J. Child Lang. **27** (2000), 367–392. Printed in the United Kingdom

Children's understanding of homonymy: metalinguistic awareness and false belief*

MARTIN J. DOHERTY
Department of Psychology, University of Stirling

(*Received* 1 *July* 1998. *Revised* 20 *September* 1999)

ABSTRACT

The aim of this study was to explain why children have difficulty with homonymy. Two experiments were conducted with forty-eight children (Experiment 1) and twenty-four children (Experiment 2). Three- and four-year-old children had to either select or judge another person's selection of a different object with the same name, avoiding identical objects and misnomers. Older children were successful, but despite possessing the necessary vocabulary, younger children failed these tasks. Understanding of homonymy was strongly and significantly associated to understanding of synonymy, and more importantly, understanding of false belief, even when verbal mental age, chronological age, and control measures were partialled out. This indicates that children's ability to understand homonymy results from their ability to make a distinction characteristic of representation, a distinction fundamental to both metalinguistic awareness and theory of mind.

INTRODUCTION

Homonyms are words with two distinct, unrelated meanings such as *bat* (flying mammal) and *bat* (sports equipment). Preschool children have difficulties correctly interpreting homonyms: they tend to interpret homonyms as the most common meaning even when contextual information indicates this is absurd (Campbell & Macdonald, 1982; Beveridge & Marsh, 1991). Even older children appear to have difficulties learning homonyms (Mazzocco, 1997).

Understanding homonymy is clearly metalinguistic: it requires children to distinguish between the linguistic medium and what it represents. It also

[*] Address for correspondence: Martin J. Doherty, Department of Psychology, University of Stirling, Stirling FK9 4LA, Scotland. tel: +44-(0)1786 466 366. fax: +44-(0)1786 477 641. e-mail: mjd1@stirling.ac.uk

requires children to understand at least one aspect of the relationship between the two: that one linguistic form can be used to represent more than one kind of object. Children's difficulties with homonymy could result from their metalinguistic awareness, or lack of it, in two possible ways:

(1) The metalinguistic deficit account: children may lack the ability to conceive of the relationship between linguistic form and meaning. As a result, they would not search for a secondary meaning when their initial interpretation was absurd.

(2) The metalinguistic one-to-one mapping account: children may be able to conceive of the relationship between linguistic form and meaning, but hold an erroneous assumption about the nature of this relationship. In relation to children's understanding of homonymy, it has been proposed that children assume each meaning is expressed by a distinct form (Slobin, 1985). This one-to-one mapping assumption might be useful in learning language, since it constrains the number of hypotheses children need to consider when trying to learn the meanings of words.

However, the one-to-one mapping account has not been spelled out in detail, and it may not require metalinguistic awareness. Thus there is a third, non-metalinguistic reason for children's difficulties with homonyms:

(3) The implicit one-to-one mapping account: rather than an explicit assumption that one form has just one meaning, the one-to-one mapping assumption may be better characterized as an implicit tendency not to assign the same form to different meanings. This would imply that children simply do not learn homonyms.

Either of the metalinguistic accounts would also predict learning homonyms would be difficult. Whichever of the three accounts is true, it is conceivable that young children learn some homonyms. Under either metalinguistic account, children might learn homonyms because they fail to notice they already know a different meaning for that particular linguistic form. For the metalinguistic one-to-one mapping account, children might relax their assumption for individual words given strong evidence that the word does in fact violate the one-to-one mapping assumption. Some homonyms might also be learned because children did not notice that they already knew a different meaning for a novel word. For the implicit one-to-one mapping account, given repeated exposure to the secondary meaning of a homonym children might eventually learn two distinct meanings for the word. However, for the implicit one-to-one mapping account, once they had learned both words, children would not be able to access both words on demand without the addition of metalinguistic awareness.

Clearly these differing explanations will be hard to tease apart. Studies to date have failed to do this. Peters & Zaidel (1980) and Backscheider & Gelman (1995) found that by at least the age of four years children possessed the metalinguistic ability to identify homonym pairs. In both studies children

were shown four pictures, two of which were homonymous. Children were required to identify them, and thus acknowledge that different kinds of object can share the same name. Peters & Zaidel used a graded series of prompts from 'find another picture that sounds exactly the same as this one but means a different kind of thing' to 'find another kind of [e.g.] bat'. They found that children of mean age 3;10 made just under 50% correct responses on this task, whereas children of mean age 4;9 gave 84% correct responses. Backscheider & Gelman, using the same task but omitting all but the final most explicit prompt, found that children of mean age 3;9 were about 75% correct on this task. They did not include younger or older children, so the development of this ability is difficult to judge.

This task seems to require children to understand the relationship between words and their referents, and thus to require metalinguistic awareness. However, it also requires children to possess the relevant vocabulary. Backscheider & Gelman (1995) found that on average 82% of homonym pairs were in children's vocabularies. Although this is similar to the 75% success rate on the experimental task, direct comparison is not possible because different children took the vocabulary and experimental tasks. Peters & Zaidel (1980) did not report the level of vocabulary performance, but found very high correlations between vocabulary performance and success on the Homonym task. Thus it is possible that younger children in these studies had the necessary metalinguistic awareness but lacked the vocabulary to demonstrate it.

To begin to resolve why children should experience difficulties with homonyms it is necessary to distinguish between vocabulary difficulties and metalinguistic difficulties. Evidence of relevant metalinguistic abilities in preschoolers has been provided by Doherty & Perner (1998). They examined children's understanding of synonymy, which like homonymy requires children to understand the relationship between linguistic form and meaning: for synonymy, one referent has two names, and for homonymy, one name has two referents. They found that the ability to produce one half of a synonym pair when given the other (Experiments 3 & 4), or to judge others' attempts to produce synonyms (Experiments 1 & 2) develop around the age of four years. Younger children were unable to do this, even though they possessed the necessary vocabulary and were able to pass control tasks of equivalent logical structure and complexity.

These difficulties with synonymy would be predicted by either the metalinguistic deficit account or the metalinguistic one-to-one mapping account. Children's error patterns favoured the metalinguistic deficit account, but the strongest evidence came from children's performance on the False Belief task. Children's ability to predict another person's mistaken belief correlated very highly with their performance on the Synonym task, with correlations remaining between $r = 0{\cdot}60$ (Experiment 4) and $r = 0{\cdot}85$

(Experiment 1) even after control measures and verbal intelligence were partialled out.

