also confirmed in other studies (e.g., Eisenmajer & Prior, 1991; Leslie & Frith, 1988; Perner, Frith, Leslie, & Leekam, 1989; Prior, Dahlstrom, & Squires, 1990; Reed & Peterson, 1990; Roth & Leslie, 1991). These findings led to a number of suggestions to explain why autistic children apparently fail to perform well on ToM tasks, in particular “false belief” tasks. Although the autistic children in the Baron-Cohen et al. (1985) study possessed a much higher mental age than the nonautistic children, their performance as a group was significantly worse than that of other children: they failed to comprehend the difference between their own beliefs and those of another person.

In response to this discovery, Frith (1992) suggested that such a persistent failure is caused by an abnormal neurological development specific to autism (Frith, 1989, 1992; Frith, Morton, & Leslie, 1991). Research using the false belief test, in clinical populations of children with linguistic, emotional, and intellectual disorders, provided some support for Frith’s theory (e.g., Baron-Cohen, 1989; Baron-Cohen et al., 1985; Leslie & Frith, 1988; Perner et al., 1989; Prior et al., 1990; Reed & Peterson, 1990; Siddens, Happé, Whyte, & Frith, 1990). An example of a false belief task is the Sally-Anne task, which involves two dolls: one doll, Sally, hides a marble in a basket and leaves the
Peterson and Siegal replicated the study of false belief, the Sally-Anne test, by Baron-Cohen et al. (1985), with Australian signing, prelingually deaf children. The results obtained from their study revealed that a majority of the deaf children failed the simple test of false belief. The results also displayed no significant difference in the performance between that of the deaf children and the autistic children tested in the Baron-Cohen et al. study. Peterson and Siegal explained their results in terms of their theory of lack of exposure to communication about mental states, which affects the development of ToM both in deaf children and in children with autism.

Baron-Cohen and his colleagues conducted another study devised to investigate mechanical, behavioral, and intentional understanding of picture-stories in autistic children and children with Down's syndrome (Baron-Cohen, Leslie, & Frith, 1986). The intentional condition was designed to require understanding of the protagonists' mental states in the stories. Results revealed that autistic children performed considerably better on the mechanical and behavioral stories; however, their performance on the intentional stories was significantly worse. The narratives collected from the autistic children in the picture-sequencing task revealed evidence that autistic children engaged in mechanical- and behavioral- but rarely mental-state language. Baron-Cohen et al. claimed that this study supported their theory of a specific cognitive deficit that causes the apparent impairment in the development of ToM in the autistic child.

By focusing on deafness, the Peterson and Siegal (1995) study has introduced the possibility that the deficit in the performance of the autistic child on false-belief tasks is caused not by a specific cognitive failing but, instead, is the indirect result of a lack of communication about mental states, which leads to a restriction or impairment in his or her development of ToM. This theory can be applied to the apparent lack of understanding of social and moral reasoning demonstrated by a majority of prelingually deaf children. According to research cited in Marschark (1993) (e.g., Bachara, Raphael, & Phelan, 1980; Cates & Shontz, 1990; Kusché & Greenberg, 1983) and Peterson and Siegal (1995), then, prelingually deaf children should perform poorly on tasks concerning mental states of others (as
also should deaf adolescents and young adults, according to Couch, 1985; DeCaro & Emerton, 1978; and other research). Recent studies have concentrated on possible mediating variables. Russell et al. (1998) gave a false-belief test of ToM to deaf children and adolescents ages 4 to 16 years and discovered that performance on the test appeared to be age-related, with the older participants performing significantly better than the younger deaf children. Russell et al. proposed that the ToM “abilities of deaf children are most appropriately described as being subject to a developmental delay.” Russell et al. emphasize the importance of deaf children receiving sufficient communication about mental states in a social context.

Peterson and Siegal extended the proposal of language environment affecting the development of ToM by conducting a series of studies (1996, 1997, 1998, 1999), which all provided apparent evidence that conversational deprivation leads to an impairment in the development of ToM abilities in deaf children. Their 1999 study investigated the performances of deaf children, autistic, and normal hearing children on a range of mental state tasks requiring representations of other people’s “inner worlds.” The deaf children in the study were categorized into three “conversational backgrounds and modality”: native signers; signers from hearing families; and oral deaf. On the tasks requesting an understanding of others’ mental states, the performances of the native signers, oral deaf, and the normal hearing children were similar. In contrast, the performance of the signers from hearing families and children with autism did not differ significantly from one another: each group was significantly inferior to the other groups. The mean chronological age of the deaf children used in this study was 9 years, 5 months, which was 5 years older than the mean chronological age of the hearing children used in the same study (4 years, 6 months).

These studies on ToM abilities in deaf children lead Marschark, Green, Hindmarsh, and Walker (in press) to conduct a study investigating stories told by groups of deaf and hearing children and adolescents (age: 9–15 years), examining the attribution of mental states to themselves and others. The results from their study appeared to contradict those of Peterson and Siegal (1995, 1996, 1997, 1998, 1999), demonstrating that the deaf participants between the ages of 9 and 15 years performed successfully and thus suggesting healthy ToM abilities. As Marschark et al. (in press) noted, however, their study also did not include deaf children younger than 9 years old.

