Distinguishing Lies from Jokes: Theory of Mind Deficits and Discourse Interpretation in Right Hemisphere Brain-Damaged Patients

Ellen Winner
Boston College and Harvard Project Zero

Hiram Brownell
Boston College and Aphasia Research Center, Department of Neurology, Boston University School of Medicine

Francesca Happé
Institute of Psychiatry, London, Great Britain

Ari Blum
University of South Florida

and

Donna Pincus
State University of New York at Binghamton

Right-hemisphere brain damaged (RHD) patients and a normal control group were tested for their ability to infer first- and second-order mental states and to understand the communicative intentions underlying ironic jokes and lies. Subjects listened to stories involving a character who had either a true or a false belief about another character’s knowledge. Stories ended either with an ironic joke or a lie by this character. In the joke stories, the speaker knew that the listener knew the truth (a true second-order belief) and did not expect the listener to believe what was said; in the lie stories, the speaker did not know that the listener actually knew the truth (a false second-order belief) and thus did expect the listener to believe what was said. RHD patients performed significantly worse than control subjects on one of this research was supported by the National Institutes of Health grants R01 NS27894 and P01 DC 00102, and the Research Service of the Department of Veterans Affairs. Portions of this research were reported at the TENNET Conference in Montreal, May, 1994.

Address reprint requests to Ellen Winner, Department of Psychology, Boston College, Chestnut Hill, MA 02167.
two measures of second-order belief, which suggests that the ability to make second-order mental state attributions is fragile and unreliable following right-hemisphere damage. RHD patients in addition performed worse than controls when asked to distinguish lies from jokes, confirming their known difficulties with discourse interpretation. For both groups, the ability to distinguish lies from jokes was strongly correlated with two measures of the ability to attribute correctly second-order beliefs. These results suggest that the fragility of RHD patients' understanding of second-order mental states underlies a portion of their difficulties in discourse comprehension, but that the underlying impairment is not restricted to right hemisphere dysfunction.

False statements can be false intentionally or unintentionally. Intentional falsehoods can be uttered to protect oneself (in the case of lies) or to cover up one's embarrassment (in the case of ironic jokes). For example, imagine a situation in which a hockey fan calls in sick at work so that he can go to a hockey game. Unbeknownst to the fan, his employer is at the same game and sees him. The next day at work, the boss asks the employee if he feels better. The employee replies, "Yes, a day of bed rest cured me." In this case, the employee has lied. He believes (alas, wrongly) that his employer does not know that he was at the hockey game, and also believes (wrongly) that his employer will believe what he says.

But now imagine the same situation, except that in this case the employee bumps into his employer at the game. The next day at work, the employer asks the same question, to which the employee makes the same response. In this case, the employee has uttered a self-mocking, ironic joke. He knows that his superior knows he was at the game, and so he tries to make light of the situation in the hopes of not getting into too much trouble. He knows that his employer will not believe what he says.

Both forms of intentional falsehood are common. An inability to distinguish the two would lead to severely impaired communication. Research with children has shown that the ability to distinguish a joke from a lie cannot be reliably demonstrated in children younger than six (Andrews, Rosenblatt, Malkus, Gardner, & Winner, 1986; Winner, Windmüller, Rosenblatt, Bosco, Best, & Gardner, 1987). Research has also shown that children's ability to distinguish a joke from a lie is related to their "theory of mind," that is, their recognition that people have beliefs that may differ from their own, that these beliefs may be false, and that these beliefs, whether true or false, guide people's behavior. More specifically, the ability to distinguish a joke from a lie has been shown to be related to one particular component of the child's theory of mind—the ability to make second-order mental state attributions—attributions about one person's knowledge about another person's knowledge. That is, when children can determine that a speaker does not know that a listener knows (an attribution of a false second-order belief), they interpret the speaker's intentional falsehood as a lie. When they can determine that a speaker does know that a listener knows (an attribution of
a true second-order belief), they interpret the speaker’s intentional falsehood as a joke (Sullivan, Winner, & Hopfield, 1995; Winner & Leekam, 1991). There is, thus, evidence from normal development for a relation between a selective aspect of an intact theory of mind (the ability to attribute second-order mental states) and a specific component of communicative competence (the ability to interpret speakers’ intentionally false utterances).

There is also evidence for this same relationship in individuals with autism. Autism is a developmental disorder of biological origin with a triad of impairments in communication, socialization, and imagination (Happe, 1995). Autistic children have been shown in numerous studies to have impaired theories of mind (Baron-Cohen, 1989; Baron-Cohen, Leslie, & Frith, 1985; Happé, 1994; Leekam & Perner, 1991; Leslie & Thaiss, 1992; Perner, Frith, Leslie, & Leekam, 1989; Tager-Flusberg, 1992) and to have the kinds of communication difficulties that also affect RHD patients (Ozonoff & Miller, 1996; Rumsey & Hanahan, 1990; Weintraub & Mesulam, 1983). And Happé (1993) has reported a strong connection, in autistic subjects, between an impaired ability to attribute mental states and an impaired ability to interpret nonliteral expressions such as irony.