Doherty & Perner (1998) argue that understanding of synonymy and of false belief are strongly associated because both metalinguistic awareness and understanding of false belief rest on a common insight that things can be represented in a certain way (Perner, 1991, 1995). In the typical False Belief task, the protagonist places an object in one location which is moved in his absence. The protagonist has a mental state which represents something (the real location) in a certain way (in this case, in a way that differs from reality). Similarly language represents states of affairs in a certain way (e.g. in terms of its formal structure). In the case of synonyms, the same state of affairs can be represented in terms of different formal structures. Since homonymy is very similar to synonymy, the argument can be extended to the case of homonyms. The metalinguistic ability to understand homonymy should develop at the same time as the abilities to understand synonymy and false belief, i.e. at roughly the same time that children pass the Homonym task devised by Peters & Zaidel (1980). The aim of the present study is to test this prediction. This was in order to distinguish between the metalinguistic deficit account and both of the one-to-one mapping accounts of children's homonym difficulties, and to further test Doherty & Perner's hypothesis that children succeed on the False Belief task because they can distinguish between what is represented and how it is represented. This was done in the following way:

(1) The Homonym task used in previous studies was improved by ensuring that failure does not result from lack of the necessary vocabulary.

(2) The Homonym task was administered together with a suitable Synonym task and the False Belief task. If children's difficulty with homonyms is metalinguistic in nature, these three tasks should be strongly associated.

(3) To avoid failure due to misunderstanding of instructions or task demands, an extensive modelling phase was included in which the experimenter provides the answers for the child if necessary, along with feedback and explanations.

(4) Peters & Zaidel (1980) and Backscheider & Gelman (1995) both suggest deficits in children's search skills as potential reasons for failure on the Homonym task. To minimize the need for these skills, instead of requiring children to identify homonyms themselves, this was 'attempted' by a puppet. Children had to judge whether Puppet was successful or not.

(5) As a further guarantee against failure due to general task demands an analogous control task was introduced. Instead of having to select a homonym, Puppet's job was to point to another picture of the same kind of object.

(6) Since performances on the Homonym, Synonym and False Belief tasks might be related simply due to a common association with verbal mental age,

a measure of verbal mental age was taken using the British Picture Vocabulary Scale, the British version of the Peabody Picture Vocabulary Scale (Dunn, Dunn, Whetton & Pintillie, 1982).

The modified Homonym task is analogous to the Synonym Judgement task used by Doherty & Perner (1998, Experiments 1 & 2) and follows the same logic. Children are shown sets of pictures which include a homonym pair, for example, *bat* (animal) and *bat* (sporting equipment). Children are asked to indicate one bat and then have to judge whether a puppet successfully indicates the OTHER bat. In three types of trial the puppet indicates the same object as the child, an irrelevant object, or the homonym. In order to successfully judge the puppet's performance on every trial children must monitor MEANING, to check that the puppet indicates a different kind of object, and VERBAL FORM, to check that the puppet indicates an object with the same name. If children monitored the difference in meaning alone they would judge incorrectly on the trial in which the puppet indicates an irrelevant object; if they monitored the sameness of verbal form, they would fail the trial in which the puppet indicates the same object as the child. Thus the appropriate criterion for success on the Homonym task – and the analogous Synonym and Object Pointing control tasks – is success on all 3 types of trial.

EXPERIMENT 1

METHOD

Participants

The participants were 48 children (20 boys and 28 girls) from a University preschool in Stirling, Scotland. Ages ranged from 2;11 to 4;7, with a mean age of 3;5 and a standard deviation of 4·5 months. For the analysis of results children were divided into two groups: a younger group (24 children from 2;11 to 3;4, mean age 3;1, S.D. = 1·5 months) and an older group (24 children from 3;4 to 4;7, mean age 3;8, S.D. = 3·5 months).

Design

Each child was tested on all four tasks: Homonym Judgement, Synonym Judgement, False Belief, and Object Pointing control. Tasks were administered over two sessions about a week apart with two tasks per session. The order of administration was counterbalanced in a 4×4 sequence balanced Latin square design. In addition the British Picture Vocabulary Scale (BPVS) long form was administered about a week after the second session.

Procedure and materials

Each child was seen in a quiet and familiar room adjacent to the nursery area. The following four tasks were administered in the order discussed in the Design section above.

Homonym Judgement task

The Homonym Judgement task comprised three phases: vocabulary check, modelling, and test phase.

Vocabulary check. In this phase children were given a vocabulary test checking on their knowledge of the homonyms used later in the modelling and test phases. It also served to alert the child to the distinctions which had to be made in the experiment. Seven A4 sheets were used, each with four pictures on: both members of a homonym pair and two distracters. See Table 1 for a list of the homonyms used.

TABLE 1. *Experimental items and percentage of children who identified both items from homonym or synonym pairs*

	Homonym 1/2	%known	Synonym pairs	%known
Modelling	Letter (grapheme/envelope)	92	Rabbit/Bunny	
	Nail (iron/finger)	96	Cup/Mug	
Test	1 Bat (sports/flying)	88	Lady/Woman	92
	2 Glasses (drinking/specs.)	98	Truck/Lorry	98
	3 Knight/Night	73	TV/Television	96
Replacements	4 Bare/Bear	91	Jacket/Coat	96
	5 Bow (ribbon/weapon)	100		

Children were shown each sheet with the homonym judged to be most familiar to the child covered by a white piece of card. The first sheet had a picture of a fish, a bicycle, a letter (the letter A), and a letter (a stamped addressed envelope). The envelope was covered, children were shown the picture and asked 'Can you point to letter?'. (Although 'a letter' would have been more natural, the indefinite article would have been less appropriate for some of the other items, such as *night*, *glasses*, and particularly *bare*, so it was excluded.) After the child responded, the card was moved to cover the other homonym, the letter A, and the child was asked 'Which one of these is letter?'. If children pointed to the card now covering their original choice the visible three pictures were indicated and they were asked 'Which one of THESE THREE is letter?'. When children had identified the second homonym,

the first was uncovered and the experimenter said, 'So this is letter, and this is letter. They've got the same name!'. The procedure was repeated with the next four sheets. On the third and fourth sheets, after the first homonym had been identified and covered, children were asked to identify one of the distracter items. This was to prevent them thinking that for the second question they should always point to the object that had just been uncovered. Any vocabulary failures were noted and these items were replaced in the modelling and test phases by the replacement items (see Table 1). To continue to the test phase children had to know at least 5 out of the 7 homonym pairs. Only one child (3;5) did not, and was replaced.

Modelling phase.The objective of this phase was to model the actual test procedure. A teddy bear glove puppet was used. The child was shown the first of the A4 sheets used in the vocabulary check and invited to point to a 'letter' (grapheme or envelope). Puppet's job was to point to 'the other letter, NOT the same one you pointed to'. Puppet made two incorrect attempts, pointing first to the same letter as the child and then to an unrelated object (a bicycle). Finally Puppet correctly pointed to the other letter. Each attempt was followed by the test question: 'Is that what he should have done?'. After a short pause the experimenter provided the answer, repeating that Puppet's job was to point to a letter, but not the same one as the child. The procedure was repeated for the other warm-up item, nail (fingernail or iron nail).