Our study involved two main components investigating the abilities of deaf children (age: 6–10 years) and adolescents (age: 11–16 years) to make social judgments as compared with their hearing peers. The first component was a replication of the Baron-Cohen et al. (1986) experiment with autistic children on mechanical, behavioral, and intentional understanding of picture-stories with deaf signing children and hearing children. We chose the Baron-Cohen et al. adaptation of Bullock’s (1984) picture-sequencing technique because it is a visual test and because of its success with very young children in the understanding of causal relations. The test was designed to assess the ability of the child to understand the mental states of the characters, which differs from false belief tasks, in which the child has to predict characters’ behavior when their beliefs conflict with reality. If the deaf children perform significantly worse on the intentional stories, as the autistic children did in the Baron-Cohen et al. study, this result will provide additional support for the results of the Peterson and Siegal (1995, 1996, 1997, 1998, 1999) studies. If the deaf children perform similarly to the hearing children, however, this result will question the nature of materials used in preceding research (i.e., did the deaf children in those studies understand the experimenters?) and will also raise the possibility that deaf children are gaining greater access to information than generally believed, through improved communication.

With the aim of increasing the validity and reliability of the social reasoning test, we included an additional assessment. The social judgment test applied in the study of young people with Asperger’s syndrome (Ellis, Ellis, Fraser, & Deb, 1994) was adopted, after small adjustments, for our research. The social judgment test was devised to reflect the difficulties frequently observed in people with Asperger’s syndrome when requested to make social judgments about people’s behavior. It poses hypothetical social judgment questions, which require knowledge of everyday cultural conventions and empathy. The test was created
with the intention of revealing any impairment or difficulties in the social judgment abilities of adolescents and young people. If the deaf adolescents perform in a manner similar to that found in young people with Asperger's syndrome, this may indicate that deaf adolescents and adolescents with Asperger’s syndrome have limited opportunities to develop ToM and thus to learn social skills.

Thus, this subsequent research replicated the Baron-Cohen et al. (1986) study with the intention of investigating the proposal of an impairment in the development of ToM in deaf children (age: 6–10 years) as indicated by Peterson and Siegal (1995, 1996, 1997, 1998, 1999) and Russell et al. (1998). In addition, the study investigates social awareness of deaf adolescents (age: 11–16 years), by replicating the Ellis et al. (1994) study, with the proposal that the ability to make social judgments indicates healthy ToM.

The hypotheses are (1) the younger deaf children will perform significantly worse than their hearing peers of the same ages and the older deaf participants on tasks requiring attribution of mental states and (2) the deaf adolescents will possess significantly poorer knowledge about social reasoning than hearing peers of the same ages.

Method

Experimental Measures

In this study, we used two experimental measures to assess the participants’ understanding of social reasoning. In the Baron-Cohen et al. (1986) picture-sequencing task, from the ordering of the pictures, one can determine how well the child understands the intended story, even in the absence of words. The picture-sequencing paradigm provides a technique for collecting narrative information from the participants that allows the child's interpretation of the picture-story sequences to be checked. This technique gives the chance of obtaining strong evidence of the child's understanding of the picture-stories since words/signs provide a more direct insight into what the child perceives and understands the intended story to be. Baron-Cohen et al. (1986) devised three types of picture story with the hypothesis that the autistic child will understand certain events well and other events comparatively poorly. The first type of picture story was mechanical, which depicted physical-causal relations (e.g., balloon flies to tree and bursts on a sharp branch); the second type was behavioral and showed people involved in various activities and interactions (e.g., boy dressing himself). Baron-Cohen et al. asserted that the second type of picture story could be perfectly understood simply by reference to behavior without reference to mental states, the critical issue behind the third type of picture story. The third type, intentional, depicted people engaged in everyday activities, and in order for the intended story to make sense as a story, a state of mind needs to be attributed to a protagonist, thus involving intuitive and immediate understanding of mental states, not just the behavior.

Baron-Cohen et al. stated that the test was designed with the purpose of the first two story types being the most “naturally” understood in mechanical and behavioral terms; and the third in intentional terms. Their results indicated that the act of mentalizing in the intentional stories demands a very different type of ability, which is involved in causal and behavioral thinking. On the picture-story sequencing task, five types of story, one for each of the five conditions, were used. The conditions were, as according to Baron-Cohen et al., (1) mechanical 1: objects interacting causally with each other; (2) mechanical 2: people and objects acting causally on each other; (3) behavioral 1: a single person acting in everyday routines not requiring attribution of mental states; (4) behavioral 2: people acting in social routines, involving more than one person, but not requiring attribution of mental states; and (5) intentional: people in everyday activities requiring attribution of mental states.

For each of the five conditions, there were three different stories (see Table 1). The pictures used in these studies were exact replicas of those used in Baron-Cohen et al. and were drawn in color onto white cards, 5 × 5 inches. Appendix 1 provides illustrative examples of the stories used: mechanical, behavioral, and intentional. Baron-Cohen et al. decided upon a standard length of four pictures since pilot studies indicated that this length was appropriate for the younger children.

