

CATHERINE RATTRAY  
University of Strathclyde, Glasgow  
ANDREW TOLMIE  
Institute of Education University of London

## YOUNG CHILDREN'S DETECTION AND DECODING OF IRONIC INTONATION

Two studies examined 3- and 4-year-olds' ability to follow the mental 'sub-text' of conversations employing ironic intonation. In Study 1, children were asked what a confederate thought was inside a tin, following an exchange in which she saw (joke conditions) or did not see (lie conditions) the contents (a stone) and heard these referred to in neutral or ironic tone as a cake. Study 2 repeated the joke conditions, with the confederate touching the stone. Amongst 4-year-olds, intonation was found to trigger complex assessment of the information available to the confederate, whilst 3-year-olds appeared confused. The data suggest that ability to track the belief implications of conversations is underpinned by substantial improvements in working memory between 3 and 4 years.

*Key words:* intonation, irony, theory of mind, pragmatics, preschool children, conversational skill

### Introduction

Two contrasting positions have emerged from debate about the development of children's theory of mind capabilities. Conceptual *change* accounts point to a radical shift in children's success on standard belief prediction tasks after the fourth birthday (e.g. Astington & Gopnik, 1991; Gopnik & Wellman, 1992), and have proposed endogenous mechanisms that might be responsible for this (e.g. the theory of mind module hypothesized by Baron-Cohen & Ring, 1994). Conceptual *continuity* accounts argue that performance outside of standard tasks indicates development proceeds more incrementally (see e.g. Mitchell, 1996), and have focused increasingly on participation in conversation as a major influence

on progress. In the context of deception, for example, Newton, Reddy, & Bull (2000) report that mothers' diaries of their 3- to 4-year-old children's attempts to deceive others showed these frequently occurred when parents were trying to obtain admissions of guilt. The authors suggest this desire for honesty may have been a catalyst to deception: unless they admitted doing wrong, children could only maintain conversational relevance (cf. Grice, 1975) by giving an altered version of events. Importantly, children recognized the conditions of ignorance required for this strategy to be successful from 3 years (cf. Hogrefe, Wimmer, & Perner, 1986; Baron-Cohen & Swettenham, 1996, on 3-year-olds' awareness of visual data as a source of information), but were poor at determining what alternative account adults would find plausible. The implication is that children only become adept at deception by acquiring a detailed grasp of the knowledge that informs the interpretation of their statements by others over subsequent exchanges.

The significance of such understanding is highlighted by Deleau & Bernard (2003). They presented pre-school and primary age children with standard false belief tasks, tests of conversational awareness (e.g. identification of the addressee of a superficially ambiguously-directed remark), and general measures of language ability. They found that false belief performance was predicted best by conversational awareness, and this relationship became stronger as children grew older. Deleau & Bernard conclude that whilst a capacity to build internal representations is a factor in theory of mind ability, what matters most is cumulative experience of conversation and the interpretative processes that underpin it. These findings are consistent with longstanding reports of relationships between parent-child dialogue and later theory of mind ability (Brown & Dunn, 1991; Deleau, Le Sourn, & Guehenneuc, 2000; Dunn, 1994; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991), and point to the requirement to engage with the 'mental sub-text' that informs utterances as the crucial aspect of parent-child exchanges. This direct experience of unraveling the implicit assumptions central to the elliptical character of everyday conversations, it is argued, fuels development of a flexible and generalized understanding of metarepresentation that is grounded in its actual usage (cf. Deleau *et al.*, 2000; Garfield, Peterson, & Perry, 2001; also Clark & Wilkes-Gibbs, 1986; Lewis, Freeman, Kyriakidon, Maridaki-Kassotaki, & Berridge, 1996; Perner, Ruffman, & Leekham, 1994; Ruffman, Perner, Naito, Parkin, & Clements, 1998; Siegal, 1999a; Siegal & Want, 2003).

At a basic level, one form that developing conversational understanding may take is awareness of the tacit conventions used to manage interaction, and the expectations that participants hold about the content of conversations as a result. Grice's 'co-operative principle' (Grice, 1975; Siegal, 1999a), for instance, calls for conversational contributions to be of clear and truthful relevance to the circumstances and to the contributions that preceded it, so that joint understanding is not unduly strained. This implies that participants are typically focused, mentally,

on the obvious aspects of the here and now. However, as Grice notes, experienced communicators frequently deviate from convention in order to convey additional, often social, information (e.g. referring to a jointly witnessed past event or making an obviously false statement as a means of underlining shared understanding). Thus children have to acquire awareness not just of the basic expectations held by participants in conversations, but also of the meaning conveyed by apparent violations. This presents more of a challenge because such violations are by definition less regular and may therefore demand more flexible use of metarepresentational knowledge. For this reason, efforts to decode them might contribute particularly to the development of more general theory of mind capabilities. Indeed, Siegal (Siegal, 1999a; Siegal & Peterson, 1994) argues that children only pass standard false belief tasks when they can resolve the apparent violations of relevance in experimenters' questions.

This raises three questions: at what age do children begin to show an ability to decode conversational violations; how do they go about doing so; and what does this reveal about their emerging theory of mind capabilities? With regard to the first question, there is evidence that they are presented with relevant experiences from a young age. Ninio & Snow (1996), for instance, report that parents spontaneously attempt to extend children's grasp of how conversations depart from basic convention from 18 months onwards, by introducing increasingly complex references to non-immediate objects. This work also points up the fact that some forms of violation are more common, and have a semi-conventional character. It seems likely therefore that children learn to interpret these first.

Amongst the forms children are most likely to encounter from an early age is the use of marked intonation, i.e. departure from standard speech via the addition of unusual emphasis. Both prosodic and intonational variation are reported to be common in adult speech to infants (McRoberts & Best, 1997), and there is evidence that they are amongst the markers used to highlight other behavior that is 'not normal', such as pretence (Lillard, 2003). Sarcastic or ironic intonation in particular (signaling that the intended meaning is the opposite of the literal, or more generally that the literal truth of a statement is not to be believed) is likely to form part of young children's conversational experience with respect to playful deceptions. Siegal (1999b) argues that adults use nonverbal signals in such contexts to help children identify the accompanying violation of quality, and to indicate that this is not to be taken seriously. Understanding of these intonational forms might therefore be acquired relatively early, and serve as a first illustration of how children attempt to decode more complex conversational signals.

Previous evidence on this is both limited and contradictory, however. Keenan & Quigley (1999) found that sarcastic intonation enabled 6- to 10-year-olds to make significantly more correct judgments of a speaker's intended meaning, and conclude that intonation is the easiest cue to sarcasm for children to grasp. Similarly, Capelli, Nakagawa & Madden (1990) report that 8- to 12-year-olds were

only able to recognize sarcastic intent when intonation was present. In its absence, the discrepant nature of a sarcastic statement (i.e. that its literal content was at variance with circumstances) was sometimes recognized, but it was typically still interpreted as either sincere, or, less commonly, as a consoling lie. In consequence, they go further than Keenan & Quigley, and argue that intonation is a necessary cue to sarcasm for children. This would suggest that it is understood primarily via an association with non-literal intent that perhaps derives from earlier play with adults. Since neither of these studies employed pre-school children, they leave the exact point of onset of sensitivity to sarcastic intonation unclear, but if it is well-established amongst primary age children, it is plausible that it is present from some time prior to this.