Test phase. For the test phase the modelling phase was continued with 3 new picture sheets and no feedback. Puppet made only one attempt to identify the homonym per sheet and then the TEST QUESTION: 'Is that what he should have done?' was asked. The sheets containing homonym pictures were always presented in a fixed order, as listed in Table 1 (items children failed to identify in the vocabulary check were replaced from replacement items 4 and 5).

For the Homonym task, and the Synonym and Object Pointing tasks, there were three trial types: in one trial Puppet incorrectly gave the same response as the child. In one trial Puppet pointed to a distracter (Homonym and Object Pointing tasks) or misnamed the object (Synonym task). In one trial Puppet correctly pointed to a different picture with the same name (Homonym), pointed to a different picture of the same object (Object Pointing), or gave a synonym (Synonym). Assignment of the three trial types to the three picture sheets was counterbalanced between participants in a 3×3 sequence balanced Latin square design. For individual participants, trials were presented in a different order for each of the three tasks. To pass the Homonym, Synonym, or Object Pointing tasks, children had to answer correctly on each of the three trials, as discussed in the Introduction.

Object Pointing control

The procedure in this task was designed to parallel the modelling and test phases of the Homonym Judgement task.

Children were shown an A4 sheet with four drawings on. Two of the drawings were the same (of the same objects used in the Synonym Judgement task, below) and two were different, e.g. two rabbits, a tree and a house. The child was asked to point to one of the identical pictures (in this case, a rabbit) and Puppet's job was 'to point to the other rabbit, NOT the one [the child] pointed to'. Children were required to judge whether Puppet's response was appropriate. The test question was: 'Is that what he should have done?'.

There were two trials of this procedure in the modelling phase, each with a different sheet of pictures. As in the homonym modelling phase, Puppet pointed to the same object as the child and then an unrelated object before pointing correctly, with the experimenter providing appropriate feedback and explanation. For the test phase the procedure continued without feedback, Puppet only making one response per sheet. The sheets were always presented in a fixed order, and the three trial types were counterbalanced as discussed above.

Synonym Judgement task

The Synonym Judgement task also designed to be as far as possible analogous to the Homonym Judgement task. It consisted of three phases: vocabulary-check, modelling and test phase.

Vocabulary check. In this phase children were given a vocabulary test checking on their knowledge of the synonyms used later in the actual test. It also served to alert the child to the distinctions which had to be made in the experiment. Four A4 sheets were used. Each of them had four pictures on it. Two of the pictures were experimental items used later (truck/lorry and woman/lady on two of the sheets; TV/television and coat/jacket on the other two). The other two items on each sheet were chosen from among a rabbit, a cat, an apple, a bird, and a daisy. Children were shown each sheet and asked to point to, e.g. a truck, and then to a lorry. If they hesitated they were given encouragement, and the question was repeated if they answered incorrectly. They were then told that the object has two names, lorry and truck. On the third and fourth sheets the first item to identify was not one of the experimental items, in order to prevent children from thinking that the same item was required for both questions on each sheet, and then pointing to the same item regardless of which word is used. Then the experimental item was asked about once with each synonym, as before.

Modelling phase. The objective of this part of the procedure was to model the actual test procedure. A white glove puppet and a hand drawn 10 × 15 cm colour picture showing a rabbit were used. The child was shown the picture

of a rabbit, told that it could be called a rabbit or a bunny, and invited to choose one of these names. The puppet's task was 'to say the other name. NOT the one that you said'. Puppet made two incorrect attempts, first using the same name as the child and then saying something unrelated (e.g. elephant or banana). Finally Puppet correctly gave the synonym. Each attempt was followed by the test question: 'Is that what he should have said?'. After a short pause the experimenter provided the answer, reminding the child that Puppet's job was to produce a correct name, but not the same one as the child. The procedure was repeated for the other warm-up item, cup/mug.

Test phase. For the actual test the modelling phase was continued with four new pictures but no feedback was given. Puppet named each item only once and then the test question: 'Is that what he should have said?' was asked. The three items were always presented in fixed order as listed in Table 1, but the assignment of response-type was counterbalanced as discussed above. Depending on response-type (same, synonym, different-meaning) the puppet used one of the following words for each of the following three items (plus replacement):

1. woman: 'woman', 'lady', 'truck'.
2. truck: 'truck', 'lorry', 'lady'.
3. television: 'television', 'TV.', 'coat'.

(4. coat: 'coat', 'jacket', 'television'.)

False Belief test

For this test a short story was acted out with two Playpeople dolls (5 cm), a marble, an opaque jar (5 cm high × 2·5 cm wide) and a box (3 cm high × 4 cm wide). In the story one of the dolls, Sally, places a marble in the box and exits. In her absence the other doll moves the marble to the jar and also leaves. Sally returns and children are asked the following questions:

Belief question:	Where will she look first for her marble?
Reality question:	Where is the marble really?
Memory question:	Where did Sally put the marble in the beginning?

British Picture Vocabulary Scale

The long form of the BPVS was administered about a week after the original test sessions.

RESULTS

All p-values are 2-tailed, except Fisher's exact p-values, which are 1-tailed. Correlations are Pearson product-moment correlations (r). Special cases of Pearson's r are used when one variable is dichotomous (point biserial coefficient, r_{pb}) or both variables are dichotomous (phi coefficient, r_{ϕ}).

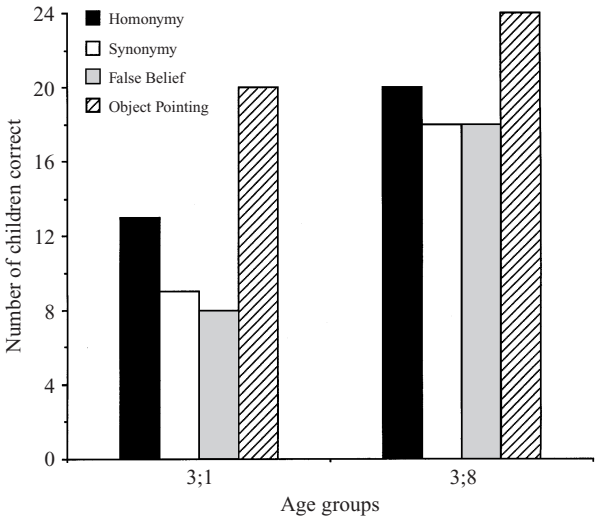

Fig. 1. Number of children showing understanding of Homonymy, Synonymy and False Belief in Experiment 1.