The presentation of the pictures was identical to that of Baron-Cohen et al.; the first of the four pictures
for each story was always placed in front of the child, as the start of the sequence, so that the child had only three pictures to arrange into a story sequence. This was to reduce the ambiguity of possible stories that could be constructed from the four pictures. The order of presentation of the condition was fixed (M1, M2, B2, I, B1), which was identical to that of Baron-Cohen et al. As in their procedure, the mechanical stories were always introduced first, in the assumption that these stories were the simplest. Baron-Cohen et al. placed the intentional condition, the critical condition, in the penultimate position, which allowed it to acquire any benefits of practice. In order to evaluate any differential effects of practice/fatigue on the groups of subjects used in their study, the behavioral 1 condition always ended the experimental session.

The social judgment test, devised by Dewey (1991) and modified by Ellis et al. (1994), was chosen as a suitable means of measuring social awareness in pupils ages 11 to 16 years. The test complements and augments the picture-story sequencing paradigm. It poses hypothetical social judgment questions, which require knowledge of everyday cultural conventions and empathy. The questions are presented in the context of short stories, each of which involves a series of behaviors to be rated on a four-point scale: A = fairly normal behavior in that situation; B = rather strange behavior in that situation; C = very eccentric behavior in that situation; and D = shocking behavior in that situation.

Ellis et al. (1994), to suit British adolescents in their research, modified the eight scenarios. Four scenarios were selected for this study and were adjusted to achieve a smooth translation of the English language into British Sign Language (BSL); in this way deaf and hearing participants received effectively identical presentation of the scenarios (see Appendix 2). In the previous studies, of course, the social judgment test was presented in a written form and the participant made his or her responses to each of the behaviors described in the stories by choosing one of the letters (A, B, C, or D) representing the scale. In our study, however, we decided to sign each story individually to the deaf participants, using sign language, and to read aloud each story to the hearing participants. Evidence indicates that deaf children, as a group, are significantly poorer at reading English language than their hearing peers of the same age (e.g., Allen, 1986; Di Francesca, 1972; Vernon, 1972; Waters and Doehring, 1990). The translation of English language into BSL was checked for accuracy following discussions with Deaf BSL teachers and interpreters.

Participants

Thirty-four severely and profoundly prelingually deaf pupils, from 6–16 years old, were selected from a local school for the deaf in Wales (UK). These participants all used BSL as their first or preferred communication
matrices test, 1983 edition), gender, and age. The background data for age (CA) and nonverbal intelligence quotient (as determined by the Raven’s Progressive Standard Matrices Test, 1983 edition) are presented in Table 2. The gender ratio was approximately 2:1 with 13 girls and 7 boys in the older groups of participants. In the group of younger deaf participants, 10 boys and 4 girls participated in the research.

Procedure

The participants were tested individually. An initial interview was designed to put them at their ease with the deaf experimenter (Rhys-Jones). They then were given the picture-sequencing test and the social judgment test. The order of presentation of the tests was randomized for the adolescents. Each task was conducted in sign language, by the experimenter, with the deaf participants and in speech, through an interpreter, with the matched hearing participants. The procedures were identical for each group.

The procedure for the picture-story sequencing
task was identical to that in Baron-Cohen et al. (1986). The order of presentation of the three story pictures within each trial condition was randomized. As stated previously, the order of the five conditions was always fixed (M1, M2, B2, I, B1). The experimenter placed the first picture of each story in front of the child and the remaining three pictures in random order above the first story picture. Each participant was given the instruction, “This is the first picture. Look at the other pictures and see if you can make a story with them.” The experimenter recorded the order of sequence chosen by the participant, after any self-corrections. Each participant was allowed to proceed at his or her own pace but was allowed only one attempt at each of the 15 picture stories. A completely correct sequence, according to the predetermined order employed by Baron-Cohen et al. (1986), was awarded 2 points.

The narration component of the picture–story sequencing test was conducted by asking the participant, on completion of the story sequence, “Can you tell me the story?” With the deaf participants, the experimenter wrote down the narration of each participant, noting the order of the signs and later translating into written English. The narration of each hearing participant was recorded on a tape-recorder operated by the interpreter, and the interpreter later transcribed the tapes.

On the social judgment test, the experimenter informed each participant, tested individually, that he or she would be told stories about the behavior of people in everyday life and would be asked what he or she thought of that person’s behavior. The participants were allowed to proceed at their own pace and were told that there were no “right” or “wrong” answers to the stories. In the stories, behaviors were underlined and after each underlined story part, there was a pair of brackets ( ). For each of these, the participant was asked to choose one of the following: A = normal behavior in that place; B = strange behavior in that place; C = very strange behavior in that place; and D = awful behavior in that place, according to how the adolescent felt about the person’s behavior if they saw it. The experimenter placed the chosen response in the bracket for each underlined story part.

Results

Sequencing

In interpreting the results of the sequencing task, we used the critical scoring procedure designed by Baron-Cohen et al. (1986). A completely correct sequence was awarded 2 points. If only the last picture card was placed correctly, 1 point was awarded. The other four potential picture orders scored zero. This scoring procedure was designed by Baron-Cohen et al. to minimize correct responses based on erroneous reasoning.

For each of the five conditions, the maximum score was 6 (three trials in each condition). The results are displayed in Table 3 and show a “ceiling effect.” From the results of Peterson and Siegal (1995, 1996, 1997, 1998, 1999), we expected that the performance of the deaf children would be significantly poorer on the sequencing of the intentional stories than that of their hearing peers. In addition, we expected that the performance of the older deaf participants (age: 11–16 years) would be significantly better than that of the younger deaf participants (age: 6–10 years).