In contrast, Ackerman (1982) concludes identification of irony or sarcasm involves two distinct and separate processes, first detection, and then interpretation. He reports that whilst 6-year-olds were aware obvious violations of conversational convention signaled that the speaker had some special intent, they performed at no better than chance level in identifying the actual meaning of intonationally-marked sarcastic statements, unless contextual information supporting a non-literal interpretation was provided straight afterwards. The implication is that even if young children can detect sarcastic or ironic intonation, they are unlikely to be capable of determining what it means, and further, that doing so depends initially at least not on simple association, but on more laborious deciphering and integration of available information. In line with Ackerman's data, Winner & Leekham (1991) found that intonation failed to help 5- to 7-year-olds differentiate between deceptive and ironic statements. Sullivan, Winner, & Hopfield (1995) report similar outcomes with 5- to 8-year-olds, and found children could only detect the attitude underlying ironic statements once they were able accurately to attribute second order intention (i.e. what the speaker wanted the listener to believe). These results suggest that far from sarcastic or ironic intonation serving as a conventional aid to interpretation of non-literal statements, it is only fully recognized once the meaning of these has been deciphered; and that understanding of conversational violations *follows* rather than informs theory of mind capability.

There are various aspects of these studies which call their evidence into question, however, and which may account for the apparent contradictions. Firstly, they all required children to interpret ironic or sarcastic statements in the context of stories, which may have underestimated their abilities. Deleau (1998, 1999) reports that direct participation in a false belief scenario leads to improved performance, as the child only has to engage with a single rather than a double reference system (i.e. experimenter-child *and* story characters). In the present context, use of narratives may entail decoding departures from standard conversational convention in *both* systems before a correct response can be arrived at (cf. Siegal & Peterson, 1994). It is possible therefore that children under the age of 5 years know what to infer from sarcastic or ironic intonation when they encounter it

directly, but cannot apply this understanding to narratives until later. Secondly, the studies where children performed least well (i.e. Sullivan *et al.*, 1995; Winner & Leekham, 1991) also employed the most demanding methodology, since children were only counted as having made correct responses if they classified a deceptive *and* an ironic scenario appropriately.

Data on the understanding of intonation in a more naturalistic setting, using simpler forms of response, would therefore provide a better assessment of a) the extent of younger children's capabilities; b) how far detection and interpretation are separate processes; and c) the way in which interpretation is made in different contexts (i.e. by conventional association or by more deliberate puzzling out of meaning). This paper reports two studies designed to provide such an assessment. The focus of these was on ironic intonation, signaling that a statement should be disbelieved, in the context of 1) a self-revealing lie, and 2) a joke about shared knowledge. Both studies employed the same paradigm, based on the deceptive box task (Perner, Leekham, & Wimmer, 1987). This was used to investigate 3- and 4-year-olds' understanding of the belief implications for a confederate of an experimenter's false statement about the contents of a tin.

Interest centered on whether belief attributions altered with intonation, indicating it had been detected, and if so, whether the direction of shift showed that it had been interpreted appropriately. The confederate's belief was the focus of questioning since it could be addressed more straightforwardly than the speaker's intent (i.e. via questions about actual conclusions, rather than underlying motives). The younger age group was chosen as a baseline sample in the light of claims that children are aware of conventional conversational practice from 3 years onwards (Siegel, 1999a); cf. also Newton *et al.*, 2000), which suggests they should be capable of at least detecting variation in intonation. The choice of 4-year-olds as the comparison group was guided by consistent findings of success on the deceptive box and related tasks after the fourth birthday (Astington & Gopnik, 1991; Mitchell, Robinson, Nye, & Isaacs, 1997; Perner *et al.*, 1987). Children of this age should thus have no intrinsic difficulty in recognizing that the confederate could have a different belief to their own, depending on the information available to them. This made it possible to examine the effects of false statements and intonation without the confounding influence of task effects.

## Study 1

Study 1 examined whether 3- and 4-year-olds are sensitive to ironic intonation as a cue that lies (verbal statements in the absence of visual information) and jokes (here, verbal statements conflicting with visual information) should be disbelieved. Interest centered not on how far performance differed for lies and jokes, but on what these contexts revealed *conjointly* about children's ability. Lies and jokes provided a useful combination because there are important differences be-

tween the situations in which they occur. In the case of typical lie scenarios, the recipient only has verbal information available, whereas in joke scenarios of the kind defined above, they have both verbal and visual. As a result, there are different baseline expectations about the belief consequences of verbal statements. Mitchell *et al.* (1997) report 4- to 6-year-olds are less likely to judge that a listener will believe a message when it contradicts information derived from sight than when the listener has no prior knowledge, though performance is by no means uniform (see also Perner & Davies, 1991).

In the present context, this meant ironic intonation had differing status depending on the situation: in the case of lies, it served as a cue uniquely defining whether or not the verbal statement should be believed, whereas in the case of jokes, it was simply an *additional* piece of information pointing in the same direction as visual data. In this sense, then, decoding of ironic jokes might be expected to be easier. In practice, however, it might be more difficult if the greater amount of information to be assessed is a factor, since here there were effectively three cues to take into account (visual information, verbal content, and intonation), whereas for lies there were only two, visual information being absent. Thus the examination of both lies and jokes provided a test not only of whether 3- and 4-year-olds could decode a basic form of ironic intonation, but also of whether they did so by evaluating it as one piece of information amongst the others available, as implied by Ackerman's (1982) two stage account – and if so, how – or formulaically, via a knowledge of what it conventionally signals (cf. Keenan & Quigley, 1999; Capelli *et al.*, 1990) – in which case there should be no difference between lies and jokes.

Children were required to indicate what a confederate of the experimenter thought was inside a tin containing a stone, which the confederate did (joke conditions) or did not see (lie conditions) prior to the experimenter stating to them that there was a cake in the tin. Within these conditions, this statement was delivered either in neutral fashion, or using ironic intonation, which exaggerated the modulation of pitch and added emphasis to the word “cake” by drawing it out longer (cf. Capelli *et al.*, 1990, who argue this is the form of sarcastic intonation most familiar to children). Given the influence of stereotypic contents on children's responses in previous versions of the deceptive box paradigm (e.g. Wimmer & Hartl, 1991; Saltmarsh & Mitchell, 1998) it should be noted that the appearance of the tin gave no clear indication of its contents, allowing the impact of intonation to be assessed without this intruding.

It was anticipated that in the lie conditions, when intonation was neutral children would say that the confederate thought there was a cake in the tin, in line with the verbal statement. When ironic intonation was present, however, if they were sensitive to its belief implications they should answer that the confederate thought there was neither a cake nor a stone in the tin, given the absence of other information. In the joke conditions, when intonation was neutral it was expected

that children would tend to answer that the confederate thought there was a stone in the tin, given the visual information available to her. With ironic intonation emphasizing the discounting of the verbal statement, they should attribute this belief to the confederate more uniformly. The type and relative scale of effect across lie and joke conditions would serve to indicate whether understanding of intonation took the form of awareness of its conventional meaning; or whether its decoding was dependent upon weighing up the implications of the different sources of information available to the confederate.

## **Method**

### *Design*

A randomized between-subjects design was employed, in which two age groups of participants (3- vs. 4-year-olds) witnessed an exchange between a confederate and an experimenter under one of four conditions, defined by whether or not the confederate saw inside a tin (and thus whether the experimenter's false statement to her about its contents constituted a lie or a joke); and whether the intonation used for this statement was neutral or ironic. Dependent measures were derived from children's responses to questions about whether the confederate had seen inside the tin; knew what was in it; and thought it contained a stone or a cake. A measure of vocabulary was taken to check that the verbal abilities of participants in each condition were comparable.