In the present study, we examined adults with right-hemisphere brain damage (RHD) following stroke to determine whether this same relationship obtained. We selected participants with right-hemisphere brain damage because these patients have demonstrated difficulties in communicative competence (e.g., Molloy, Brownell, & Gardner, 1990) and with irony in particular (Brownell, Carroll, Rehak, & Wingfield, 1992; Kaplan, Brownell, Jacobs, & Gardner, 1990). We sought to examine these patients’ difficulty determining whether an intentionally false utterance is a joke or a lie, and, if so, whether such a difficulty is related to an impairment in attributing second-order mental states.

Research in our laboratory had suggested to us that RHD patients have some impairment in attributing, and making judgments based on, belief states. In a pilot study, 9 RHD patients heard three folk tales in which one character held both a false first-order and a false second-order belief. For example, in one story, a lazy son stayed home every day while his mother went out to the fields to earn money. She had told him not to eat up all their food while she was gone. But every day he gained another pound. Eventually the mother became suspicious and, instead of going to the fields, she spied on him through a hole in the floor from the basement. She saw that he was secretly eating food he had hidden. Patients were first asked a first-order false belief question, “Where does the son think the mother is?”

On this fairly simple false first-order belief task, no RHD patient had consistent difficulty. Patients typically stated correctly that the son thought the mother was in the fields, as did normal control patients. This pilot finding that RHD patients were able to attribute false first-order beliefs is consistent with a recent report by Siegal, Carrington, and Radel (1996), who found that
RHD patients have no difficulty inferring how someone with a false first-order belief would act as long as the pragmatic purpose of the question is made unambiguous.

Patients were then asked a second-order false belief question: ‘If you asked the son, ‘Does your mother know that you are eating all of that food?’ what do you think he would say?’ A correct answer, ‘No,‘ would indicate that the subject thinks that the son does not know that the mother knows. Two of the RHD patients had consistent difficulty with this type question following all three folk tales. While two out of nine is a small proportion, we decided to investigate further RHD patients’ apparently variable understanding of mental states.

In one study (Brownell, Pincus, Blum, Rehak, & Winner, 1997), we demonstrated that RHD patients have difficulty judging whether speakers should refer to a third party formally (e.g., Mr. Smith) or informally (e.g., Bob). The information on which this judgment was to be made was whether the speaker knew the person well, and whether the speaker knew that the listener also knew the person well (hence, a second-order belief judgment). And in another study (Happe, Brownell, & Winner, 1997), we found that RHD patients have difficulty making use of false first-order belief judgments (but not physical anomalies) in determining whether a cartoon is funny, and in providing mentalistic (but not physical) inferences to explain people’s behavior.

A difficulty understanding mental states would be consistent with clinical evidence. A failure to recognize when a speaker does or does not know what is in a listener’s mind could, for instance, lead to a confusion between joking and lying, as the opening example about the hockey fan and his employer shows. An inability to recognize when a speaker does or does not know what is in a listener’s mind would also be consistent with clinical observations of RHD patients’ difficulty in carrying on a coherent and focused conversation (Molloy et al., 1990).

The purpose of the present study was two-fold. We sought to determine whether RHD patients have difficulty in distinguishing jokes from lies; and we sought to determine whether any such difficulty is associated with a difficulty in making second-order mental state attributions.

We tested RHD patients’ ability to attribute second-order true beliefs and second-order false beliefs. A true second order belief pattern underlay our ironic jokes (the speaker knows that the listener knows), while a false second-order belief pattern underlay our lies (the speaker does not know that the listener knows). We also investigated RHD patients’ ability to attribute appropriate or inappropriate second-order expectations—the liar’s inappropriate expectation that the listener will believe what is said, vs. the joker’s appropriate expectation that the listener will not believe what is said. By having two measures of second-order mental state understanding, we could determine whether patients were consistent in their second-order responses.
Next, we investigated patients’ ability to determine whether specific utterances were lies or jokes. And finally, we investigated whether a relationship obtained between either of our measures of second-order understanding and the ability to interpret and distinguish jokes from lies.

METHOD

Subjects

Thirteen unilateral RHD right-handed stroke patients were tested at least four months post onset of stroke. (See Appendix 1 for details.) The mean time post onset of stroke varied considerably across patients (mean years, months post onset = 5.6; SD = 5.5; range = 0.7–21.0). The RHD group consisted of seven women and six men (mean age = 59.5 years, SD = 12.2, range 39.9–84.1) who were recruited from rehabilitation settings. Lesion descriptions are based on patients’ medical records. Localization, verified by means of CT or MRI, varied considerably within the right hemisphere. As shown in the Appendix, two patients’ lesions were predominantly striatocapsular, but these individuals’ lesions extended outside of the basal ganglia and, therefore, most likely affected cortical function. (See, for example, Alexander, 1992, and Naeser et al., 1994, for related discussion concerning aphasia resulting from subcortical lesions in the left hemisphere.)