A 2 (group) × 2 (age) × 5 (conditions) analysis of variance (ANOVA) was conducted on the results. No effect of group was observed in any of the five conditions, indicating that the deaf participants performed similarly to the hearing participants: mechanical 1, F(1, 70) = 1.76, p > .05; mechanical 2, F(1, 70) = 1.74, p > .05; behavioral 1 F(1, 70) = 1.60, p > .05); behavioral 2, F(1, 70) < 1; and intentional, F(1, 70) = 1.50, p > .05. The effect of age was not observed in mechanical 1, F(1, 70) = 1.76, p > .05; mechanical 2, F(1, 70) = 1.74, p > .05; behavioral 1, F(1, 70) = 1.74, p > .05; or behavioral 2, F(1, 70) = 3.06, p > .05, revealing that the younger children’s performance was not significantly different from the adolescents on these conditions. However, the effect of age was observed in the intentional condition, F(1, 70) = 10.07, p < .002, revealing that the younger children, in general, performed significantly worse than the older participants. The interaction between group and age, however, failed to be significant for any of the conditions.

An ANOVA was conducted separately on the results of the older participants and the 2 (groups) × 5 (conditions) interaction yielded no significant differ-
ences. The ANOVA revealed no significant difference between the performance of the older deaf participants and the performance of their hearing peers on any of the five conditions, $F(1, 39) < 1$. No significant difference was found between the performance of the younger deaf children and their hearing peers on mechanical 1, $F(1, 33) = 3.20, p > .05$, and mechanical 2, $F(1, 33) = 3.17, p > .05$; behavioral 2, $F(1, 33) < 1$, and the intentional, $F(1, 33) = 2.73, p > .05$, conditions. However, the deaf children performed significantly worse than their hearing peers on the sequencing task in behavioral 1 condition, $F(1, 33) = 4.03, p < .05$.

The 2 (age) $\times$ 5 (conditions) ANOVA was conducted on the scores of the deaf participants and revealed that the younger deaf children performed significantly worse on the behavioral 1 condition, $F(1, 33) = 6.06, p < .02$, and on the intentional condition, $F(1, 33) = 8.81, p < .01$, than the older deaf participants. There was no significant difference between the performance of the younger deaf children and the older deaf participants on mechanical 1, $F(1, 33) < 1$, mechanical 2, $F(1, 33) = 3.17, p > .05$, and on behavioral 2, $F(1, 33) = 1.25, p > .05$.

An ANOVA indicated no significant difference between the performance of the hearing participants on mechanical 1, $F(1, 39) = 3.89, p > .05$, and mechanical

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mec. 1</th>
<th>Mec. 2</th>
<th>Beh. 1</th>
<th>Beh. 2</th>
<th>Int.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaf (Secondary) ($n = 20$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>6</td>
<td>6</td>
<td>5.6</td>
<td>5.25</td>
<td>5.80</td>
</tr>
<tr>
<td>SD</td>
<td>—</td>
<td>—</td>
<td>.60</td>
<td>.85</td>
<td>.41</td>
</tr>
<tr>
<td>Hearing (Secondary) ($n = 20$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>6</td>
<td>6</td>
<td>5.5</td>
<td>5.4</td>
<td>5.80</td>
</tr>
<tr>
<td>SD</td>
<td>—</td>
<td>—</td>
<td>.61</td>
<td>.75</td>
<td>.41</td>
</tr>
<tr>
<td>Deaf (Primary) ($n = 14$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>6</td>
<td>5.86</td>
<td>4.79</td>
<td>4.86</td>
<td>4.79</td>
</tr>
<tr>
<td>SD</td>
<td>—</td>
<td>.54</td>
<td>1.48</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>Hearing (Primary) ($n = 20$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>5.90</td>
<td>6</td>
<td>5.45</td>
<td>4.95</td>
<td>5.35</td>
</tr>
<tr>
<td>SD</td>
<td>0.31</td>
<td>—</td>
<td>1.05</td>
<td>.99</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Narration

Participants in each of the four groups were completely cooperative in providing signed (deaf participants) or verbal (hearing participants) descriptions for all of the 15 picture sequences.

The narratives were scored in one of the following categories; causal, descriptive, or mental state, and the scoring procedure was identical to the procedure devised by Baron-Cohen et al. (1986). In scoring, the narratives were not judged as to whether they were “correct” or whether the participant related well to a specific event in a story. Instead, the scoring procedure examined only the surface form of the descriptions, and, for each picture story, the narration, regardless of length, was categorized into causal, descriptive, or mental state. Each narration was scored for the presence or absence of causal and mental state expressions and, if neither of these was found, it was categorized as descriptive. The following examples are considered as sources for causal, mental state, or descriptive language terms, according to Baron-Cohen et al. They defined the distinctive aspects of the three categories:
The Causal category was applied only when the utterance contained at least one of the following: because clauses, e.g. The egg broke because it fell off the table (M.1.2); explicit mention of agent-causal verb-object, or passive construction with by-phrase, e.g. The boulder broke the tree (M.1.1). The man was hit by the rock (M.2.2); causal verb phrase e.g. made . . . happen: The rock made the man fall down (M.2.2).

A few narrative examples were judged to be too ambiguous to be categorized as causal, considering the definition of causal by Baron-Cohen et al. (e.g., “It went near the tree and it went ‘bang’ [M1.3]; It rolls, then stops at the tree and the tree fell down [M1.1]”).