### *Participants*

The participants were 128 pre-school children attending two nursery schools with a mixed social class intake in Glasgow, Scotland. The sample was divided equally into two age groups. The younger group (36 girls, 28 boys) ranged in age from 3 years 1 month to 4 years (mean age 3.57 years). The older children (31 girls, 33 boys) were aged from 4 years 1 month to 4 years 9 months (mean age 4.35 years). Children were assigned at random in equal numbers from these age groups to each of the four conditions. All children had standard English as the language of the home.

### *Materials*

Vocabulary was assessed using the long form of the British Picture Vocabulary Scale (BPVS) 1<sup>st</sup> Edition (Dunn, Dunn, Whetton & Pintilie, 1982). Whilst superseded by a 2<sup>nd</sup> Edition, this was preferred since it had the advantage of providing more fine-grained scores in the participating age groups, and was only required as a comparative measure. A plain flat rectangular tin, which opened on a hinge was used as the focus for the exchange between the confederate and the experimenter. The tin contained a round, quite flat, smooth stone. Children's responses were recorded using an audio-cassette recorder.

### *Procedure*

Children were taken individually by the female experimenter to a small room off the main nursery area, where they were invited to sit at a rectangular table opposite her, with their back to the door. This was left ajar, to allow a confederate sitting just outside to hear the discourse and so know when to step in. The experimental session consisted of four stages, beginning with the child being administered the BPVS using standard procedures. Throughout this first stage the tin containing the stone was on the table with its lid shut. Once the material for the BPVS had been removed, children were engaged in conversation about its possible contents. They were first asked whether they could guess what was inside, and were then invited to open the tin and look. Children were encouraged to touch the stone, and most spontaneously named it as such. The remainder were prompted so that all participants explicitly recognized the object was a stone before progressing. The tape recorder was then switched on and left running for the rest of the session.

At this point the procedure varied as a function of experimental condition. The exact sequence of events in each condition was as follows:

*Condition 1 – Lie/neutral intonation:* Subsequent to the child naming the stone, the tin was closed. The confederate entered, carrying paper and pencil, ostensibly to take a note of the name of the child. The following discourse took place.

C1) Confederate: *Who is this then?*

E1) Experimenter: *This is [name of child].*

[Confederate writes down name.]

C2) Confederate: *That's a nice tin you've got there* [touching the closed lid of the tin].

E2) Experimenter: *There's a really nice cake in my tin.*

C3) Confederate: *I hope you're going to give [name of child] some.*

The confederate then left the room. The intonation of utterance E2 was kept as neutral as possible, in line with the entire discourse.

*Condition 2 – Lie/ironic intonation:* This condition was identical to Condition 1, except that utterance E2 alone was spoken with ironic intonation.

*Condition 3 – Joke/neutral intonation:* After the child named the stone, the tin was left open. The discourse which followed the entry of the confederate was the same as for Condition 1, the sole difference being that the confederate looked inside the tin before partially closing it so that she could touch the lid while stating utterance C2. The confederate then left the tin open as she had found it. As in Condition 1, the intonation of the entire sequence was neutral.

*Condition 4 – Joke/ironic intonation:* This was the same as Condition 3, except that utterance E2 was given ironic intonation in the same way as in Condition 2.

After the confederate left the room, children were asked four questions:

1. Does she know what is in the tin?
2. Does she think there is a stone in the tin?



3. Does she think there is a cake in the tin?
4. Did she see what was in the tin?

On occasion, children responded to Question 1 with the word "stone". In such cases it was unclear whether they were referring to their own knowledge or to that of the confederate, so a supplementary question, "*Is that what she thinks is in the tin?*" was asked. The order of Questions 2 and 3 was counterbalanced within conditions. Once they had responded to all four questions, children were thanked for their participation and re-united with their class. The session lasted approximately ten minutes in total.

### *Scoring and reliability*

*BPVS responses.* Individual test records for the BPVS were scored in the standard way and a raw score (i.e. of absolute vocabulary level) derived for each child.

*Scoring of dependent measures.* Responses to Questions 1 ("Does she know what is in the tin?") and 4 ("Did she see what was in the tin?") were scored separately according to whether children gave a <yes> or <no> response. Question 1b ("*Is that what she thinks is in the tin?*") was only used to determine which category children's responses should be put under when they answered <stone> to Question 1. Responses to Questions 2 ("Does she think there is a stone in the tin?") and 3 ("Does she think there is a cake in the tin?") were scored according to the pattern of answer across these questions. Thus, there were four possible categories of response: <stone-yes/cake-no>, <stone-no/cake-yes>, <stone-yes/cake-yes> and <stone-no/cake-no>.

*Reliability of the manipulation of intonation.* Differences between conditions in the intonation applied to E2 were checked to ensure that a) these were sufficiently obvious to be reliably detected; and b) the interpretation placed upon the neutral and ironic versions by experienced communicators was as intended. Three adults, blind to the purpose and design of the study listened to tape recordings of E2 taken from three trials in each condition, selected at random, and presented in fixed random sequence without their surrounding context. Raters answered two questions for each recording: 1) "What did I say was in the tin?"; and 2) "Did I mean it?". Question 1 was included to check on the audibility of the statements; in all cases raters answered this question correctly. Responses to Question 2 for each recording were collated with the intended meaning, and analyzed for overall agreement rate. There was a high level of agreement ( $Kappa = 0.89$ ,  $z = 7.38$ ,  $p < 0.01$ ; percentage agreement = 92%), with raters exhibiting significant consensus both with each other and with the intended meaning. Only one of the 12 trials, involving ironic intonation, produced any disagreement. Whilst this meant that detection and interpretation of ironic intonation was less than perfect amongst the adult raters, they did not have the advantage the children had of hearing the preceding dialogue and the contrast the ironic statement made to this.

## Results

Prior to analysis of the dependent measures, two-way (condition x age group) ANOVAs were carried out on children's age in months and their BPVS raw scores. These analyses showed main effects of age group ( $F(1,120) = 411.24, p < 0.001$ ;  $F(1,120) = 66.64, p < 0.001$  respectively), but confirmed there were no differences between conditions nor interactions between conditions and age group. Standardized vocabulary scores for the sample varied from 78 to 125, with a mean of 96.41 and a standard deviation of 9.02, indicating that children were for the most part in the average to low average range.

### *Awareness of information available to confederate*

For children to appreciate the belief implications of the scenario they witnessed, it was important they recognized whether or not the confederate had seen inside the tin, and the state of knowledge or ignorance of its contents that would follow from this. As Table 1 shows, 30 (100% of valid responses) of the 4-year-olds and 29 (91%) of the 3-year-olds correctly judged that the confederate had not seen the contents of the tin in the lie conditions. In contrast, in the joke conditions 28 (88%) of the 4-year-olds and 29 (94%) of the 3-year-olds said she *had* seen inside the tin. Hierarchical log linear analysis using backward elimination procedures with four factors (see response – <yes> vs. <no>; age group – 3- vs. 4-year-olds; condition – lie vs. joke; and intonation – neutral vs. ironic) showed frequencies could be accounted for solely by the interaction between see response and lie vs. joke condition (for adequacy of fit of this model, likelihood ratio  $\chi^2 = 8.18, df = 12, p = 0.77^1$ ).