A control group of 6 women and 14 men (mean age = 66.5 years, SD = 8.2, range 40.0–77.3) with no history of significant neurological illness was also tested. Control subjects were recruited from a pool of volunteers maintained by the Boston University School of Medicine Aphasia Research Center.

The mean number of years of education of the RHD group was 14.5 (range = 11–19 years, SD = 2.6); the mean education for the control group was 14.2 (range = 12–18 years, SD = 2.4). A t test showed no difference between the two groups in terms of education (t(31) = .44, p = .66).

The slightly younger average age of the RHD group compared to the control group was shown by a t-test to be marginally significant, t(31) = 1.99, p = .06. However, as described below, correlational analyses between age and scores on our key dependent variables all proved nonsignificant. In addition, there was one considerably older patient in the RHD group. This patient made almost no errors on any tasks and thus raised the means for the RHD patient group. Thus, her advanced age should not be of concern, as she clearly had no obvious memory or cognitive deficits that interfered with her task performance.

Prior to testing, all subjects were interviewed to rule out any significant history of alcohol or drug abuse and previous neurological or psychiatric illness.

Materials

Sixteen short stories (approximately 250 words in length) were created. The stories all described situations in which one person witnesses another person doing something sneaky (e.g., stealing money). In the lie stories, the “wrong-doer” does not realize that he/she has been seen (as in the example of the man at the hockey game spotted by his boss). The wrong-doer then utters a lie to the witness in order to conceal the sneaky act. In the joke stories, the wrong-doer realizes that he/she has been seen and utters an ironic joke in order to cover up his/her embarrassment at being caught. The only structural difference between the lie and joke stories was in whether the wrong-doer knows that the listener was also a witness and thus knows the truth. In the lie stories, the wrong-doer does not know that the listener knows the truth. In the joke stories, the wrong-doer knows that the listener knows the truth. The final utterances were spoken in neutral intonation so that intonation alone could...
Table 1

Sample Lie Story: Oscar and the Hockey Game

Oscar did not want to go to work. He wanted to go to a hockey game with his buddies. So Oscar called in sick at work.

Fact Question: Was Oscar really sick? Yes or no?

At the game, Oscar was having so much fun that he didn’t notice his boss sitting a few rows behind him. The boss recognized Oscar.

First-Order Belief Question: Did Oscar’s boss realize that Oscar was not at home sick? Yes or no?

Oscar did not see his boss at all throughout the game. Oscar’s buddy asked him during the game, “Hey, does your boss know you are at this game?”

Second-Order Belief Question: What do you think Oscar told him?

a. “Yes, my boss knows that I am at the game.”
b. “No, my boss does not know that I am at the game.”

Second-Order Follow up Question: Did Oscar think that what he said was really true? Yes or no?

The boss decided not to let Oscar know that he had seen him at the game. He decided he would talk to Oscar the next day in private.

The next day at work, the boss called Oscar into his office. He asked Oscar, “Did you get a lot of rest yesterday?”

Oscar replied, “Yes, and that day of bed rest cured me.”

Second-Order Expectation Question: When Oscar said that to his boss, did he think his boss would believe him? Yes or no?

Interpretation Question: When Oscar said, “Yes, and that day of bed rest cured me,” was he: a. lying to avoid getting caught, or b. joking to cover up his embarrassment?

not serve as a clue as to whether the utterance was a lie or a joke. One of the joke stories had to be deleted from the analysis due to an error in how the item was constructed. Thus, there were eight lie and seven joke stories.

Table 1 presents the lie story about the employee who sneaked off to a hockey game. Table 2 presents a sample joke story, in which Betty tells her husband Jack not to eat the brownies she has just baked because he is on a strict diet. Jack eats the brownies when he thinks (wrongly) that his wife is not looking. She confronts him and catches him in the act. When she asks him later if he was having a hard time sticking to his diet, Jack jokes: “I haven’t eaten anything fattening all day.” Since intonation did not vary between lies and jokes, the only variable that determined whether the statement was to be taken as a joke or a lie was the speaker’s second-order belief (true vs. false) about the listener’s knowledge of the truth.

Procedure

Subjects were tested in two 45 minute sessions within the same month. In each session, subjects heard four lie and either three or four joke stories (stories presented in a pseudo-random order). They were given a written version of the stories, in large type, and were asked to read along silently while listening to the tape. They were told that the tape would be stopped periodically and questions posed and that they could look back and reread parts of the story whenever they wished. Six questions were presented on tape. After each question, the tape was stopped and the answer was recorded. Tables 1–2 show the precise location of each question. These questions are discussed below, using the wording of the questions from the Brownie Story in Table 2 as illustration:

Fact Question: Did Jack eat some brownies? Yes or no?
TABLE 2
Sample Joke Story: Jack and the Brownies

Betty baked some brownies for the church bake sale. She told her husband Jack not to eat a single one because he was on a strict diet. Then she went out to the store.