Vocabulary checks

Table 1 shows the performance on the homonym and synonym vocabulary checks. Twenty-six out of the 48 children knew all 5 pairs of homonyms, and with the inclusion of replacement items all but one child knew at least 5 homonym pairs. This child was replaced as noted in the Method. All the children knew at least 3 synonym pairs. Thus for the Synonym and Homonym tasks, all test pairs were demonstrably in children's vocabularies.

False Belief control questions

Two children failed the memory question on the False Belief task, and failed the belief question. Two children failed the reality question, and passed the false belief question. All four children were included in the analysis.

Performance and comparison of experimental tasks

Figure 1 shows the number of children in the two age groups passing the four tasks: Homonym, Synonym, False Belief and Object Pointing. Clearly, performance on the Object Pointing control is very good for both age groups. The difference between the two groups' performances is marginally significant, Fisher's exact, $p = 0{\cdot}055$. For the Homonym, Synonym, and False Belief tasks the age differences are more pronounced and significant: Homonym task, Fisher's exact, $p = 0{\cdot}03$; Synonym task, Fisher's exact, $p = 0{\cdot}006$, False Belief task, Fisher's exact, $p = 0{\cdot}004$.

https://doi.org/10.1017/S0305000900004153 Published online by Cambridge University Press

TABLE 2. (a) *Ranges, means, standard deviations and intercorrelations between variables of Experiment 1.* (b) *Correlations between Homonym, Synonym, and False Belief tasks, partialling out verbal mental age (BPVS) and age*

(a)

Variable	Range	Mean	S.D.
Age	2;11–4;7	3;5	4·5 m.
BPVS	1;8–6;5	3;9	12·5 m.
Homonym	0–1	0·69	0·47
Synonym	0–1	0·56	0·50
Object Pointing	0–1	0·92	0·28
False Belief	0–1	0·54	0·50

	Synonym	False Belief	Object Pointing	BPVS	AGE
Homonym	0·67***	0·55***	0·28	0·42**	0·33*
Synonym	—	0·71***	0·34*	0·59***	0·42**
False Belief		—	0·33*	0·60***	0·48***
Object Pointing			—	0·32*	0·28
BPVS				—	0·27

(b)

	Synonym	False Belief
Homonym	0·48***	0·35**
Synonym	—	0·54***

*$p < 0{\cdot}05$; **$p < 0{\cdot}01$; ***$p < 0{\cdot}001$.

However, even for the younger group, performances on the language tasks were above the level expected if children were simply guessing ($p = 0{\cdot}125$): Homonym task, Binomial, N = 24, k = 13, $p < 0{\cdot}001$; Synonym task, Binomial, $n = 24$, $k = 9$, $p < 0{\cdot}001$. It is inappropriate to compare performance on the false belief task to chance because children without false belief understanding typically make the systematic error of looking in the object's actual location (for an illustration of the probability of passing the false belief task at different mental and chronological ages, see Happé, 1995).

Table 2 gives further details of the experimental variables, and the correlations between them. Correlations between the Homonym, Synonym, and False Belief tasks are substantial and highly significant. However, performance on each of these tasks is also significantly related to both age and BPVS score. Thus the relationship between them might be attributable to a common relationship with verbal mental age and other more general age related abilities. Table 2 (b) shows that after age and BPVS score have been partialled out, the Homonym, Synonym and False Belief tasks remain significantly correlated. The near ceiling performance on the object pointing

task means that the correlations between it and the other experimental variables were low.

Despite the relationship between the tasks, from Figure 1 it appears that for the younger group performance on the Homonym task is superior to performance on the other two experimental tasks. In order to investigate this possible difference between the Homonym task on the one hand, and the Synonym and False Belief tasks on the other, the following analysis looks at the relationship between the three tasks within each age group. Table 3 gives details of the experimental variables and intercorrelations for the older 24 children. All three tasks are strongly correlated; after age and BPVS score are partialled out, the Synonym–False Belief correlation just falls short of conventional significance ($p < 0{\cdot}07$), but the correlation between the Synonym and Homonym tasks remains remarkably high, and the correlation between the Homonym task and the False Belief task is substantial and significant.

The younger group shows quite a different pattern, as shown in Table 4. The Homonym task is significantly correlated to the Synonym and False Belief tasks, although less strongly than for the older children. However,

TABLE 3. (a) *Ranges, means, standard deviations, and intercorrelations between variables of Experiment 1 : older age group*. (b) *Correlations between Homonym, Synonym, and False Belief tasks, partialling out verbal mental age (BPVS) and age*

(a)

Variable	Range	Mean	S.D.
Age	3;4–4;7	3;8	3·5 m.
BPVS	2;5–6;5	4;1	12 m.
Homonym	0–1	0·83	0·38
Synonym	0–1	0·75	0·44
Object Pointing	0–1	1·00	0·00
False Belief	0–1	0·75	0·44

	Synonym	False Belief	BPVS	AGE
Homonym	0·78***	0·52**	0·20	0·19
Synonym	—	0·56**	0·49*	0·23
False Belief		—	0·46*	0·34
BPVS			—	0·14

(b)

	Synonym	False Belief
Homonym	0·79***	0·47*
Synonym	—	0·40

*$p < 0{\cdot}05$; **$p < 0{\cdot}01$; ***$p < 0{\cdot}001$.

https://doi.org/10.1017/S0305000900004153 Published online by Cambridge University Press

TABLE 4. (a) *Ranges, means, standard deviations, and intercorrelations between variables of Experiment 1 : younger age group.* (b) *Correlations between Homonym, Synonym, and False Belief tasks, partialling out verbal mental age (BPVS) and age*

(a)

Variable	Range	Mean	S.D.
Age	2;11–3;4	3;1	1·5 m.
BPVS	1;8–6;5	3;5	12 m.
Homonym	0–1	0·54	0·51
Synonym	0–1	0·38	0·49
Object Pointing	0–1	0·83	0·38
False Belief	0–1	0·33	0·48

	Synonym	False Belief	Object Pointing	BPVS	AGE
Homonym	0·54**	0·47*	0·26	0·49*	0·08
Synonym	—	0·73***	0·35	0·59**	0·21
False Belief		—	0·32	0·62***	0·16
Object Pointing			—	0·36	0·16
BPVS				—	−0·19

(b)

	Synonym	False Belief
Homonym	0·30	0·19
Synonym	—	0·50**

*$p < 0·05$; **$p < 0·01$; ***$p < 0·001$

once age and BPVS scores are partialled out, the correlations become small and non-significant. By contrast, the correlation between the Synonym and False Belief tasks remains substantial and significant. The mean level of performance on the Synonym and False Belief tasks is roughly comparable for each age group. While performance on the Homonym task is superior for both groups, this is particularly marked for the younger children

DISCUSSION

The results appear to show that understanding of homonymy is just developing between the ages of three and four years. Unlike previous studies, failure on the Homonym task cannot be attributed to the absence of homonyms from children's vocabularies, since all children correctly identified both items from the homonym test pairs in the vocabulary check. The excellent performance on the Object Pointing control task indicates that

failure is not attributable to a lack of search skills or any more general difficulties with the format or understanding of task demands.