As can be seen from Table 1, however, children were less clear about the implications of seeing inside the tin for knowledge of its contents, with 20% responding <see-yes> and <know-no>. Hierarchical log linear analysis on know responses, with other factors as before, revealed the presence of interactions between know response and age group, and between know response and lie vs. joke condition (for the model containing only these effects, likelihood ratio  $\chi^2 = 8.67, df = 10, p = 0.56$ ). As Table 1 shows, the 3-year-olds were in general divided on the question of whether the confederate knew what was in the tin, whereas the 4-year-olds tended to favor <no> responses, and this was true for both the lie and joke conditions. Conversely, <yes> responses were more common for the joke conditions than for the lie condi-

---

<sup>1</sup> In hierarchical log-linear analysis, the object is to test how well models of the data fit actual observations. This is done via a process of backward elimination, starting with a model that includes all possible main and interaction effects, and then examining whether excluding each effect leads to a significant departure from adequate fit to the observed data. Effects whose exclusion does not result in such a departure are dropped from consideration. This process continues until no remaining effects can be excluded without impairing fit. These effects constitute the final model, whose adequacy of fit is reported in terms of its associated likelihood-ratio  $\chi^2$ . This value is always non-significant (i.e. greater than .05), since it reflects degree of fit, not departure from it.

Table 1. Frequency of responses given to Study 1 Questions 4 ("Did she see what was in the tin?") and 1 ("Does she know what is in the tin?") by condition and age group (excluding don't know responses).

	See response		Know response	
	Yes	No	Yes	No
3-year-olds				
1. Lie/neutral intonation	1	15	9	7
2. Lie/ironic intonation	2	14	5	10
3. Joke/neutral intonation	14	2	11	4
4. Joke/ironic intonation	15	0	9	7
4-year-olds				
1. Lie/neutral intonation	0	16	1	15
2. Lie/ironic intonation	0	14	1	13
3. Joke/neutral intonation	14	2	8	7
4. Joke/ironic intonation	14	2	4	11

tions, and this was true for both age groups. Overall, then, among the 3-year-olds, understanding of the relationship between sight and knowledge was an uncertain one, though when the contents of the tin were seen, they were more likely to conclude that this would result in knowledge. The 4-year-olds in contrast, were clear that *not* seeing resulted in ignorance, but were more ambivalent about the effects of seeing, especially in the ironic joke condition.

#### *Effects of intonation on perceptions of the confederate's belief*

In the neutral lie condition, <stone-no/cake-yes> responses to the questions about what the confederate thought was in the tin would indicate children thought the confederate believed the experimenter's false statement, whereas <stone-yes/cake-no> would suggest they failed to distinguish the confederate's knowledge from their own. This was also true for the ironic lie condition, whilst <stone-no/cake-no> would indicate sensitivity to the implication of intonation, coupled with recognition of the absence of other information to inform the confederate's belief. In the neutral joke condition, <stone-yes/cake-no> would suggest the child favored direct observation over the verbal statement as the source of the confederate's belief, whilst <stone-no/cake-yes> would indicate the opposite. In the ironic joke condition, <stone-yes/cake-no> would in addition be consistent with the implication of intonation. Overall, then, if children were sensitive to the conventional implications of ironic intonation, this should have been revealed by a greater incidence of <stone-no/cake-no> and <stone-yes/cake-no> responses in the ironic lie and ironic joke conditions respectively, relative to the relevant neutral intonation conditions.

Table 2. Frequency of responses given to Study 1 Questions 2 (“Does she think there is a stone in the tin?”) and 3 (“Does she think there is a cake in the tin?”) by condition and age group.

	Pattern of response			
	stone-no cake-yes	stone-yes cake-no	stone-yes cake-yes	stone-no cake-no
3-year-olds				
1. Lie/neutral intonation	12	1	3	0
2. Lie/ironic intonation	3	2	9	2
3. Joke/neutral intonation	2	3	11	0
4. Joke/ironic intonation	2	8	4	2
4-year-olds				
1. Lie/neutral intonation	11	3	1	1
2. Lie/ironic intonation	5	2	2	7
3. Joke/neutral intonation	4	8	3	1
4. Joke/ironic intonation	8	5	1	2

Table 2 shows that the observed response pattern was more complex than this. Hierarchical log linear analysis, with response pattern, age group, lie vs. joke condition and neutral vs. ironic intonation as factors, identified a significant 3-way interaction between response pattern, lie vs. joke condition, and neutral vs. ironic intonation, and a 2-way interaction between response pattern and age group (for this model, likelihood ratio  $\chi^2 = 9.36$ ,  $df = 12$ ,  $p = 0.67$ ). To take the latter effect first, it can be seen from Table 2 that the 3- and 4-year-olds differed most in the incidence of <stone-yes/cake-yes> responses compared to other patterns (27 vs. 37 for the 3-year-olds, compared to 7 vs. 57 for the 4-year-olds; Fisher Exact < 0.01, one-sided). The 3-year-olds also made fewer <stone-no/cake-no> responses (4 vs. 60 as opposed to 11 vs. 53; Fisher Exact = 0.05, one-sided). Cutting across these trends were systematic shifts in response according to lie vs. joke condition and neutral vs. ironic intonation. For the neutral lie condition, the modal response (23 vs. 9) was <stone-no/cake-yes>, indicating that children in general thought the confederate believed the experimenter’s statement under these circumstances. When intonation was introduced, however, there was a shift away from <stone-no/cake-yes> towards <stone-no/cake-no>, as anticipated (23 vs. 1 for the neutral condition, compared to 8 vs. 9 for the ironic; Fisher Exact < 0.01, one-sided), signaling that some children at least were sensitive to the belief implications of the intonation. There was also a shift towards <stone-yes/cake-yes>, though (23 vs. 4 for the neutral condition, against 8 vs. 11 for the ironic; Fisher Exact < 0.01, one-sided). These were exactly the patterns on which the 3- and 4-year-olds most differed, and as can be seen, the two age groups were almost completely differentiated in which shift they exhibited.

For the neutral joke condition, <stone-yes/cake-yes> was the most prevalent response, but this was nearly matched in frequency by <stone-yes/cake-no> (14 vs. 11 respectively), consistent with a number of children recognizing that the confederate's beliefs would be informed by what they saw. It was the 3-year-olds who were again most likely to give <stone-yes/cake-yes> responses, and the 4-year-olds <stone-yes/cake-no>. With intonation, <stone-yes/cake-yes> responses dropped, and <stone yes/cake-no> responses increased slightly, as predicted, though this shift was not quite significant (14 vs. 11 for the neutral condition, compared with 5 vs. 13 for the ironic; Fisher Exact = 0.06, one-sided). Surprisingly, though, <stone-no/cake-yes> responses also increased slightly as <stone-yes/cake-yes> responses dropped, suggesting that the effect of intonation was to make some children consider the statement *more* likely to be believed (14 vs. 6 for the neutral condition, against 5 vs. 10 for the ironic; Fisher Exact = 0.07, two-sided). More surprisingly still, it was the 4-year-olds who showed this trend, and the 3-year-olds who exhibited the *expected* shift towards <stone-yes/cake-no>.

Across the neutral lie, ironic lie and neutral joke conditions then, the pattern of responses indicated that whilst both age groups appreciated the belief implications of a verbal statement in the absence of other knowledge, only the 4-year-olds recognized in any numbers the belief implications of ironic intonation accompanying such a statement, or the likely interpretation of a false statement combined with visual information. The 3-year-olds appeared confused under these conditions, giving large numbers of <stone-yes/cake-yes> responses. Flying in the face of these trends, when it came to the ironic joke condition, the 3-year-olds gave fewer <stone-yes/cake-yes> responses and <stone-yes/cake-no> became modal, suggesting that here intonation helped them decode the belief implication. It was the 4-year-olds who gave the apparently *less* advanced <stone-no/cake-yes> responses – in line with their previously noted tendency when intonation was present to doubt that knowledge of the tin's contents would result from seeing inside it.