While she was gone, her husband’s friend came over. Jack was hungry and couldn’t stick to his diet. When his friend left to go to the bathroom, Jack started eating the brownies.

**Fact Question:** Did Jack eat some brownies? Yes or no?

Meanwhile, Betty had forgotten something and came back home. Just as she was about to open the door, she saw Jack through the kitchen window, biting into a brownie.

**First-Order Belief Question:** Did Betty realize that Jack was eating a brownie? Yes or no?

Betty walked into the kitchen. She looked angrily at Jack as he was chewing and holding a half-eaten brownie in his hand. Betty walked out of the room. Jack’s friend returned from the bathroom and asked Jack, “Hey, does Betty know that you are breaking your diet?”

**Second-Order Belief Question:** What do you think Jack told his friend?

a. “Yes, Betty knows that I am breaking my diet.”

b. “No, Betty does not know that I am breaking my diet.”

**Second-Order Follow-up Question:** Did Jack think that what he told his friend was really true? Yes or no?

Betty came back into the kitchen. She asked Jack, “Are you having a hard time sticking to your diet?” Jack replied, “I haven’t eaten anything fattening all day.”

**Second-Order Expectation Question:** When Jack said that to Betty, did he think that Betty would believe him? Yes or no?

**Interpretation Question:** When Jack said, “I haven’t eaten anything fattening all day,” was he: a. lying to avoid getting caught, or b. joking to cover up his embarrassment?

This question was asked to insure that subjects retained enough information to realize that Jack’s final statement was false.

**First-Order Belief Question:** Did Betty realize that Jack was eating a brownie? Yes or no?

This question was asked to insure that subjects realized that Betty knew the truth. This “first-order” belief question required participants to describe a person’s belief about the world, in contrast to the second-order belief question (below), which required participants to describe one person’s belief about another person’s beliefs. Because of the way the stories were constructed, the correct first-order attribution was always a true first-order belief, since in both lie and joke stories, the speaker knew the truth. As mentioned, this was necessary so that the only difference between lie and joke stories was in terms of the speaker’s second-order belief—whether the speaker knew that the listener knew.

**Second-Order True or False Belief Question:** Jack’s friend returned from the bathroom and asked Jack, “Hey, does Betty know that you are breaking your diet?” What do you think Jack told his friend?

a. “Yes, Betty knows that I am breaking my diet;” or

b. “No, Betty does not know that I am breaking my diet.”

Jack’s friend had been absent when Betty confronted Jack, and so could not know that Betty knew, and could not know that Jack knew that Betty knew. This question was called a “second-order” belief question because it required the subject to indicate what Jack believed about his wife’s knowledge. We did not phrase the question directly (Did Jack think his wife knew/did not know that he was breaking his diet) because this form is doubly embedded and thus is difficult to process (Sullivan, Zaitchik, & Tager-Flusberg, 1994).

The correct answer to the second-order belief question was “Yes” for the joke stories and
“No” for the lie stories. However, a subject with an intact understanding of second-order belief might predict that the speaker in a lie story would lie to cover up when asked the Second-Order Belief Question. For instance, one might predict that the hockey fan would lie out of embarrassment and say that his employer knew he was at the game. In order to capture such a response as potentially indicative of understanding second-order belief, we asked a second-order belief follow up question.

Second-Order Belief Follow-Up Question: Did (name of speaker) think that what he said was really true?

In those rare instances in which subjects responded incorrectly to the Second-Order Belief Question but responded to the Second-Order Belief Follow Up question by saying that the speaker did not think what he/she said was true, we gave credit for understanding second-order belief.

After the final utterance, we asked two further questions.

Second-Order Expectation Question: When Jack said that to his wife, did he think that his wife would believe him? Yes or no?

This question assessed subjects’ reasoning about the speaker’s second-order mental state after they had heard the final utterance.

Interpretation Question: When Jack said, “I haven’t eaten anything fattening all day” was he

a: lying to avoid getting caught, or
b: joking to cover up his embarrassment?

This question was asked to determine whether subjects could discriminate between joking and deceptive intentionally false utterances. We included the speaker’s potential motivation in each choice in order to make sure that subjects did not fail because they simply interpreted the meaning of the word “lying” or “joking” differently.

RESULTS

Fact errors. Errors in answering the Fact question would indicate difficulty following the story. Fact errors were never made by the control subjects, and were rare for the RHD patients. Four patients made one fact error each (mean proportion of errors = .03, SD = .40).