The narratives categorized as mental state had to contain at least one of the following definitions as proposed by Baron-Cohen et al.: (a) a mental state expression (want, believe, pretend, wish, etc.), e.g. He wanted to buy sweets (I.1); the boy didn’t know that she pinched his chocolate (I.3); (b) an implicit attribution of a mental state, e.g. The boy was surprised cuz he couldn’t find his chocolate (I.3); an attribution of an utterance to the protagonist appropriate to his or her mental state often marked by special intonation, e.g. He’s shouting, ‘Where’s my sweet gone?!’ (I.1).

The narrative examples not categorized as mental state, as a consequence of being equivocal in terms of intentional language, were scored as descriptive (e.g., “She put the teddy bear down, then the boy picks the teddy up and walks away with it when she is picking the flowers and then when she turns around, the teddy’s gone [I.2]; He goes out to play and his mother eats the sweet in the box and when he goes back, the box is empty [I.3]; A boy buys some sweets and then he comes out of the shop and there’s a hole in the bag, so when he gets home, there’s nothing left in the bag [I.1]”).

In the scoring procedure of Baron-Cohen et al., the descriptive category was applied clearly in narrative examples such as “A little girl goes into a sweet shop and she buys some sweets and she comes out with her sweets (B2.2); The man is making a pizza, he puts it in the oven and when he takes it out, the children eat it (B2.2); The girl is walking towards the boy with an ice-cream, she sits down on the bench then takes the ice-cream and she walks off (B2.3); The boy turns on the shower and has a wash, then comes out and dries himself (B1.1).” Those narrative examples contained no causal or mental state expressions and thus were categorized as descriptive.

In the scoring procedure, the ratings for each participant were converted into proportions relative to the total number of trials where narratives were collected. Table 4 shows the overall results for the four groups of participants. With the adolescents, in the behavioral condition, the narratives not categorized as behavioral were categorized as mental state; and in the intentional condition, the narratives not categorized as intentional were categorized as descriptive. In the mechanical condition, 8 narratives of the older deaf participants were classified as descriptive and 13 were classified as mental state. With the older hearing participants, in contrast, four narratives were classified as descriptive and nine as mental state.

With the younger participants, the complement of the proportions was always mental state for the behavioral condition and descriptive for the intentional condition. In the mechanical condition, the complement was descriptive, with one exception of a narration from a deaf child that was classified as mental state.

The results are particularly interesting when comparing the responses in the behavioral condition and in the intentional condition, the critical condition. The data were sufficiently normally distributed to permit ANOVAs (groups × age × appropriate narratives). On the mechanical condition, there was no significant difference when comparing the performance of the participants in all four groups: older participants, $F(1, 39) < 1$; younger participants, $F(1, 33) = 1.38, p > .05$; deaf participants, $F(1, 33) < 1$; and the hearing participants, $F(1, 39) = 3.80, p > .05$. This reflects both the success and the consistency of their performance on sequencing in this condition and indicates that the four groups of participants possess understanding of physical causality. No effect of group or age was observed in the mechanical condition, revealing that all four groups performed similarly.

In the behavioral condition, an effect of group was observed, $F(1, 70) = 15.93, p < .001$, indicating that
the hearing participants, in general, performed significantly more appropriately than the deaf participants. The effect of age was also significant, $F(1, 70) = 7.99, p < .01$, revealing that the younger children, in general, performed significantly different from the older participants. However, the interaction between group and age failed to be significant for the behavioral condition, $F(1, 70) < 1$.

In comparisons of the means of the narratives in the behavioral condition, the older hearing participants used significantly more descriptive terms in their narratives than did the older deaf participants, $F(1, 39) = 8.82, p < .01$; and the younger hearing children used significantly more descriptive terms in their narratives than either their deaf peers, $F(1, 33) = 7.26, p < .01$, or the older hearing participants, $F(1, 39) = 4.43, p < .05$. There was no significant difference in the narratives of the younger deaf children and the older deaf participants in the behavioral condition, $F(1, 33) = 3.64, p > .05$.

In the critical, intentional condition, a significant difference was revealed between the performance of the older hearing and deaf participants. An effect of group was observed, $F(1, 70) = 31.07, p < .0001$, indicating that the hearing participants actually performed significantly worse than the deaf participants. The effect of age was also significant, $F(1, 70) = 5.69, p < .05$, indicating that, in general, the younger participants performed significantly worse than the older participants. The group × age interaction was also found to be significant, $F(1, 70) = 7.99, p < .01$, revealing that the intentional condition interacted with the group and the age variables.

In comparisons of the means of the adolescents, it would seem that the hearing adolescents seldom used mental state expressions, in contrast to the huge proportion of mental state terms in the narratives of the deaf adolescents, $F(1, 39) = 39.07, p < .0001$. This result is inconsistent with the earlier predictions that the deaf adolescents would perform significantly worse than their hearing peers, according to previous researches (e.g., Couch, 1985; DeCaro & Emerton, 1978). It supports the results of Marschark et al. (in press).

In analyzing the responses of the younger participants in the intentional condition, we found no significant difference between the narratives of the younger hearing children and their deaf peers, $F(1, 70)$...
The younger hearing children also did not differ significantly from the older hearing participants, $F(1, 39) < 1$. However, in comparing the responses of the younger deaf participants with the older deaf participants, we found a significant difference in their narratives, $F(1, 33) = 12.38, p < .001$, indicating that the older deaf participants used more mental state words in their narratives than the younger deaf children.