## Discussion

These results are, on the face of it, perplexing. On the one hand, there was clear evidence contra Winner & Leekham (1991; see also Winner, 1988) and Sullivan *et al.* (1995) that intonation affected belief attribution in both 3- and 4-year-olds, indicating that children of this age are at least capable of detecting it. This was despite the fact that there was no sign that the verbal abilities of the present sample were especially precocious. These variations in belief attribution cannot be held simply to reflect some form of difficulty with the basic task: both 3- and 4-year-olds were aware whether or not the confederate saw the contents of the tin, and whilst they were seemingly less sure about the effects of this on knowledge, neither age group had any problem attributing a different belief from their own to the confederate when she only had the experimenter's statement to inform her.

The effects of intonation were not uniform, however, supporting Ackerman's (1982) hypothesis that, for children, interpretation is a distinct process from detection, subject to contextual fluctuation. In the ironic lie condition, for example, nearly half the 4-year-olds made an interpretation in line with convention, whereas in the ironic joke condition half seemed to treat intonation as a cue that the statement was *more* likely to be believed. It is the nature of these variations in interpretation that is puzzling. The fact that the 4-year-olds' responses were more consistent with expectation in the lie conditions, where there were fewer cues to consider, suggests that informational load influenced belief attribution. The apparent confusion of the 3-year-olds in the ironic lie and neutral joke conditions also indicates this, and indeed their performance in the former calls into question whether they were capable of anything beyond detection of intonation. A straightforward account in terms of informational load is ruled out, however, by the 3-year-olds' apparent ability to interpret intonation appropriately (in conventional terms) in the ironic joke condition, where the number of cues to attend to was greatest. Since it is implausible that the 3-year-olds should show their most advanced performance at the point where the 4-year-olds seemed to do worst, a more complex set of effects is implied.

One potential clue to the nature of the 4-year-olds' performance in the ironic joke condition might lie in the shift they exhibited here in the perceived equation between sight and knowledge. Whilst a majority (albeit a slender one) thought that knowledge would follow from sight in the neutral joke condition, there was a clear majority *against* in the ironic joke condition. This suggests the 4-year-olds felt there was some ambiguity about the perceptual qualities of the stone that was especially highlighted by the presence of intonation. In other words, they behaved as if they thought it was this ambiguity to which attention was being drawn, rather than the veracity of the verbal statement. This appeared to be sufficient to tip the balance towards thinking in turn that the confederate would not know the object was a stone simply from seeing it, and she would therefore be more likely to believe it *was* a cake. Consistent with this, <stone-no/cake-yes> was the modal response amongst those 4-year-olds in the ironic joke condition who answered <see-yes/know-no>.

In as much as it implies that the 4-year-olds were evaluating different elements of information in piecemeal fashion, the same account might explain the responses of the 3-year-olds, although in a different way. It has already been noted that the 3-year-olds seemed to suffer confusion in the ironic lie and neutral joke conditions, where two pieces of information had to be assessed. One interpretation of this confusion is that they responded by trying to fit in all the information they had, rather than thinking through and resolving the competing belief implications, hence their <stone-yes/cake-yes> responses. It is possible, then, that when the informational load was even higher, in the ironic joke condition, the demands of keeping track of each element became unmanageable, and they simply reverted

to ascribing their own knowledge of the contents of the tin to the confederate, giving rise to a response that was apparently consistent with the conventional interpretation of ironic intonation.

## Study 2

One test of this account of response patterns in the Study 1 joke conditions would be to modify the scenario so that participants witnessed the confederate touching the stone as well as looking at it. Under these circumstances, ironic intonation should still lead 4-year-olds to think the stone's appearance is ambiguous, but the disambiguating information provided by touch should also cause them to think the confederate will in fact conclude it is a stone. Amongst 3-year-olds, if management of informational load is the primary influence on performance, adding a further element should result in an even stronger tendency to disregard all cues and attribute personal knowledge to the confederate.

In order to assess these predictions, Study 2 replicated the joke conditions from Study 1 in both the original and an alternative version, in which the confederate touched the stone prior to the experimenter referring to it as a cake using either neutral or ironic intonation. Children aged 3 and 4 years viewed one of these scenarios, and were then asked the same four questions as in Study 1, plus a final question about whether the stone looked like a cake, to establish whether it was indeed perceived to be ambiguous. It was anticipated that 4-year-olds would judge the stone to look more like a cake when the experimenter used ironic intonation, but to think the confederate would believe it *was* a cake only if she did not touch it. The addition of touch was expected to lead 3-year-olds to make even more <stone-yes/cake-no> responses in the ironic joke condition than in the original version.

## Method

### *Design*

The design was as for Study 1, except that the four conditions were now defined by whether or not the experimenter's joking statement that the tin contained a cake was given neutral or ironic intonation; and whether or not the confederate touched the stone. The dependent measures were also as before, with the addition of children's responses as to whether the stone looked like a cake. A measure of vocabulary was taken as previously to check that participants' verbal abilities were approximately the same in each condition.

### *Participants*

Participants were 64 3-year-olds (33 girls and 31 boys; range 3 years 0 months to 4 years, with a mean age of 3.61 years) and 64 4-year-olds (23 girls and 41

boys; range 4 years 1 month to 4 years 11 months, with a mean age of 4.36 years) drawn from three Glasgow nurseries with a mainly middle class intake. Sixteen children from each age group were assigned at random to each of the four conditions. As for Study 1, all of the children had standard English as the language of the home and all had parental consent to participate.

### *Materials and procedure*

The materials were identical to those used in Study 1. As before, the test session took place in a room adjacent to the child's classroom, and consisted of four stages. Stages 1 and 2 involved administration of the BPVS and discussion with the child about the contents of the tin as in Study 1. The only variations from Study 1 were at the point when the confederate entered the room, and in the questions that the child was asked. On this occasion, the tin was left open at the start of Stage 3 in all four conditions. The ensuing sequence of events and dialogue between confederate and experimenter was as in the Study 1 joke conditions, except that in two of the four conditions, the confederate touched the stone at the point when she looked inside the tin, before she partially closed the lid, and utterances C2 and E2 (the test statement) were made. Within the touch and no touch conditions, E2 was stated, as previously, with neutral intonation for half the children, and ironic for the other half. The sequence of questions at Stage 4 was the same as in Study 1, save that children were asked a final, additional question, "Does the stone look like a cake?".

### *Scoring*

Individual test records for the BPVS were scored as before, and raw scores derived for each child. Responses to the original four experimental questions were scored as in Study 1. Answers to the additional question about the stone's appearance were scored according to whether the child responded <yes> or <no>.

## **Results**

Two-way (condition x age group) ANOVAs were carried out on children's age in months and their BPVS raw scores, which revealed differences between age groups ( $F(1,120) = 256.51, p < .001$ ;  $F(1,120) = 10.64, p = 0.001$  respectively), but no differences between conditions, nor interactions between condition and age group. Standardized vocabulary scores ranged from 86 to 149, with a mean of 106.63 and a standard deviation of 11.05, indicating that the sample was primarily in the average to high average range.

### *Awareness of information available to confederate*

Table 3 presents children's responses to Questions 4 ("Did she see what was in the tin?") and 1 ("Does she know what is in the tin?"). As it shows, irrespective of age group, intonation, or whether the confederate touched the stone, children



Table 3. Frequency of responses given to Study 2 Questions 4 ("Did she see what was in the tin?") and 1 ("Does she know what is in the tin?") by condition and age group (excluding don't know responses).