First-order true belief errors. Errors in answering the First-Order True Belief question would indicate difficulty making the inference that knowledge follows from perceptual access. That is, failure would show that the subject does not infer that if a person sees X (perceptual access), the person also knows X, or that if a person does not see X, the person does not know X.

RHD patients made more first-order true belief errors (mean = .09, SD = .16) than did control participants (mean = .02, SD = .04). Three of the 20 control subjects made 1 error, and one made 2 errors. In contrast, six of the 13 RHD patients made errors: one RHD patient made 7 errors, one made 6, one made 3, one made 2, and one made 1. However, a chi-square test comparing number of RHD and normal control subjects who made at least one error vs. those who made no errors revealed no difference between groups. This result is not surprising since the first-order questions always concerned true beliefs: one can respond correctly to a true first-order belief question by reporting one’s own belief about the story facts.
Our main interest in this study was in patients’ second-order understanding. We thus went on to analyze responses to the Second-Order Belief, Second-Order Expectation, and Interpretation questions. Responses to these questions were included in the analysis only if a subject had answered both the Fact and the First-Order True Belief questions correctly for the corresponding story. Responses following incorrect responses to either Fact or First-Order True Belief questions were excluded so that we could be confident that problems with the more complex questions were not due to a problem following the story (as shown by an incorrect assessment of the story facts) or to a problem in making a simpler inference (as shown by an incorrect assessment of the speaker’s true belief about the facts). Each subject then received a score for the proportion of correct responses out of all scoreable responses (since not all subjects had the same number of scoreable responses).

Second-order belief and follow-up responses. Subjects were considered to have demonstrated the ability to attribute a second-order belief on a particular item if they showed one of two possible patterns of response to the Second-Order Belief question and its follow-up:

1. Replying correctly to the Second-Order Belief question, and then replying to the Follow-Up question by saying that the speaker thought what he said was true;
2. Replying to the Second-Order belief question incorrectly, but then replying to the follow-up question by saying that the speaker did not think what he said was true. The latter pattern was rare for both subject groups.

RHD patients had a higher mean proportion of second-order belief errors than did control subjects on both lie and joke stories, as shown in Table 3.

<table>
<thead>
<tr>
<th>Subject Group:</th>
<th>RHD</th>
<th>NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>Lie 2-OB Errors</td>
<td>.35 (.29)</td>
<td>.17 (.21)</td>
</tr>
<tr>
<td>Joke 2-OB Errors</td>
<td>.35 (.41)</td>
<td>.13 (.16)</td>
</tr>
<tr>
<td>Lie Expectation Errors</td>
<td>.32 (.34)</td>
<td>.19 (.32)</td>
</tr>
<tr>
<td>Joke Expectation Errors</td>
<td>.22 (.28)</td>
<td>.19 (.26)</td>
</tr>
<tr>
<td>Lie Interpretation Errors</td>
<td>.38 (.42)</td>
<td>.11 (.22)</td>
</tr>
<tr>
<td>Joke Interpretation Errors</td>
<td>.31 (.40)</td>
<td>.22 (.28)</td>
</tr>
</tbody>
</table>

* Any story on which participants failed either the Fact or the First-Order True Belief Question were excluded from this table and subsequent analyses.

* Numbers in parentheses are standard deviations.
A Story Type (Lie, Joke) × Subject Group (RHD, Controls) mixed-design ANOVA showed that RHD patients displayed a significant deficit in attributing second-order beliefs, $F(1, 31) = 5.621$, MSe $= .107$, $p = .02$. There was no effect of Story Type (RHD patients had as much difficulty attributing the true second-order beliefs underlying the jokes as the false ones underlying the lies), and no interaction of Subject Group and Story Type. Although RHD patients performed worse than the normal controls, there was considerable variability in performance within both groups. Not all the RHD patients were impaired, and a few normal subjects also performed poorly.

Second-order expectation responses. RHD patients performed no worse than controls on the Second-Order Expectation question. Although RHD patients performed worse on the lie stories (where the speaker’s second-order expectation was inappropriate) than on the joke stories (where the speaker’s second-order expectation was appropriate), a Story Type × Subject Group ANOVA revealed no effects of Story Type or Subject Group, and no interaction. Thus, after hearing the final utterance, RHD patients typically responded correctly to a question about the speaker’s expectation about the listener’s belief. Mean proportion of errors by each group are shown in Table 3. As with the Second-Order Belief responses, there was considerable variability in performance within each group.

Interpretation responses. RHD patients had a significantly higher mean proportion of errors on the Interpretation Question than did control subjects, for both lie and joke stories, as shown in Table 3. There was a main effect of Subject Group, $F(1, 31) = 7.186$, MSe $= .070$, $p = .01$, but no effect of Story Type, and no interaction.