Furthermore, the Homonym task was associated strongly and significantly with both the Synonym task and the False Belief tasks. These tasks also associated strongly with each other, replicating the findings of Doherty & Perner (1998; Experiments 1 & 2). This suggests common factors underlie these tasks over and above any common association with chronological age and verbal mental age. However, this conclusion is weakened by the fact that for the youngest group, once chronological and verbal mental age are partialled out, there was no longer any specific relationship between the Homonym task and either the Synonym or the False Belief task.

One possible explanation for this difference between the age groups is that the Homonym Judgement task, along with the tasks used by Peters & Zaidel (1980) and Backscheider & Gelman (1995), is prone to false positives. The task was designed to measure the ability to compare a difference in the object with sameness of word. This would require children to conceive of the relationship between linguistic form and what it represents. However, the way the task is implemented means that difference of object also always corresponds to difference of location. Children could therefore succeed on this task by comparing a difference in LOCATION with sameness in word: they could realize that it was incorrect for Puppet to choose the same location as they did (or precisely the same object). When Puppet chooses an object in a different location, children could give the correct reply by simply asking themselves (e.g.) 'is this a *bat*?'. This non-metalinguistic strategy would not be available for the Synonym task.

A few children employing this strategy could account for the difference between the experimental tasks of Experiment 1. The older children were generally successful at all tasks, so the possibility of false positives would have little effect on their overall results. However, the effect of false positives on the younger group's performance on the Homonym task would have been more marked. Furthermore, the children able to devise this alternative strategy during the modelling phase, would tend to be more intelligent. This would explain why partialling out BPVS scores, as well as age, caused the correlations with the Synonym and False Belief tasks to dwindle to non-significance. The aim of Experiment 2 is to confirm the overall findings of Experiment 1 by eliminating the possibility of false positives. In Experiment 2, the initial exemplar of a homonym is no longer one of the response options.

EXPERIMENT 2

The judgement task used in Experiment 1 is quite conservative: children only succeed if they produce a series of correct responses, so brief lapses in attention can result in failure, and the task is unlikely to show transitional

performance. The following experiment allowed a range of performance by using a selection procedure based on that used by Peters & Zaidel (1980) and Backscheider & Gelman (1995). In their procedure children were shown four pictures, two of which were homonymous; the experimenter selected one, and children had to find the other. This is prone to false positives, as discussed above: if children realize they are not allowed to select the same picture as the experimenter, they need only consider whether any of the other three pictures is, e.g. a bat. To avoid this problem, children were shown a single picture, and required to select a homonym from four additional pictures, which comprised an identical picture, the target homonym, and two distracters. Children had to avoid the identical picture, thus attending to the meaning of the word, but pick something with the same name, thus also attending to the form of the word. Therefore this task requires the ability to understand the relationship between meaning and linguistic form

However, this modification renders the task prone to other false positives: children could achieve a degree of success by selecting the object corresponding to the most common meaning of the word. This can be avoided by making each of the two objects the target on separate trials, and requiring successful choice on both trials. This makes the task analogous to the Synonym Production task used by Doherty & Perner (1998, Experiments 3 & 4), but by doubling the number of trials, it makes the Homonym task longer than the task used in Experiment 1. Because preschool children have limited attention spans, and because the main comparison of interest in this study is between understanding of homonymy and false belief, a Synonym task was not included in Experiment 2. Instead, a second False Belief task was included to give a more consistent picture of children's false belief understanding.

The homonym vocabulary phase was modified so that each sheet only contained one homonym. Although the version used in Experiment 1 helped alert children to the distinctions they had to make, it may also have produced false negatives: being asked the same question twice, but having to select a completely different object the second time is pragmatically rather odd, and many children did initially seem to suspect they were being asked a trick question. The object pointing control task was also modified to make it more similar to the Homonym task. Now, just as in the Homonym task, children had to avoid the identical picture, and pick a picture of a different object. Instead of being a different kind of object with the same name, however, it had to be a different object of the same kind (e.g. a different TV).

METHOD

Participants

The participants were 24 children (14 girls and 10 boys) from a state preschool with a predominantly working class intake in Stirling, Scotland. Ages ranged from 3;6 to 5;2, with a mean age of 4;2 and a standard deviation of 6 months. For the analysis of results children were divided into two groups: a younger group (12 children from 3;6 to 4;0, mean age 3;9, S.D. = 2 months) and an older group (12 children from 4;0 to 5;2, mean age 4;7, S.D. = 4·5 months).

Design

Each child was tested on all four tasks: Homonym Selection, Object Selection control, 'Sally' False Belief, and 'Puppet' False Belief. Tasks were administered over two sessions up to a week apart with two tasks per session. The order of administration was counterbalanced in a 4×4 sequence balanced Latin square design. In addition the BPVS short form was administered at the end of the session containing the Object Selection control (because this was the shorter session).

PROCEDURE AND MATERIALS

Homonym Selection task

The Homonym Selection task again comprised three phases: vocabulary check, modelling, and test phase.

Vocabulary phase. Fourteen A4 sheets were used, each with a picture of one of the seven homonyms to be used in the experiment, and 3 distracters. The first sheet had a picture of a fish, a sofa, a cake, and a cricket bat. Children were asked 'which one is bat?'. If they chose incorrectly they were told the right answer, then the next sheet was presented. The target on each of the first seven sheets was one half of the seven homonym pairs. The other half of each pair appeared in the same order on the last seven sheets.

Modelling and test phases. Children were shown an A4 sheet on which were four pictures, both items from a homonym pair and two distracters. The first sheet had a picture of a (metal) nail, a (finger) nail, a sofa, and a rabbit. Above the A4 sheet was placed a 10×15 cm card on which was a picture identical to one of the homonym pictures on the sheet. For half the children it was a finger nail, and for half a metal nail. The card was pointed out and the child was told: 'Look, here's a nail. But, can you show me a DIFFERENT kind of nail?'.

If children pointed correctly they were praised and the experimenter pointed to both pictures and said 'yes, look, this is a nail and this is a nail,

but they're different aren't they?'. If children pointed to the identical picture, the experimenter pointed out that it was the same kind of nail, and stressed they were looking for a *different* kind of nail. If the child still did not point to the other half of the homonym pair, the experimenter indicated it himself and then gave the feedback as for successful choice. This procedure was repeated for the next two modelling items, bare/bear and knight/night.