Social Judgment

In the “Story” social judgment test, the responses of the adolescents in this study were scored using the scoring procedure devised by Ellis et al. (1994), who gave the test to 36 adolescents and young adults. Their modal responses were taken and represented the “correct” responses to the scenarios in the test. The scoring criteria used by Ellis et al. were

[a] deviation of one category was scored as 1, a two category deviation as 2, etc. For example, where most control subjects gave the response “B” to a particular behavior, a reply of either “A” or “C” was scored as 1 error and a reply of “D” was scored as 2 errors.

Ellis et al. acknowledged that the scoring method is crude but expedient. The responses of the adolescents in this study were scored using the procedure devised by Ellis et al. and appear in Table 5. An independent $t$ test and a 2 (groups) $\times$ 1 (condition) ANOVA was conducted on the results; and no significant difference in the performance of the participant groups was obtained from either statistical test, $t(39) < 1$.

Table 5

<table>
<thead>
<tr>
<th>Groups</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaf (Secondary)</td>
<td>6.0</td>
<td>2.49</td>
</tr>
<tr>
<td>(n = 20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hearing (Secondary)</td>
<td>4.65</td>
<td>2.94</td>
</tr>
<tr>
<td>(n = 20)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

These results, while contradicting previous research on deaf adolescents 11–16 years old (e.g., Couch, 1985; DeCaro & Emerton, 1978), provide support for the recent study by Marschark et al. (in press). This study provides potential support for the hypothesis following from the proposal of Peterson and Siegal (1995, 1996, 1997, 1998, 1999) for an impairment in the development of ToM in the majority of prelingually profoundly deaf children (age: 6–10 years). The results also appear to support Russell et al. (1998), in that the development of ToM abilities in the majority of deaf children seems to be age-related with the older deaf participants performing significantly superior than the younger deaf children. The revelation of significant and nonsignificant findings raises a number of implications and possible explanations.

The finding that the older deaf participants and the older hearing participants performed similarly in the sequencing of the picture-stories suggests that each group of participants possessed an understanding of physical causality, which is further supported by the lack of any difference in their performance in the mechanical condition. In the behavioral and intentional condition, the deaf adolescents used a significantly greater proportion of mental state expressions in their narratives, whereas the hearing adolescents notably used more descriptive utterances on the same stories. There was no significant difference on the social judgment test.

The results from our research provide evidence for the ability of deaf adolescents (age: 11–16 years) both to empathize with and to attribute mental states to the protagonist in the picture-sequencing stories. These data thus support the idea of the existence of a well-developed ToM in the deaf adolescents. The results of better performance by the deaf adolescents than their hearing peers also provide evidence for the assumption, based upon studies on short-term memory, that deaf and hearing individuals may possess qualitatively different ways of organizing their experiences as a consequence of the lack of oral-aural experience in deaf people. Accordingly, an inference drawn from numerous studies on the memory abilities of deaf and hearing people is that, perhaps as an effect of an absence or im-

foundly deaf children (age: 6–10 years). The results also appear to support Russell et al. (1998), in that the development of ToM abilities in the majority of deaf children seems to be age-related with the older deaf participants performing significantly superior than the younger deaf children. The revelation of significant and nonsignificant findings raises a number of implications and possible explanations.

The finding that the older deaf participants and the older hearing participants performed similarly in the sequencing of the picture-stories suggests that each group of participants possessed an understanding of physical causality, which is further supported by the lack of any difference in their performance in the mechanical condition. In the behavioral and intentional condition, the deaf adolescents used a significantly greater proportion of mental state expressions in their narratives, whereas the hearing adolescents notably used more descriptive utterances on the same stories. There was no significant difference on the social judgment test.

The results from our research provide evidence for the ability of deaf adolescents (age: 11–16 years) both to empathize with and to attribute mental states to the protagonist in the picture-sequencing stories. These data thus support the idea of the existence of a well-developed ToM in the deaf adolescents. The results of better performance by the deaf adolescents than their hearing peers also provide evidence for the assumption, based upon studies on short-term memory, that deaf and hearing individuals may possess qualitatively different ways of organizing their experiences as a consequence of the lack of oral-aural experience in deaf people. Accordingly, an inference drawn from numerous studies on the memory abilities of deaf and hearing people is that, perhaps as an effect of an absence or im-
pairment in oral-aural experience, deaf people develop a high reliance on visual-spatial short-term memory codes (cf. Whorf, 1956). Studies by O’Connor and Hermelin (1972, 1973a, 1973b, 1976), Hermelin and O’Connor (1973, 1975), and Belmont, Karchmer, and Bourg (1983) all demonstrated that, in comparison with their hearing peers, deaf children depend more heavily on visuo-spatial codes than on temporal-sequential codes. Consequently, in the picture-sequencing test, the deaf adolescents, perhaps, relied on the facial expressions of the protagonist and the contextual cues in the pictures to gain information about what was happening; in contrast, the hearing adolescents needed auditory cues as well as visual attributes to gain the same extent of understanding of the mental state of the protagonist.