Relationship between second-order belief and interpretation responses. One should only be able to determine whether a specific utterance is a joke or a lie if one can determine and hold in mind what the particular speaker thinks that the listener knows (the speaker’s second-order belief). Speakers do not normally lie about something when they think their listener is already privy to the truth. To determine whether a relationship between these two abilities is seen in RHD patients, we correlated each subject’s overall proportion of Second-Order Belief and Interpretation errors. As mentioned above, both groups showed considerable variability in performance on both measures. The Pearson $r$s were strong for the control group ($r = .57$, df $= 18$, $p < .01$), the RHD patient group ($r = .71$, df $= 11$, $p < .025$), and for the two groups combined ($r = .71$, df $= 31$, $p < .01$). The relationship between the two skills was similar for both groups: high error scores on second-order belief predicted high error scores on interpretation. The stronger correlation for the RHD patients was most likely due to the greater range of variation within this group. Both subject groups may thus be using their understanding of the speaker’s second-order belief to decide whether the speaker had uttered a joke or a lie. Because RHD patients were on average more impaired
in their ability to attribute second-order beliefs, they were consequently more impaired in their ability to tell a joke from a lie.

**Relationship between second-order expectation and interpretation responses.** To determine whether there is a relationship between an understanding of the speaker’s second-order expectation vis-à-vis the listener’s beliefs and the ability to interpret each utterance, we correlated subjects’ mean proportion of Expectation and Interpretation errors. For both RHD and control subjects, Expectation errors were strongly correlated with Interpretation errors ($r = +.72$, $df = 11$, $p < .025$ for the RHD patients, $r = +.89$, $df = 18$, $p < .01$ for the controls), and for the two groups together ($r = +.79$, $df = 31$, $p < .01$).

**Relationships between demographic variables and second-order belief and interpretation responses.** To assess whether demographic variables such as age, education, time post onset, or gender might affect interpretation of our major results, we performed two parallel series of analyses using proportion of interpretation errors and proportion of Second-Order Belief errors as dependent measures. (Lie and Joke data were combined for these analyses to provide more stable measures.) We performed t-tests for gender for all subjects together ($N = 33$) and for each group separately, bivariate correlations for age and for education for all subjects together and for each group separately, and we performed bivariate regressions for time post onset for just the RHD patients. Neither age, nor gender, nor time post onset (RHD patients only) ever reached even a marginal level of reliability in any of the analyses. Education level never reached even a marginal level of reliability in any of these analyses with the following one exception. For controls subjects alone, there was a marginally significant correlation between education and performance on the interpretation question suggesting that more education was associated with fewer errors, $r = -.38$, $df = 18$, $p < .10$. Finally, we performed separate simultaneous model multiple regressions for interpretation and Second-Order Belief as dependent measures. These analyses included group (RHD, controls), gender, age, and education as the four predictor variables. The most important finding was that the effect of the group variable was reliable in both analyses while effects of gender, age, and education were all not significant in either analysis, which reinforces conclusions based on group differences.

**DISCUSSION**

The results of this study demonstrate, first of all, that the ability to attribute second-order mental states is related to the ability to tell a joke from a lie. This relationship obtained for two different measures of ability to attribute second-order mental states: our Second-Order Belief question, and our Second-Order Expectation question. The results also show that on one of
these measures, RHD patients have a moderate impairment: sometimes they were able to respond correctly to the Second-Order Belief question, and sometimes not. At best, they seem to have a fragile ability to determine what one person thinks another person knows. This difficulty was correlated with their performance in distinguishing jokes from lies, and could thus account for their problems in determining whether a false utterance is, in fact, a joke or a lie.

Although RHD patients performed worse than control subjects on the Second-Order Belief question, they performed no worse than the control subjects on the Second-Order Expectation question. There was no effect of Story Type on responses to the expectation question. However, examination of the scores in Table 3 shows that RHD patients’ relatively good performance on the expectation question may have been carried by the joke stories: RHD patients failed about a third of the Second-Order Belief questions for both lies and jokes, and also failed about a third of the Expectation questions for lies. For jokes, however, RHD patients failed only about a fifth of the Expectation questions, bringing their performance on joke expectation questions close to that of the control subjects. One possible reason for this is that the joker holds an appropriate second-order expectation (the joker rightly does not expect to be believed), while the liar holds an inappropriate second-order expectation (the liar wrongly expects to be believed). Thus, RHD patients may have more difficulty attributing second-order expectations when these expectations are inappropriate. The discrepancy between performance on the two second-order questions may also reflect the variability of RHD patients’ ability to attribute beliefs and expectations to one person about another’s mental state. Sometimes and on some measures they succeed, at other times and on other measures they fail.