For the test phase the procedure continued without feedback or prompting for 2 sets of four new sheets and cards. Each experimental homonym pair appeared once in each set, in the order: letter, bow, bat, glasses. Half of the children received set A first, and half set B first. The cards presented with set A/set B depicted:

Letter (grapheme/envelope)
Bow (ribbon/weapon)
Bat (flying/cricket)
Glasses (spectacles/drinking).

Object Selection control

The procedure in this task was designed to parallel the modelling and test phases of the Homonym selection task. The only difference was that instead of pictures of two items in a homonym pair, two pictures of different exemplars of the same kind of thing were used. For example, the first sheet had a TV (with four legs), another TV (on a pedestal, with an aerial on top), a cat and a block of cheese. A picture of one of the TVs was placed above the A4 sheet, pointed to, and the child was told:

'Look, here's a TV. But, can you show me a DIFFERENT kind of TV?'.

Feedback and explanations were given as for the Homonym Selection task. The modelling phase was continued for two more trials with a sheet with pictures of two different cups and a sheet with pictures of two different coats. For the test phase the procedure continued without feedback or prompting for 2 sets of four new sheets and cards. As for the Homonym task, each pair appeared once in each set in the order: truck, tree, chair, house, and one picture was on the cards in set A, and the other in set B. Half the children received set A first, and half received set B first.

'*Sally*' *False Belief task*

This was the same task used in Experiment 1.

'*Puppet*' *False Belief task*

For this task children were introduced to Puppet, who was holding a key. He put his key in a 9 × 9 cm black box, and then went to have a nap in the experimenter's bag. While he was asleep the experimenter announced he was

going to play a trick on Puppet, and moved the key from the box to underneath the experimenter's scoring sheets. Then Puppet 'woke up', and as the experimenter reached into his bag to retrieve him, the child was asked the following questions:

Belief question:	Where will Puppet look first for his key?
Reality question:	Where is the key really?
Memory question:	Where did Puppet put the key in the beginning?

RESULTS

Vocabulary check

Vocabulary performance was good: 14 children identified all four experimental pairs and the remaining 10 children identified three of the four experimental pairs. Eight out of the ten failures were on bow/bow, 1 was on (flying) bat and 1 was on letter (envelope). Immediately following a vocabulary failure the correct object was pointed out to children. In the Homonym Selection task, 6 of the 10 children failing a pair on the vocabulary check went on to pass that pair and the remaining 4 children failed all 4 pairs, despite identifying 3 of them all in the vocabulary test. Therefore lack of vocabulary cannot explain poor performance on the Homonym Selection task.

Homonym Selection task and Object Selection control task

Figure 2 shows the number of children selecting 0 to 4 homonym pairs. Children selected a mean of 2·67 pairs out of 4 (67%) on the Homonym Selection task, and 3·13 pairs out of 4 (78%) on the Object Selection control task. Performance on both tasks improved with age: for the Homonym task, the younger group selected 50% and the older group selected 83% of homonym pairs; for the Object Selection task, the younger group selected 65% and the older group selected 92% of object pairs. Performance of both age groups on both tasks was significantly above the level expected by chance (one sample t-tests, df = 11, $p < 0{\cdot}01$ in each case). An analysis of variance was carried out over the number of pairs selected for the Homonym and Object Selection tasks, with the two age groups as a between subjects factor and tasks as a within subjects factor. Age group was significant ($F(1, 22) = 6{\cdot}38$, $p = 0{\cdot}019$) but condition was not ($F(1, 22) = 1{\cdot}74$, $p = 0{\cdot}20$) and there was no interaction ($F(1, 22) = 0{\cdot}13$, $p = 0{\cdot}722$). (Using a MANOVA with tasks as dependent variables produces precisely the same result, since there are only two tasks.)

Despite the lack of significant difference between overall level of performance on the two tasks, the pattern of responses was quite different. For the Homonym Selection task, 21 out of 24 children selected either all four

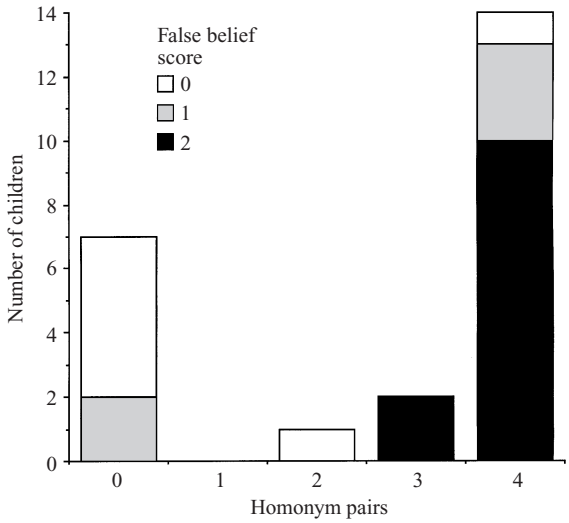

Fig. 2. Number of children selecting 0–4 homonym pairs and showing false belief understanding in Experiment 2.

pairs or none (see Figure 2). This suggests that children's difficulties with this task are conceptual. For the Object Selection control task, however, 20 out of 24 children selected 3 or 4 pairs, and the 8 children who selected 3 pairs only failed one of the 8 trials. Nevertheless, performance on the control task will be partialled out from the later analysis relating performance on the false belief and Homonym Selection tasks.

False Belief tasks

Thirteen children passed the 'Sally' False Belief task (54%) and 16 children passed the 'Puppet' False Belief task (67%). Although the 'Puppet' task seems to have been slightly easier, this was not significant (Binomial, $n = 5$, $k = 1$, $p = 0{\cdot}376$). Each task shows non-significant improvement with age: 'Sally' task, younger children 42% correct, older children 67% correct, Fisher's exact, $p = 0{\cdot}2068$; 'Puppet' task, younger children 58% correct, older children 75% correct, Fisher's exact, $p = 0{\cdot}33$. The youngest child failed the 'Puppet' memory control question, and failed all other tasks. No child failed the 'Sally' memory control question. Since the two tasks were highly correlated ($r = 0{\cdot}59$, $p = 0{\cdot}002$) and did not differ significantly, for comparison with the Homonym Selection task they were combined to give a False Belief score from 0 (fail both) to 2 (pass both).

https://doi.org/10.1017/S0305000900004153 Published online by Cambridge University Press

Comparison of tasks

Figure 2 also compares performance on the False Belief tasks with the number of homonym pairs selected. Table 5 gives further details of the

TABLE 5. (a) *Ranges, means, standard deviations, and intercorrelations between variables of Experiment 2.* (b) *Correlations between Homonym, Synonym, and False Belief tasks, partialling out verbal mental age (BPVS) and age*

(a)