	See response		Know response	
	Yes	No	Yes	No
3-year-olds				
1. Neutral/no touch	12	3	11	4
2. IroniC/no touch	12	3	11	5
3. Neutral/touch	13	3	8	8
4. IroniC/touch	12	4	8	6
4-year-olds				
1. Neutral/no touch	14	2	10	5
2. IroniC/no touch	15	1	5	7
3. Neutral/touch	14	1	8	8
4. IroniC/touch	13	3	8	5

were likely to say that the confederate did see the contents of the tin with 105 children (84% of valid responses) answering <yes> to Question 4 and 20 (16%) answering <no>. These results are very similar to those in the comparable conditions in Study 1. Hierarchical log linear analysis using backward elimination procedures with four factors, see (<yes> vs. <no>), intonation (neutral vs. ironiC), touch (touch vs. no touch) and age group (3- vs. 4-year-olds), confirmed the data could be accounted for simply in terms of a main effect of see response (likelihood ratio  $\chi^2 = 4.25$ ,  $df = 14$ ,  $p = 0.99$ ).

As in Study 1, however, children were less certain about whether seeing inside the tin resulted in the confederate knowing what it contained. Hierarchical log linear analysis on the know responses (<yes> vs. <no>), with other factors as before, found no significant effects (for the null effect model, likelihood ratio  $\chi^2 = 9.83$ ,  $df = 15$ ,  $p = 0.83$ ). There was, however, a marginally significant main effect of know response (change in likelihood ratio  $\chi^2$  when removed from the model = 3.79,  $df = 1$ ,  $p = 0.05$ ), with <yes> responses more likely overall than <no> responses (69 vs. 48). Consistent with this, 3-year-olds in both the neutral and ironiC intonation/no touch conditions, and 4-year-olds in the neutral/no touch condition were all more likely to conclude on balance that the confederate knew as well as saw what was in the tin. These response patterns were similar to those in the corresponding conditions in Study 1.

Moderating this trend, however, were a number of fluctuations, although none quite achieved conventional levels of significance. First of all, 4-year-olds in the ironiC/no touch condition were more likely to conclude that the confederate did *not* know what was in the tin, as in the equivalent Study 1 condition. Although this

effect was not as marked as before, this was partly due to a relatively high number of <don't know> responses, suggestive of greater uncertainty than in the neutral/no touch condition and in this sense still consistent with heightened sensitivity to the stone's ambiguity. Responses in the touch conditions also moderated the trend towards answering <yes> to the question on knowledge, with 3-year-olds being less certain here about whether the confederate knew what was in the tin, regardless of intonation. The same was true of 4-year-olds in the neutral/touch condition. In the ironic/touch condition, however, responses shifted back towards concluding on balance that the confederate *did* know what was in the tin, consistent with touch being seen as a disambiguating cue by this age group, as predicted.

*Effects of intonation and touch on perceptions of the confederate's belief*

If touch was seen as a disambiguating cue when coupled with ironic intonation, 4-year-olds' modal responses to Questions 2 and 3 under these conditions should have returned to <stone-yes/cake-no>, as in the Study 1 neutral joke condition. For 3-year-olds, on the other hand, the addition of a further cue was expected to increase <stone-yes/cake-no> responses relative to the ironic/no touch condition. Table 4 presents the relevant data. Hierarchical log linear analysis with four factors, response pattern, age group, neutral vs. ironic intonation and touch vs. no touch identified the two-way interaction between pattern and age group as the only significant effect (for this model, likelihood ratio  $\chi^2 = 23.66$ ,  $df = 24$ ,  $p = 0.48$ ). There were also, however, marginally significant effects for the three-way interactions between pattern, intonation and touch (change in likelihood ratio  $\chi^2$  when removed from the model = 7.66,  $df = 3$ ,  $p = 0.05$ ), and between pattern, intonation and age group (change in likelihood ratio  $\chi^2$  when removed = 7.19,  $df = 3$ ,  $p = 0.07$ ). The implication was that the differences in response pattern between the 3- and 4-year-olds acted to drown out more subtle interactions within age groups.

As can be seen, responses for the 3-year-olds in the no touch conditions were very similar to those in the equivalent Study 1 joke conditions. When intonation was neutral, <stone-yes/cake-yes> was again the modal response, indicating confusion, but this shifted to <stone-yes/cake-no> when ironic intonation was present. However, counter to prediction, the same pattern was reproduced almost identically in the touch conditions, as if touch had no effect at all. A separate log linear analysis on the 3-year-olds' data identified a significant main effect of pattern (for this model, likelihood ratio  $\chi^2 = 8.25$ ,  $df = 12$ ,  $p = 0.76$ ), reflecting the overall prevalence of <stone-yes/cake-yes> and <stone-yes/cake-no> responses relative to the other categories, but also a marginal two-way interaction between pattern and intonation (change in likelihood ratio  $\chi^2$  when removed = 6.16,  $df = 3$ ,  $p = 0.10$ ). Further analysis showed this to be attributable to a significant shift in modal response from <stone-yes/cake-yes> to <stone-yes/cake-no> across the intonation conditions (20 vs. 7 in the neutral condition, compared with 12 vs. 16 in the ironic; Fisher Exact = 0.02, one-sided).

Table 4. Frequency of responses given to Study 2 Questions 2 ("Does she think there is a stone in the tin?") and 3 ("Does she think there is a cake in the tin?") by condition and age group (excluding responses from children who answered don't know to one or other question).

	Pattern of response			
	stone-no cake-yes	stone-yes cake-no	stone-yes cake-yes	stone-no cake-no
3-year-olds				
1. Neutral/no touch	1	4	9	1
2. Ironic/no touch	2	8	6	0
3. Neutral/touch	1	3	11	1
4. Ironic/touch	1	8	6	1
4-year-olds				
1. Neutral/no touch	4	5	4	2
2. Ironic/no touch	8	3	4	0
3. Neutral/touch	5	8	1	2
4. Ironic/touch	1	6	5	2

The picture for the 4-year-olds was quite different. As with the younger age group, responses in the no touch conditions mirrored those in Study 1, although <stone-yes/cake-yes> answers were slightly more frequent than before. Thus, with neutral intonation, the most common response was once more <stone-yes/cake-no>, but this shifted to <stone-no/cake-yes> when ironic intonation was present, consistent with this having increased sensitivity to the stone's ambiguity, and in line with the know responses in this condition. When touch was added, in the neutral intonation condition <stone-yes/cake-no> became more clearly modal. Crucially, though, there was now no shift towards <stone-no/cake-yes> responses when ironic intonation was used, and <stone-yes/cake-no> remained the most common pattern of answers, as predicted. A separate log linear analysis on the 4-year-olds' data confirmed these effects, identifying a significant three-way interaction between response pattern, intonation and touch (change in likelihood ratio  $\chi^2$  when this effect is removed from the saturated model = 8.17,  $df = 3$ ,  $p = 0.04$ ).

#### *Effects of intonation and touch on perceptions of the stone*

Table 5 shows the frequency of <yes> vs. <no> answers to Question 5 ("Does the stone look like a cake?"). As can be seen, few 4-year-olds considered that the stone looked like a cake in the neutral intonation conditions, even if the confederate did not touch it. In contrast, more than half of those in the ironic intonation conditions *did* think the stone looked like a cake. As hypothesized, this belief was not affected by the confederate touching the stone, suggesting that children who

Table 5. Frequency of responses given to Study 2 Question 5 (“Does the stone look like a cake?”) by condition and age group (excluding don’t know responses).