The neural pathways involved in theory of mind ability and discourse comprehension in normal individuals are far from established. The radiological summaries for our patients were sufficient to establish neurological involvement of the right hemisphere in all of our patients, but were not uniformly precise enough for more specific claims. While the right hemisphere is clearly implicated in theory of mind, there is also some reason to suspect that the frontal lobes may be involved in the conceptualization of mental states. (See McDonald, 1993, for a discussion of the difficulty of distinguishing between the roles of the right hemisphere and the frontal lobes, given that most studies are not able to compare right frontal vs. right posterior lesions.) For instance, Damasio (1995) noted a striking similarity between RHD patients and ones with prefrontal damage: both have difficulty making appropriate judgments in social situations. Recent evidence from functional brain imaging has also implicated the frontal lobes in theory of mind, though the studies conflict about whether the activated area is on the left or the right side of the brain. One study showed that the right orbito-frontal region was involved when subjects judged whether words related to the mind (Baron-
Cohen, Ring, Moriarty, Schmitz, Costa, & Ell, 1994); another study found that an area in the left frontal region was activated when subjects processed stories requiring mental state inferences (Fletcher, Happè, Frith, Baker, Dolan, Frackowiak, & Frith, 1995). In the case of autism, a deficit which has been shown to be associated with difficulty in making mentalistic attributions (Baron-Cohen, Tager-Flusberg, & Cohen, 1993), there is as yet little agreement as to the localization of brain damage (e.g., Bauman & Kemper, 1994; Gillberg & Coleman, 1992). The study of the brain basis of the human tendency to see behavior in terms of invisible mental states has only just begun.

Interpretation of the associations between right-sided damage due to stroke, impaired theory of mind, and reduced discourse interpretation ability demonstrated in the present study must also reflect the variability across the control and patient groups. There appear to be individual differences in ability in this area among normal as well as brain-damaged individuals. Not all of our RHD patients were equally impaired, a few were not measurably impaired at all. Somewhat more surprising is that some control subjects consistently performed poorly. An impaired ability to conceptualize others’ mental states may thus be a non-specific marker for various conditions, including but not limited to focal right-hemisphere damage.

An admitted limitation of our study was that we were not able to administer our stories to a left-hemisphere damaged (LHD) control group. Most left-hemisphere patients have some form of aphasia, and the linguistic nature of our task made it impossible for them to comprehend the task. Nonetheless, a comparison between RHD patients and normal controls demonstrates clearly that for RHD patients, the ability to distinguish lies from jokes can be fragile and unreliable. This impairment in communicative competence may well be due to an equally fragile ability to make second-order mental state attributions. We recently administered a simpler, forced-choice, nonlinguistic measure of mental state attribution ability to both RHD and LHD groups (Happè, Brownell, & Winner, 1997). This study showed that the “mentalizing” deficits found in the present study are due to right-hemisphere brain damage in particular, rather than to brain damage in general.
## APPENDIX
Patient Information and Individual Performances

