Metacognitive development and the disambiguation effect in monolingual and bilingual children

Cornelia Gollek

Department of Psychology

University of Stirling

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Table of Contents

Acknowledgements

When writing my undergraduate thesis, the advice I was given was not to spend hours contemplating who to thank in the acknowledgements. Gratitude should be shown in person and as this is a report and not an Oscars speech, I would like to heed this advice and keep it simple.

Thank you, Martin, for always being an approachable supervisor. Your input was invaluable, your guidance has been key to my progression and your assistance has helped me greatly to finish my thesis.

The time I spent at the University of Salzburg was delightful. I would like to thank everyone who took time to help me, especially Eva.

I am very lucky to work in a research area where participants are always cheerful and motivated, so thank you to all the children. I would also like to thank all the nurseries who worked with me over the years, regardless of how busy or understaffed they were. A special thank you goes to all pre-school staff asking for advice or showing some interest in my work.

Abstract

Background

Research suggests that children are only able to flexibly apply more than one label (e.g. mouse and animal) in one situation with one conversational partner after they pass standard false belief tasks. Both abilities have been attributed to the understanding of perspective

Aim

The aim of the studies was to extend previous research to examine the disambiguation effect, children's tendency to select an unfamiliar object in the presence of another but familiar object as referent for a novel word. Theoretical considerations suggest this effect initially results from a lack of understanding perspective.

Method

Five studies were conducted in Scotland and Austria, involving 243 children between the ages of 2.5 and 6.5. Studies 1 to 3 compared the standard disambiguation task with a task in which a strong pragmatic cue indicates the familiar object is the correct referent. Performances on these tasks were compared with performances on the false belief task, the alternative naming task, as well as tests of executive functioning. Studies 4 and 5 extended these methods to examine bilingual children's metacognitive abilities in relation to word learning.

Results

Children become able to suspend the disambiguation effect when presented with strong pragmatic cues at the same time as they pass false belief and alternative naming tasks (Experiment 1). This can neither be attributed to impulsivity or the ability to inhibit a response, nor order effects of pragmatic cues and novel words (Experiment 2). Children's ability to apply two labels to one object in a correction task also related to their perspectival understanding. Previous findings that suggested that younger children could produce multiple labels in a misnaming paradigm were not replicated (Experiment 3 a, b). The developmental change in children's metalinguistic behaviour was demonstrated to follow the same trajectory in monolinguals, bilinguals and children exposed to another language (Experiment 4 and 5). Bilinguals show a marginally better ability to recall novel foreign language labels.

Conclusion

The disambiguation effect is the result of cognitive immaturity in young children. Older children show a change in behaviour at the same time as they present more metacognitive maturity. Common development with theory of mind and metalinguistic abilities is attributed to an understanding of perspective.

Introduction

The following thesis addressed the relationship between metarepresentation and word learning. Children between the ages of 3 to 5 were of special interest as changes in metarepresentational abilities have been observed at this time in development. The studies presented investigated possible changes in metalinguistic development to address if these occur at the same time. The mutual exclusivity bias, a proposed word learning phenomenon, was examined. Two different theoretical accounts, namely lexical principles and sociopragmatic approach, and their explanation around mutual exclusivity were analysed. The perspectival account of metacognition was applied to children's word learning. Based on this account, it was suggested that mutual exclusivity is firstly due to conceptual immaturity, of which any word learning benefits are an incidental consequence. The immaturity is overcome once children gain a level of metacognitive understanding which is indicated by passing false belief tasks. Later observations of mutual exclusivity behaviour can be explained by logical reasoning.

The second part of the thesis investigated bilingual children's metarepresentational and metalinguistic development. Previous research around bilingual theory of mind and the disambiguation effect was presented. Possible advantages bilingual children might have due to constant exposure to multiple labelling were addressed.

The conclusion was that the disambiguation effect in young children is the result of cognitive immaturity which restricts children from considering multiple labels for one object, even when given strong pragmatic cues to do so. Once children understand perspective they are able to accept two labels for one object, when there are pragmatic cues to this being appropriate. This ability is demonstrated by their performance on the pragmatic cue (PC) task.

The findings were well supported by the perspectival account's predictions and account for word learning phenomena between the age of 3 and 4 for monolingual and bilingual children equally.

Metarepresentation and theory of mind.

Metarepresentations were defined as "The process of ´representing the representational relation itself` (Pylyshyn, 1978), or representing a representation as a representation (Perner, 1991)." (Doherty, 2009, p. 214).

A representation can take the form of a statement or a thought about something or a picture of something. A metarepresentation is then, for example, a statement about a statement, a thought about a thought or a picture of a picture. Metarepresentation is most commonly discussed in the context of theory of mind as the critical ability required to pass the standard false belief task. In relation to this it is important to note that representations include not only mental but also public representations (like utterances, pictures, etc.) (Sperber, 2001). Hence, the mature ability to metarepresent includes both understanding of thoughts and understanding of non-mental representations such as language. The thesis extended this comparison.