However, an alternative explanation for the results can be provided, namely that they arose primarily because the experimenter who carried out the study is herself profoundly deaf. Anecdotally, many deaf children and adolescents become uninhibited and feel freer to express themselves when meeting a deaf rather than a hearing adult. The deaf form a minority group and, traditionally, deaf children and adolescents rarely meet deaf adults, with the exception of approximately 10% of the deaf population born into deaf families. It is worth mentioning that on first entering the school for the deaf to meet the teachers and pupils, the experimenter was treated with caution from most pupils, who assumed that she was “hearing” and on learning that she was, in fact, profoundly deaf “like them,” the attitudes of many pupils changed to that of pleasant surprise. There was an immediate loss of wariness and suspicion. The pupils expressed warmth and motivation in the tests, and a strong impression was gained that if the experimenter had been hearing, the attitude of the pupils would have been quite different.

In contrast, the hearing pupils had not met a deaf person prior to the research conducted in their school. For the majority of the pupils, it was a new and unsettling experience to communicate to a deaf person through an interpreter. Although care was taken to spend time preceding the administration of each test to ensure that the individual child was at ease and understood that there was no right or wrong responses in the tests; the presence of a deaf experimenter and an interpreter may have had an inhibiting influence on the responses made by the majority of the pupils. This leads to the impression that if the experimenter had been hearing, the hearing participants would have responded differently and improved their concentration on the test. This explanation further leads to the question of whether the same happened with deaf participants faced with a hearing experimenter in previous research.

There is no indication that the investigators, in the previously cited studies on the deaf, were themselves deaf signers. Russell et al. (1998) had a hearing experimenter, a teacher for the deaf, who signed the story in their study. Peterson and Siegal (1995, 1999) employed an interpreter in their tests on the false-belief task and tests requiring an understanding of others’ mental states, in deaf children. The studies cited in this research, we believe, involved either interpreters or oral methods of communication. As for the exception of the study by Marschark et al. (in press), the experimenters did not record the stories told by the deaf and hearing participants; they used videotapes that had already been created. The researchers do not say whether the participants were recorded by a person known to them; thus, the results of their study cannot be discussed in this context.

The idea that the deaf participants in previous experiments may have felt intimidated or inhibited and, therefore, reluctant to express themselves fully could provide an explanation for the poor performance reported in the former studies, excluding Marschark et al. (in press). It also provides a potential reason for the deficit in the performance of the hearing participants in this research. This possible factor merits empirical examination.

In contrast to the picture-sequencing test, the social judgment test failed to reveal any significant difference in the performance of the deaf adolescent participants compared with the performance of the matched hearing participants. The reason may be that the test did not require individual narratives as to what the participant thought of each fictional social scenario; instead, the requirement of the test was to respond by rating each behavior on a given scale. Thus, the partici-
pants did not have to give their personal views and thoughts; instead, the test was based on a fixed response structure in contrast to the picture-sequencing test, which required the participant to disclose thoughts about each picture-story. The nonsignificant difference between the results from the older deaf children and their hearing peers indicates that the deaf adolescents are not impaired in the development of social awareness, as suggested by other researchers (e.g., Couch, 1985; DeCaro & Emerton, 1978), and supports the conclusion of Marschark et al. (1999). The social judgment of the deaf adolescents is both normal and superior to that of young people with Asperger’s syndrome.

Finally, there were the results with the younger deaf participants. In the sequencing task in the mechanical condition, there was no significant difference in the performance of the younger deaf children when compared to the performance of their hearing peers and the older deaf participants. The younger deaf children could be described as good intuitive physicists. These stories did not require the attribution of mental states, but instead they could be clearly understood by reference to physical-causal relations.

In the behavioral and intentional conditions, the younger deaf participants performed significantly worse than the deaf adolescents on the sequencing task; however, the younger hearing children performed similarly to the older hearing participants. There was no significant difference in the sequencing performance of the younger children, with the exception of behavioral 1 condition, the reason for which is unclear. It could be explained by the suggestion that the deaf children lost interest in the task quicker than their hearing peers, since behavioral 1 always ended the task. As described previously, the majority of the deaf children and adolescents expressed pleasure on meeting a deaf adult and the deaf participants wanted to ask the experimenter several questions at the completion of the task. The adolescents understood the importance of the completion of the task, whereas the majority of the younger children expressed desire to converse with the experimenter by the end of the task, which could have resulted in decreased concentration. However, this point needs to be examined further. In the behavioral condition, the younger hearing children used significantly more descriptive terms in their narratives than did the older hearing participants and their deaf peers; whereas there was no significant difference in the narratives of the deaf participants. In the critical intentional condition, there was no significant difference in the narratives of the younger children, whereas the deaf adolescents used a significantly greater proportion of mental state words than the younger deaf children. There was no significant difference in the narratives of the hearing participants on the intentional condition.

The result of no significant difference in the general performance of the younger deaf children and their hearing peers is interesting since the older deaf adolescents performed significantly better than the younger deaf children but there was no difference in the performance of the hearing participants. An explanation for this disparity may be that the younger hearing children had not met a deaf person before the study; for the majority of the children, it appeared to be distracting to see a deaf adult and an interpreter at work. Despite care to establish rapport with each child, the presence of a deaf adult and an interpreter likely resulted in an inhibiting influence on their performance, a repeat of the experience with the older hearing adolescents. This could explain why there was no difference in the performance of the hearing participants.