Variable	Range	Mean	S.D.
Age	3;6–5;2	4;2	6 m.
BPVS	2;2–7;3	4;9	18 m.
Homonym	0–4	2·67	1·81
Object Selection	0–4	3·13	1·23
False Belief:			
'Sally'	0–1	0·54	0·51
'Puppet'	0–1	0·67	0·48
Combined	0–2	1·21	0·88

	Object Selection	False Belief	BPVS	AGE
Homonym	0·45*	0·73***	0·33	0·46*
Object Selection	—	0·46*	0·37	0·43*
False Belief		—	0·21	0·39*
BPVS			—	0·24

(b)

	Synonym	False Belief
Homonym	0·26	0·66***
Object Selection	—	0·33

*$p < 0{\cdot}05$; **$p < 0{\cdot}01$; ***$p < 0{\cdot}001$.

experimental variables, and the correlations between them. Clearly performance on the Homonym and False Belief tasks is strongly related, $r = 0{\cdot}73$, df = 22, $p < 0{\cdot}0001$. As shown in Table 5(b), it remains substantially and significantly related even after performance on the BPVS and Age have been partialled out. This is also the case if performance on the Object Selection control task is partialled out in addition to age and BPVS score, $r = 0{\cdot}63$, df = 19, $p < 0{\cdot}01$.

The levels of children's performance on the two tasks are very similar: children were 61 % successful on the False Belief task, and 67 % successful on the Homonym Selection task. Because they are on different scales the two tasks cannot be compared directly, but if scores of 2 or above on the

Homonym task and 1 or 2 on the False Belief task are counted as success, the two tasks do not differ in difficulty: Binomial, $n = 4$, $k = 2$, $p = 0{\cdot}688$.

In Experiment 1, performances on the Homonym and False Belief tasks were only associated for the older children, the children who performed well on the False Belief task. In the present experiment, however, the association was equally strong for each age group. For the older children the correlation between the Homonym and False Belief tasks was $r = 0{\cdot}74$, df = 10, $p < 0{\cdot}01$. When BPVS score and age were partialled out, the correlation remained significant, $r = 0{\cdot}70$, df = 8, $p < 0{\cdot}05$. For the younger children, the correlation between the two tasks was $r = 0{\cdot}69$, df = 10, $p < 0{\cdot}05$. When BPVS score and age were partialled out, the correlation again remained significant, $r = 0{\cdot}68$, df = 8, $p < 0{\cdot}05$. This comparison is meaningful since, although children in Experiment 2 were older than children in Experiment 1, the age groups are roughly matched in terms of performance on the 'Sally' False Belief task: the younger groups were 33 and 42% correct in Experiments 1 and 2, respectively, and the older groups were 75 and 67% correct respectively.[1] Thus, with the improved methodology of Experiment 2, even for groups which perform poorly on the False Belief task, the Homonym and False Belief tasks are associated.

Discussion

The results of Experiment 2 confirm those of Experiment 1: there is rapid improvement in children's ability to understand homonymy around the age of four years. Again, the failure of younger children cannot be attributed to lack of vocabulary. Furthermore, the possibility of false positives inherent in Experiment 1 and previous studies was removed. The improved methodology of Experiment 2 requires children to coordinate sameness of linguistic form with difference in meaning. In doing so, children must represent the relationship between linguistic form and what it represents. This ability is central to definitions of metalinguistic awareness, discussed below.

In Experiment 1, the younger children performed poorly on the False Belief task, but less poorly on the Homonym task, and for the younger group the two tasks no longer correlated significantly when age and BPVS scores were partialled out. This raised two possibilities: (1) that younger children had some understanding of homonymy prior to understanding false belief; (2) that there was a source of false positives on the Homonym task inflating performance especially amongst younger children and weakening the as-

[1] It is common for there to be variations in the age at which children pass the false belief task, depending on where the sample was drawn from. For example, in the very first false belief study, Wimmer & Perner (1983) found only 50% of five-year-olds were successful. In the present study, variations in ability were probably a result of the children in Experiment 1 being predominantly middle class, and the children in Experiment 2 being predominantly working class.

sociation with the False Belief task. When this possibility of false positives was removed in Experiment 2, performances on the Homonym and False Belief tasks were more comparable and associated more strongly, even in the younger group (roughly matched in terms of false belief performance with the younger group of Experiment 1). This suggests that the ability to understand homonymy is not present prior to the ability to understand false belief.

The less than perfect performance on the Object Selection control task is slightly surprising. However, although performance was not significantly better than on the Homonym task, the patterns of performance were quite different. Since most children passed all trials or only failed one, the problems with this task seem to result from occasional lapses in attention. Thus this task served its purpose as a control – since almost all the children could clearly understand the task demands and make the necessary discriminations, these difficulties cannot account for failure on the Homonym task. Furthermore, the association between the control task and the Homonym task was very low, and did not persist beyond a common association with chronological and verbal mental age.

GENERAL DISCUSSION

The main findings of this study are as follows:

(1) The ability to select homonyms, or judge others' selection of homonyms, arises around the age of four years. This accords well with the findings of Peters & Zaidel (1980) and Backscheider & Gelman (1995). This study extends previous research by showing that lack of the necessary vocabulary is not the cause of children's selection difficulties. The addition of analogous control conditions also shows that children's problems were not to do with the process of selection – for example, poor search skills or more general difficulties with the format or understanding of the task demands. The methodology developed in Experiment 2 also precludes a possible source of false positives inherent in the design of Experiment 1 and previous studies.

It should be borne in mind, however, that the experimental design provides children with extensive assistance. They are primed with both referents of homonyms in the vocabulary checks, and they are simultaneously shown both exemplars in the test phase. In everyday life, without such aids, identifying homonyms may be more difficult.

(2) The ability to select homonyms does not arise in isolation. It is strongly related to the ability to select synonyms, and more surprisingly to the ability to pass the false belief task. These relationships persist beyond a common association with chronological and verbal mental age.

The results help discriminate the possible accounts of children's difficulty with homonymy. Three general accounts were outlined: the metalinguistic

deficit account, in which younger children lack the conceptual understanding necessary to understand homonymy; the metalinguistic one-to-one mapping account, in which children possess the conceptual understanding but make an erroneous assumption about the relationship between words and their referents; the implicit one-to-one mapping account, in which children simply fail to learn homonyms.

The implicit one-to-one mapping account cannot account for the results of the study. Because the experimental homonyms were demonstrably in children's vocabularies, poor performance cannot be a result of a tendency not to learn homonyms. It may nevertheless be true that children (and even adults) ARE biased not to learn homonyms, but once they have learned them the younger children have some additional source of difficulty.