	Looks like a cake	
	Yes	No
3-year-olds		
1. Neutral/no touch	7	7
2. Ironic/no touch	8	8
3. Neutral/touch	9	7
4. Ironic/touch	9	7
4-year-olds		
1. Neutral/no touch	3	12
2. Ironic/no touch	8	7
3. Neutral/touch	2	14
4. Ironic/touch	9	7

gave <stone-yes/cake-no> responses in the ironic/touch condition did so because they considered that touch would have a disambiguating effect. In line with this, the modal responses to Questions 1, 2 and 3 amongst those in this condition who thought the stone looked like a cake were <know-yes> and <stone-yes/cake-no>. In the ironic/no touch condition, the corresponding modes were <know-no> and <stone-no/cake-yes>. The 3-year-olds’ responses to Question 5 revealed no sign of any effect of intonation, being evenly split between <yes> and <no> answers regardless of condition. This suggests they were responding at random, since the expected default answer would be <no>. Confirming these patterns, hierarchical log linear analysis with four factors, response (yes vs. no), neutral vs. ironic intonation, touch vs. no touch, and age group, identified the three-way interaction between response, intonation and age group as the only significant effect (for this model, likelihood ratio  $\chi^2 = 0.79$ ,  $df = 8$ ,  $p = 1.00$ ).

**Discussion**

The pattern of the 4-year-olds’ responses to questions about the confederate’s knowledge and beliefs regarding the tin’s contents, and whether the stone looked like a cake, were all consistent with the hypothesis that ironic intonation increased their awareness of the ambiguity of the stone’s appearance, but that they recognized the disambiguating effect the confederate’s act of touching the stone would have upon her perceptions. No corresponding awareness of ambiguity or disambiguation was anticipated on the part of the 3-year-olds, and none was in evidence. The predictions made under the informational load account were not borne out, however, since the addition of the touch cue had no apparent impact on belief attributions. There were signs that the 3-year-olds noticed this cue, given

the slight shift towards greater uncertainty about the confederate's knowledge in the touch conditions, suggesting increased confusion. The effect was very marginal, however, and it may be that to children struggling to keep up with events, the touch cue seemed incidental. In general, the informational load account, and the tendency to fall back on personal knowledge when overloaded, still seems the most plausible interpretation of the 3-year-olds' performance, which in all other respects mirrored that in Study 1. Their apparently random responses to Question 5, for instance, are indicative of an inability to follow proceedings fully. In addition, they were reluctant throughout to give <stone-no> responses (87% of belief attributions included <stone-yes>), as if their own knowledge was never far from their thoughts, even though they are plainly capable of discounting this under favorable conditions: in the Study 1 neutral lie condition, 75% of the 3-year-olds responded <stone-no/cake-yes>.

## General discussion

The data from Study 2 indicate strongly that the apparently anomalous performance of 4-year-olds in the Study 1 ironic joke condition was in fact the product of a sophisticated process in which the belief of the confederate was determined via the decoding and weighing up of competing cues from sight, language and intonation. Indeed, when the confederate touched the stone, further increasing the complexity of the situation, many 4-year-olds were still able to coordinate the various implications and arrive at coherent patterns of response. Some signs of such capacity have been reported previously by Mitchell *et al.* (1997), who found that 4- to 6-year-olds were able to alter belief attribution depending on the information available to a story character, but not under such complex conditions.

This is not to claim that 4-year-olds in the present research showed uniform ability to resolve conflicting cues, but the underlying trends are real enough, and sufficiently similar across Studies 1 and 2 to suggest the effects are stable. In contrast, the apparent ability of 3-year-olds to appreciate the significance of intonation in the Study 1 ironic joke condition seems to have been, on balance, simply a manifestation of an inability to manage the same set of information that the 4-year-olds successfully unraveled, and a consequent default to reporting personal knowledge. The implication is that there is some shift after the fourth birthday in the ability to decode complex cues to the beliefs of others, including ironic intonation, although this takes longer to become evident in some children. On this point, it should be noted that the increase in apparently confused <stone-yes/cake-yes> responses among 4-year-olds in Study 2, despite their higher vocabulary scores, suggests the emergence of this ability is not strongly wedded to general language skills. This echoes Deleau & Bernard's (2003) finding of separation between conversational awareness and wider language ability, with the former being more predictive of performance on false belief tasks.

As far as the specific ability to interpret ironic intonation is concerned, these results differ considerably from those reported by Sullivan *et al.* (1995) and Winner & Leekham (1991), who found that substantially older children were unable to decode this cue. Similarly, whilst they are more consistent with the findings of Capelli *et al.* (1990) and Keenan & Quigley (1999) that sarcastic intonation was well understood by primary school children, they show that onset of its appreciation occurs at a much younger age. In this respect, they are more consistent with Yasui & Lewis (2003), who report that 4-year-olds who witnessed a scene involving use of intonationally marked sarcasm were able both to grasp and explain the non-literal nature of the language employed. As suggested earlier, it would appear that aspects of the methodology employed in previous research, including the use of narratives rather than direct experience (cf. Deleau, 1998, 1999), have served to obscure children's capabilities. This said, the 4-year-olds in the present research evidently did not grasp the conventional meaning of ironic or sarcastic intonation as understood by adults (cf. the reliability check in Study 1). In particular, whilst the modal response of <stone-no/cake-no> in the ironic lie condition in Study 1 seems to imply recognition that such intonation signals a verbal statement should be disbelieved, the idea that 4-year-olds' understand this in rule-like terms is called into question by the fact that they treated ironic intonation so differently in the joke conditions.

It is possible then that the 4-year-olds engaged in piecemeal interpretation of intonation in both the lie and the joke conditions. This would still lead to the observed outcome in the ironic lie condition if they regarded intonation here as *over-emphasis*, likely to result in disbelief, given there was no reason for the confederate not to be swayed by the experimenter's statement anyway. In the joke conditions, where other information was available to the confederate, they might have regarded intonation as solely emphasizing the verbal statement in the face of competing cues, i.e. seeing the experimenter's statement as a form of trick, rather than a joke about shared knowledge. The believability of this claim would then have to be assessed in terms of how feasible it would be to think the stone might be a cake, resulting in it being noticed that its appearance was in fact ambiguous.

This account implies 4-year-olds are capable of surprising interpretative subtlety, but something of this kind must have occurred in the joke version of the experimental task; the extension to the lie condition merely adds consistency. It also fits in with Ackerman's (1982) hypothesis that, for young children, detection of intonation is distinct from its interpretation. The implication is that 4-year-olds notice intonation is present, conclude that it is of conversational significance, and try to work out what it might signify. Rather than relying on a knowledge of convention which they do not yet possess, their interpretation rests on the contrast between the statement marked out in this way and its unmarked background, plus an assumption of relevance (i.e. that *something* is intended). Since this assumption is arguably the foundation for adult interpretations of more idiosyncratic vio-

lations of conversational convention (cf. Grice, 1975), even if 4-year-olds are unaware of what ironic intonation usually implies, they are nevertheless well advanced in terms of working out what such markers might signify for the mental sub-text of conversations.

The principal difference amongst 3-year-olds is that whilst they too seem capable of detecting intonation (cf. Siegal, 1999a), they fail to engage in the calculus exhibited by 4-year-olds, and instead struggle to come to terms with coordinating the different cues available. That 3-year-olds might have difficulty ascribing representations on the basis of multiple sources of information squares with longstanding reports of problems at this age with multiple representations (e.g. Flavell, 1988). However, the present research indicates the difficulty is essentially one of information management, and offers a different account of the widely replicated finding that 3-year-olds often ascribe their own knowledge to others in many false belief tasks (Mitchell, 1996). It is not so much that they cannot grasp alternative representations to their own, as that they cannot resolve information when there is too much of it (cf. their performance in the neutral lie condition, where there was effectively only one cue to assess).