In contrast, the failure of the younger deaf children to perform similarly to the deaf adolescents in the intentional condition implies poor understanding of how people interact and react socially, using mental state expressions. This result can be viewed as providing some additional support for the Peterson and Siegal (1995) conversational account of the development of ToM in deaf children. Peterson and Siegal proposed that, in the case of deaf children, early deprivation of communication and language can cause delayed understanding of mental states and sharing of emotions. Deaf children born into families that do not communicate openly and spontaneously with their deaf child usually grow up unaware of the naturalness of conversations about mental states, the sharing of thoughts and emotions, which thus impedes the ethical development of ToM in the growing deaf child. Generally, for deaf children born into such families, not until the child starts school and meets older deaf children or deaf peers who have achieved a natural and healthy develop-
ment of ToM, usually through sign language, will he or she finally gain access to the world of mental states, the sharing of thoughts and emotions.

The significant impairment in the performance of the younger deaf children when compared with the performance of the older deaf adolescents indicates that the older children have achieved a healthy and unconstrained development of ToM, whereas the younger children demonstrate some evidence of a deprivation in their development in ToM. The results thus are consistent with Peterson and Siegal’s hypothesis that deprived access to communication about mental states at an early age leads to impairment in the development of ToM. A possible reason for this striking difference is that the older children have improved communication resulting from a longer exposure to sign language, which is being recognized as an effective and natural communication method for prelingually deaf children (e.g., Kyle & Woll, 1985; Marschark, 1993; Sacks, 1989). The adolescents can thus freely share information and thoughts among themselves, whereas the younger children are comparatively restricted in awareness about mental states as a consequence of a deprivation in early communication. Thus, ToM abilities may not be age-related as Russell et al. (1998) propose but instead are linked with language fluency. As mentioned previously, the fluency of BSL varied with the younger deaf children in this study. The majority of the deaf children tested in this study started the school for the deaf with prior deficient exposure to language, signed or spoken. It would be interesting in future studies to investigate whether ToM in deaf children is related to language fluency.

Additionally, as Marschark et al. (in press) emphasized, tests described as measuring ToM may not incorporate the full complexities of ToM abilities. The fact that different results are produced with different tests also needs to be taken into consideration.

The development of ToM, then, appears not to be a function of nonverbal intelligence, so even if the deaf child achieves a high score on the nonverbal intelligence scale, it will not be beneficial in terms of a development in social awareness and ToM.

In conclusion, this research did not provide evidence for the hypothesis, following from previous studies, that deaf adolescents would possess significantly poorer knowledge about social reasoning than hearing peers of the same ages, a conclusion supported by Marschark et al. (in press). The results from this research did present additional support for Peterson and Siegal’s (1995) conversational account of the development of ToM in the younger deaf children. This research emphasizes the importance for the prelingually deaf child to gain full access into the world of social awareness, understanding, and sharing of emotions that leads to a healthy and natural development of a ToM. Without at least one fluent conversational partner in the early and crucial years, the prelingually deaf child shows evidence of an impediment in the course of development of a ToM. It could be said that without a ToM, a person is not completely developed as a human being.
Appendix 1

Mechanical

Behavioral

Intentional
Keith knew that when his baby brother cried, it sometimes meant that a pin in his nappy had opened and was pricking the baby. Keith decided not to bother the baby’s mother and he quickly felt inside the baby’s clothing to see if he could feel an open pin.

STORY 3: In the shop.

David always did his shopping in a shop which had a small sign on the door that said “BARE FEET NOT ALLOWED IN THIS SHOP.”

One summer day, David saw a pretty girl go in the shop and she had no shoes on her feet. The girl looked about 20 years old, the same as David. She had long hair and she was wearing a long old-fashioned dress.

David wanted to tell her about the sign on the shop door but he was afraid to talk to her. Bad things happened to him if he tried to talk to girls he didn’t know.

David thought that he might be able to hide her feet from being seen by the shop owner. He pushed his shop trolley close behind her and followed her around in the shop. Sometimes the girl looked back at him with a cross face.

Quickly the girl pushed her trolley into the fast check-out lane with 12 things in her trolley but the sign in the check-out lane said “ONLY 10 THINGS OR LESS.”

David felt more upset now. He thought that the girl was asking for trouble by breaking another shop rule.

When the check-out person let the girl through without saying anything, David started to relax. David followed the girl out of the shop. Outside she turned around and said to David, “I don’t know why you are following me but clear off or I will call the police!”

STORY 4: The lunch-time nap.

Frank, a young man, found a job looking after people’s gardens. When he worked, Frank would take his lunch with him in a box.

In the summer, at lunch-time, Frank would wash his hands in the water from the hose-pipe and sit in a shady part of the garden to eat his lunch.

He was allowed one hour for his lunch and sometimes he would have a quick nap by curling up behind a bush.

One day he was working in a lady’s garden and it
started to rain at lunch-time. Frank went to her house and knocked on the door, asked if he was allowed to eat inside. The lady said he could come inside. She was busy with her children and Frank decided not to bother her.

He found the bathroom by himself and washed his hands ( ).

After his lunch in the kitchen, he cleaned the crumbs from the table and looked around the house for a place to have his rest ( ).

The carpet in the living room was thick, so he decided to curl up for his nap behind a large chair ( ).

References