The other two accounts are harder to distinguish between, since both can account for the relationship between understanding of homonymy and false belief. The metalinguistic deficit account predicted this association based on Doherty & Perner's (1998) hypothesis that each of these tasks requires a distinction between what is represented and how it is represented (Perner, 1991, 1995). According to this hypothesis, younger children fail both tasks because they are unable to make this distinction. However, this association could potentially also be explained by the metalinguistic one-to-one mapping account. For example, Flavell (1988) suggested that theory of mind tasks and mutual exclusivity tasks both call for the understanding that objects can be represented in more than one way. Prior to this children could assume a one-to-one correspondence between words and their referents, and between thoughts and the world. This assumption could account for the association between performance on the false belief task and the homonym task.

If children really did think that one thing could be represented in only one way, they would be expected to have difficulty with the vocabulary checks. However, in the synonym vocabulary check, all children were willing on the same occasion to point to an item when identified under one and then the other synonym; in the homonym vocabulary check of Experiment 1, children were willing to point to different objects identified under the same name. Had children been using a one-to-one mapping assumption, they might be expected to reject one label in the synonym vocabulary check, and reject one of the referents in the homonym vocabulary check.

However, it is also possible to argue that during the vocabulary checks children with a one-to-one mapping assumption were put in a position where they had no choice other than to accept violations of that assumption. During the test phase they were free to reassert this assumption and consequently selected or judged incorrectly. The vocabulary data are therefore not conclusive.

The one-to-one mapping assumption should also have lead to characteristic error patterns in the test phases of the Homonym and Synonym Judgement

tasks. In the Homonym Judgement task, assuming that the homonym only applies to one of the objects would lead children to endorse Puppet's choice of the object they themselves indicated and to reject Puppet's choice of the homonym or an unrelated item. Only one child showed this pattern. Similarly for the Synonym task, children would be expected to endorse Puppet's repetition of the term they themselves had provided and reject Puppet's use of a synonym or a misnomer. Only three children showed this pattern. In the Homonym Selection task, a one-to-one mapping assumption would lead children to select the 'correct' homonym regardless of the picture on the card. Thus children failing the task would correctly select 4 single homonyms but no pairs. However, of the 7 children who selected no pairs, 3 selected no single homonyms, 3 selected only 1, and 1 child selected 3 single homonyms.

Given the small number of children failing the homonym tasks in both experiments, however, it would be unwise to make strong conclusions based on this error data. Although the present study suggests that children do not employ an explicit one-to-one mapping assumption, more data are needed to convincingly distinguish between the metalinguistic deficit and metalinguistic one-to-one mapping accounts. Direct comparison of them using homonyms is difficult, since the two theories make very similar predictions. If children are confronted with homonymy in an experimental situation, it is always possible to argue that they have a one-to-one bias but relax it (if only temporarily) given strong evidence that it is inappropriate in any given setting.

A more promising way of distinguishing them is to look for parallel developments of other metalinguistic abilities. The distinction between what is represented and how it is represented is characteristic of many abilities considered metalinguistic. One example is grammatical awareness, which requires children to understand the link between grammatical form and meaning. Data suggest that children begin to succeed on tests of grammatical awareness around the age of four years; for example, Smith & Tager-Flusberg (1982) found that whilst 78% of four-year-olds successfully judged the grammaticality of short sentences, only 22% of three-year-olds could do so.

If this ability were found to relate to children's understanding of false belief, or their understanding of homonymy and synonymy, it would provide further evidence for the development of metalinguistic awareness around the age of four years. According to the one-to-one mapping account, however, younger children have metalinguistic awareness coupled with assumptions about the exclusivity of the relationship between meaning and linguistic form. Children ought to be able to recognize ungrammatical sentences, since, apart from anything else, they violate the normal relationship between (for example) word order and meaning. Unfortunately, as many authors have

pointed out (e.g. Bowey, 1988; Gombert, 1992), successful judgements of grammaticality could be based on semantic rather than syntactic judgements: ungrammatical sentences are harder to understand. This confounding factor would have to be removed before one could be sure that children were making a metalinguistic judgement.

Regardless of whether children employ a metalinguistic one-to-one mapping account or not, their word learning may still conform to one-to-one mapping or mutual exclusivity biases. Even adults, in the absence of evidence to the contrary, are likely to employ such biases when learning novel words However, clearly words, such as homonyms and synonyms, that violate these assumptions ARE learned, so these biases do not preclude the learning of words that violate the one-to-one mapping assumption.

CONCLUSION

The results of the present study show that despite possessing the necessary vocabulary, until the age of roughly four years children are unable to identify both items from a homonym pair. More importantly, the ability to understand homonymy develops at the same time, and is closely associated with the ability to understand false belief. This extends and replicates Doherty & Perner's (1998) finding that the ability to understand synonymy is closely associated with the ability to understand false belief. The results suggest that prior to this age children do not possess the metalinguistic awareness necessary to represent the relationship between words and their referents, which is needed to understand homonymy. However, it is also possible that children have the necessary metalinguistic awareness but make a simplifying assumption that words and their referents bear a one-to-one relationship, and so discount the possibility of homonymy. More general research on the development of metalinguistic awareness should help to further distinguish these two accounts.

REFERENCES

Backscheider, A. G. & Gelman, S. A. (1995). Children's understanding of homonyms. *Journal of Child Language* **22**, 107–27.

Beveridge, M. & Marsh, L. (1991). The influence of context on young children's understanding of homophonic words. *Journal of Child Language* **18**, 459–67.

Bowey, J. A. (1988). *Metalinguistic functioning in children*. Geelong, Australia: Deakin University.

Campbell, R. N. & Macdonald, T. B. (1982). Text and context in early language comprehension. In M. D. Donaldson, R. Grieve & C. Pratt (eds), *Early child development and education*. Oxford: Blackwell.

Doherty, M. J. & Perner, J. (1998). Metalinguistic awareness and theory of mind: just two words for the same thing? *Cognitive Development* **13**, 279–305.

Dunn, L. M., Dunn, L., Whetton, C. & Pintillie, 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. W. Astington, P. L. Harris & D. R. Olson (eds), *Developing theories of mind*. New York: Cambridge University Press.

Gombert, J. E. (1992). *Metalinguistic development*. London: Harvester Wheatsheaf.
Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development* **66**, 843–55.
Mazzocco, M. M. M. (1997). Children's interpretations of homonyms: a developmental study. *Journal of Child Language* **24**, 441–67.
Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press/ Bradford.
Perner, J. (1995). The many faces of belief: reflections on Fodor's and the child's theory of mind. *Cognition* **57**, 241–69.
Peters, A. & Zaidel, E. (1980). The acquisition of homonymy. *Cognition* **8**, 187–207.
Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (ed.) *The crosslinguistic study of language acquisition*. Vol. 2. Hillsdale, NJ: Erlbaum.
Smith, C. L. & Tager-Flusberg, H. (1982). Metalinguistic awareness and language development. *Journal of Experimental Child Psychology* **34**, 449–68.
Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–28.