This begs the question of what change during the short space between 3 and 4 years of age might bring about such marked differences in performance. The data in this respect are reminiscent of conceptual change accounts, but rather than suggesting the activation of a dedicated module for handling information about belief states (cf. Baron-Cohen & Ring, 1994), they indicate that the fundamental progress may be in short-term or working memory capacity (cf. Case, 1992, on the growth of free space for mental manipulation as cognitive processes become more efficient). This would be consistent with evidence (e.g. Saltmarsh & Mitchell, 1998; Wimmer & Hartl, 1991) that the presence of a physical counterpart for a false belief assists the shift amongst children as young as 3 years away from reporting knowledge of reality: such counterparts would serve to strengthen the memory trace of a potential false belief and enable children to keep better track of this possibility.

The implication is that whilst there would indeed appear to be a jump in capability between 3 and 4 years, this is based more on the ability to attend to and manipulate multiple sources of information than to comprehend fully the psychological states underpinning complex conversational exchanges. As far as the latter is concerned, change is perhaps likely to be more gradual (3-year-olds did exhibit *some* capacity to do this, after all). It is also plausible that there is to some degree a dialectical relationship between the development of working memory capacity and ability to track the mental sub-text of conversations. The former is needed to engage in the latter, but practice at the latter may stretch the former. The present data are in this respect consistent with conceptual continuity accounts of theory of mind development and the emphasis these place on the role of language. What improved memory management allows 4-year-olds to do is to start *thinking* about

the information contained in conversational exchanges and what it signifies about the intentions and interpretations of others. The use of naturalistic settings and a focus on children's actual interpretation of conversational information within such contexts appears to capture exactly this process in operation.

## References

- Ackerman, B. (1982). Contextual integration and utterance interpretation: The ability of children and adults to interpret sarcastic utterances. *Child Development*, 53, 1075-1083.
- Astington, J.W. & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, 9, 7-32.
- Baron-Cohen, S. & Ring, H. (1994). A model of the mindreading mechanism system: Neuropsychological and neurobiological perspectives. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 183-207). Hillsdale, NJ: Erlbaum.
- Baron-Cohen, S. & Swettenham, J. (1996). The relationship between SAM and TOMM: Two hypotheses. In P.K. Smith & P. Carruthers (Eds.), *Theories of theories of mind*. Cambridge: Cambridge University Press.
- Brown, J.R. & Dunn, J. (1991). 'You can cry, mum': The social and developmental implications of talk about internal states. *British Journal of Developmental Psychology*, 9, 237-256.
- Capelli, C., Nakagawa, N., & Madden, C. (1990). How children understand sarcasm: The role of context and intonation. *Child Development*, 61, 1824-1841.
- Case, R. (1992). *The mind's staircase*. Hillsdale, NJ: Erlbaum.
- Clark, H.H. & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.
- Deleau, M. (1998). Entre l'acte et la pensée: de l'activité conjointe aux symboles et aux croyances. *Enfance*, 1, 37-47.
- Deleau, M. (1999). Assessing links between conversational awareness and cognitive development. *Developmental Science*, 2, 16-18.
- Deleau, M. & Bernard, S. (2003). Conversational awareness and ToM development. XIth European Conference on Developmental Psychology, Milan.
- Deleau, M., Le Sourn, S., & Guehenneuc, K. (2000). Conversations and the representation of mind in preschool children. III Conference for Sociocultural Research, Sao Paulo.
- Dunn, J. (1994). Changing minds and changing relationships. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 297-310). Hillsdale, NJ: Erlbaum.
- Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, 62, 1352-1366.



- Dunn, L.M., Dunn, L.M., Whetton, C., & Pintilie, D. (1982). *British picture vocabulary scale*. London: NFER-Nelson.
- Flavell, J.H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. Washington, P.L. Harris, & D.R. Olson (Eds.), *Developing theories of mind* (pp. 244-267). Cambridge: Cambridge University Press.
- Garfield, J.L., Peterson, C.C. & Perry, T. (2001). Social cognition, language acquisition and the development of the theory of mind. *Mind & Language*, 16, 494-541.
- Gopnik, A. & Wellman, H.M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, 7, 145-171.
- Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics*, Vol 3, (pp 41-58). New York: Academic Press.
- Hogrefe, G., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, 57, 567-582.
- Keenan, T.R., & Quigley, K. (1999). Do young children use echoic information in their comprehension of sarcastic speech? A test of echoic mention theory. *British Journal of Developmental Psychology*, 17, 83-96.
- Lewis, C., Freeman, N.H., Kyriakidon, C., Maridaki-Kassotaki, K., & Berridge, D.M. (1996). Social influences on false belief access: specific sibling influences or general apprenticeship? *Child Development*, 67, 2930-2947.
- Lillard, A. (2003). Discriminating pretend and real snacks: A fundamental problem in early cognition. British Psychological Society Developmental Section Conference, Coventry.
- McRoberts, G.W. & Best, C.T. (1997). Accommodation in finding zero during mother-infant and father-infant vocal interactions: A longitudinal case study. *Journal of Child Language*, 24, 719-736.
- Mitchell, P. (1996). *Acquiring a conception of mind: A review of psychological research and theory*. Hove: Psychology Press.
- Mitchell, P., Robinson, E.J., Nye, R.M., & Isaacs, J.E. (1997). When speech conflicts with seeing: Young children's understanding of informational priority. *Journal of Experimental Child Psychology*, 64, 276-294.
- Newton, P., Reddy, V. & Bull, R. (2000). Children's everyday deception and performance on false-belief tasks. *British Journal of Developmental Psychology*, 18, 297-317.
- Ninio, A. & Snow, C. (1996). *Pragmatic development*. Boulder, Colorado: Westview Press.
- Perner, J. & Davies, G. (1991). Understanding the mind as an information processor: "Do young children have a copy theory of mind?" *Cognition*, 39, 51-69.
- Perner, J., Leekam, S. & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125-137.

- Perner, J., Ruffman, T., & Leekam, S. (1994). Theory of Mind is contagious: You catch it from your sibs. *Child Development*, 65, 1228-1238.
- Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology*, 34, 161-174.
- Saltmarsh, R. & Mitchell, P. (1998). Young children's difficulty acknowledging false belief: Realism and deception. *Journal of Experimental Child Psychology*, 69, 3-21.
- Siegal, M. (1999a). Language and thought: the fundamental significance of conversational awareness for cognitive development. *Developmental Science*, 2, 1-34.
- Siegal, M. (1999b). Beyond methodology: Frequently asked questions on the significance of conversation for development. *Developmental Science*, 2, 29-34.
- Siegal, M. & Peterson, C.C. (1994). Children's theory of mind and the conversational territory of cognitive development. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 427-455). Hove: Erlbaum.
- Siegal, M. & Want, S.C. (2003). Early sensory impairment and performance on theory of mind tasks. XIth European Conference on Developmental Psychology, Milan.
- Sullivan, K., Winner, E., & Hopfield, N. (1995). How children tell a lie from a joke: The role of second-order mental state attributions. *British Journal of Developmental Psychology*, 13, 191-204.
- Wimmer, H. & Hartl, M. (1991). Against the Cartesian view on the mind: Young children's difficulty with own false beliefs. *British Journal of Developmental Psychology*, 9, 125-138.
- Winner, E. (1988). *The Point of Words: Children's Understanding of Metaphor and Irony*. Cambridge, MA: Harvard University Press.
- Winner, E. & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speaker's second order intention. *British Journal of Developmental Psychology*, 9, 257-270.
- Yasui, M. & Lewis, C. (2003). The irony of false belief: Young children's understanding of sarcasm and mental states. British Psychological Society Developmental Section Conference, Coventry.