<table>
<thead>
<tr>
<th>Lesion Description (right hemisphere)</th>
<th>Sex</th>
<th>Edu.</th>
<th>Age</th>
<th>Time Post Onset (years)</th>
<th>2nd-Order Beliefaranstic</th>
<th>Expectationaranstic</th>
<th>Interpretationaranstic</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHD Patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>middle cerebral artery distribution (motor and sensory strips), anterior and posterior periventricular white matter, striatum (putamen)</td>
<td>f</td>
<td>16</td>
<td>40</td>
<td>.8</td>
<td>.38</td>
<td>0.00</td>
<td>0.12</td>
</tr>
<tr>
<td>middle cerebral artery distribution (superior division), striatum striatum (putamen), internal capsule, anterior and posterior periventricular white matter, fronto-parietal white matter striatum (caudate out to insula), internal capsule, frontal horn of lateral ventricle, some cortical parietal involvement</td>
<td>m</td>
<td>18</td>
<td>48</td>
<td>8.4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>middle cerebral artery distribution (parietal lobe)</td>
<td>f</td>
<td>12</td>
<td>50</td>
<td>.7</td>
<td>0.00</td>
<td>.83</td>
<td>.62</td>
</tr>
<tr>
<td>middle cerebral artery distribution (parietal lobe)</td>
<td>f</td>
<td>11</td>
<td>54</td>
<td>6.6</td>
<td>.50</td>
<td>.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Distribution</td>
<td>Gender</td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
<td>Median</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------</td>
<td>----</td>
<td>------</td>
<td>------</td>
<td>--------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Middle Cerebral Artery Distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Frontal and Parietal Lobes)</td>
<td></td>
<td>f</td>
<td>16</td>
<td>54</td>
<td>1.4</td>
<td>.38</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m</td>
<td>12</td>
<td>59</td>
<td>6.6</td>
<td>.38</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m</td>
<td>16</td>
<td>61</td>
<td>21.0</td>
<td>.50</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m</td>
<td>19</td>
<td>62</td>
<td>4.2</td>
<td>.50</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>f</td>
<td>12</td>
<td>69</td>
<td>.8</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m</td>
<td>16</td>
<td>70</td>
<td>7.3</td>
<td>.50</td>
<td>.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m</td>
<td>16</td>
<td>74</td>
<td>8.6</td>
<td>.38</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>f</td>
<td>14</td>
<td>84</td>
<td>3.2</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td>14.5</td>
<td>59.5</td>
<td>5.6</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>Standard Deviations</td>
<td></td>
<td></td>
<td>2.6</td>
<td>12.2</td>
<td>5.5</td>
<td>.29</td>
<td>.41</td>
</tr>
<tr>
<td>Lesion Description (right hemisphere)</td>
<td>Sex</td>
<td>Edu.</td>
<td>Age (years)</td>
<td>Post Onset</td>
<td>2nd-Order Belief</td>
<td>Expectation</td>
<td>Interpretation</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>lie</td>
<td>joke</td>
<td>lie</td>
</tr>
<tr>
<td>Control Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m 14 40</td>
<td>m</td>
<td>14</td>
<td>40</td>
<td>—</td>
<td>0.00</td>
<td>.29</td>
<td>0.00</td>
</tr>
<tr>
<td>m 18 54</td>
<td>m</td>
<td>18</td>
<td>54</td>
<td>—</td>
<td>.13</td>
<td>.14</td>
<td>.12</td>
</tr>
<tr>
<td>m 12 60</td>
<td>f</td>
<td>12</td>
<td>60</td>
<td>—</td>
<td>0.00</td>
<td>.14</td>
<td>0.00</td>
</tr>
<tr>
<td>f 16 64</td>
<td>m</td>
<td>16</td>
<td>64</td>
<td>—</td>
<td>0.00</td>
<td>0.00</td>
<td>.12</td>
</tr>
<tr>
<td>m 14 64</td>
<td>m</td>
<td>14</td>
<td>64</td>
<td>—</td>
<td>.25</td>
<td>0.00</td>
<td>.12</td>
</tr>
<tr>
<td>m 12 65</td>
<td>f</td>
<td>12</td>
<td>65</td>
<td>—</td>
<td>.25</td>
<td>.27</td>
<td>.12</td>
</tr>
<tr>
<td>f 16 66</td>
<td>f</td>
<td>16</td>
<td>66</td>
<td>—</td>
<td>.88</td>
<td>.50</td>
<td>1.00</td>
</tr>
<tr>
<td>f 12 66</td>
<td>f</td>
<td>12</td>
<td>66</td>
<td>—</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>m 16 66</td>
<td>f</td>
<td>16</td>
<td>66</td>
<td>—</td>
<td>.13</td>
<td>.14</td>
<td>0.00</td>
</tr>
<tr>
<td>f 12 67</td>
<td>f</td>
<td>12</td>
<td>67</td>
<td>—</td>
<td>0.00</td>
<td>0.00</td>
<td>.12</td>
</tr>
<tr>
<td>m 12 67</td>
<td>f</td>
<td>12</td>
<td>67</td>
<td>—</td>
<td>.33</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>f 12 69</td>
<td>m</td>
<td>12</td>
<td>69</td>
<td>—</td>
<td>.13</td>
<td>.43</td>
<td>.62</td>
</tr>
<tr>
<td>m 12 70</td>
<td>m</td>
<td>12</td>
<td>70</td>
<td>—</td>
<td>0.00</td>
<td>0.00</td>
<td>.38</td>
</tr>
<tr>
<td>m 12 71</td>
<td>m</td>
<td>12</td>
<td>71</td>
<td>—</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>m 18 71</td>
<td>m</td>
<td>18</td>
<td>71</td>
<td>—</td>
<td>.38</td>
<td>.29</td>
<td>0.00</td>
</tr>
<tr>
<td>f 18 72</td>
<td>m</td>
<td>18</td>
<td>72</td>
<td>—</td>
<td>.25</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>m 12 72</td>
<td>f</td>
<td>12</td>
<td>72</td>
<td>—</td>
<td>0.00</td>
<td>.14</td>
<td>0.00</td>
</tr>
<tr>
<td>m 13 74</td>
<td>f</td>
<td>13</td>
<td>74</td>
<td>—</td>
<td>.38</td>
<td>.33</td>
<td>1.00</td>
</tr>
<tr>
<td>m 14 75</td>
<td>m</td>
<td>14</td>
<td>75</td>
<td>—</td>
<td>.13</td>
<td>0.00</td>
<td>.12</td>
</tr>
<tr>
<td>m 18 77</td>
<td>m</td>
<td>18</td>
<td>77</td>
<td>—</td>
<td>.25</td>
<td>0.00</td>
<td>.12</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.2</td>
<td>66.5</td>
<td>.17</td>
</tr>
<tr>
<td>Standard Deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.4</td>
<td>8.2</td>
<td>.21</td>
</tr>
</tbody>
</table>

<sup>a</sup> Scores are proportions of errors.
REFERENCES