The term "theory of mind" was first discussed in connection with research of chimpanzees and the question whether this species reasons over other's differing beliefs, thoughts and desires (Premack & Woodruff, 1978). This lead to the investigation of theory of mind in developmental psychology and to the false belief task which assess children's understanding of other's mental states. The original version for children by Wimmer and Perner (1983) has been adapted many times; Translated from the German original, a story is acted out with several toys as follows:

Maxi has a chocolate bar. He puts his chocolate bar in the kitchen cupboard and leaves the kitchen to go and play. His mother comes into the kitchen. She needs some chocolate for baking. She takes Maxi's chocolate bar from the cupboard and puts it into the drawer after she has finished baking. She then leaves the kitchen and Maxi comes back for his chocolate bar. - Children are then asked: "Where will Maxi look first for his chocolate?" (after Wimmer & Perner, 1983).

Children need to understand that Maxi's mental state or belief about the location of the chocolate bar differs from reality to pass the false belief task and answer the question correctly. While Maxi thinks the chocolate is in the cupboard, he is misrepresenting the location of the chocolate. When children are able to recognise that Maxi can have a belief that differs from reality, they display metarepresentational understanding. They understand that other's thoughts or representations can be seen as that and evaluated against reality. Children around the age of four years seem to be able to do this and pass the false belief task, predicting belief based behaviour correctly. Hence, the ability to represent someone else's belief in one's own mind and make belief based judgements require metarepresentational skills.

Metalinguistic awareness on the other hand indicates the understanding of language as a carrier of meaning (Doherty & Perner, 1998). As reviewed in Doherty (2009) there are traditionally two different hypotheses regarding the development of metalinguistic awareness. The interaction hypothesis assigns an initial awareness of language as a carrier of meaning to children (e.g., Clark, 1978). Children need this awareness to learn to speak and develop it further while growing their understanding of language. The autonomy hypothesis declares that children do not need an understanding of language when they start to speak (e.g., Gombert, 1990/1992). Metalinguistic awareness develops later, around the age of five to seven.

Doherty (2009) pointed out that most definitions combine conceptual understanding of language as a carrier of meaning and the executive functioning required to exercise this understanding. According to Doherty (2009, p. 82): “The fundamental difference between theory of mind and metalinguistic awareness may be only in the domain of application: language or mind”. The traditional false belief task assesses the understanding of the representational character of mental states. In the unexpected-transfer task, the understanding that the protagonist’s belief misrepresents reality is necessary. Assessing Doherty’s claim requires a task that assesses the understanding of the representational character of words.

Doherty and Perner (1998) investigated the connection between metalinguistic awareness and theory of mind. Both metalinguistic and metarepresentational awareness might follow the same developmental pattern and be due to a common conceptual change. The authors turned to the use of synonyms in the search of a task examining linguistic metarepresentation in children. Two different names for one object like “bunny” and “rabbit” can refer to the same thing. Being aware that two different labels can be applied to represent the same referent indicates metalinguistic awareness. Doherty and Perner (1998) presented three to five year old children with a synonym judgement task. Children were told two names for a picture (for example mug and cup) and then asked to choose one. A puppet was then asked to say the other name for it then and children had to judge whether puppet was correct. Children's ability to correctly judge puppet's performance strongly correlated with their ability to pass a standard false belief task. In their second study, children were asked to produce a second name for the picture, which was also strongly associated with false belief understanding. This was seen as indication for metalinguistic abilities, indicated by the understanding that two labels can represent one referent, and metarepresentational ability, indicated by the understanding that different beliefs can represent and misrepresent reality at the same time, to be grounded in a shared conceptual understanding.

These findings were later extended with categorical labels (e.g. rabbit & animal; Perner, Stummer, Sprung, & Doherty, 2002). Children performed a similar task, but had to produce and judge puppet's production of alternative names for a referent. Results were similarly to the synonym tasks. Children were able to produce alternative names at the same time in development as they were able to pass false belief. Categorical labels do not share the same meaning in the way synonyms do, though. The application of metarepresentation in this example needed to be revised.

Perspectival understanding.

Perner (2000) explained this as an understanding of perspective necessary in both theory of mind and metalinguistic awareness. Perspective is a property of representation. Every representation of an object represents the object in a specific way. Different representations of the same object always differ in the perspective on that object. The easiest explanation comes from visual perspectives. Whether an object is seen from the front or the back depends on the point of view. Different places afford different points of view. Hence, many different perspectives are possible on the same object in one moment in time. The representational understanding of the mind became a perspectival understanding of the mind.

In the false belief task, both protagonists have different beliefs about the location of the chocolate bar or different perspectives on looking at reality. The shift from metarepresentational to perspectival understanding was therefore applicable here. The understanding of perspective is also relevant to the understanding of naming. Psycholinguists claim that different labels for a referent always differ in perspective (Clark, 1987; Tomasello, 1999).

This is evident when categorical labels are considered. If a referent is, for example, labelled as a mouse, the label restricts the possible judgements which can be made. It could possibly be described as “big” compared to other mice. Applying the label “animal” and comparing it to other animals, it is still quite small. So the mouse/animal is big or small depending on which label is chosen, which perspective is taken (Doherty & Perner, 2013). The understanding of alternative naming therefore requires the understanding of perspective.

Perner (2000) explains how children develop from taking a perspective to switching a perspective and finally to understanding perspective. The first step is simply having a mental state - representing something in thought, utterance or other way. The next step is switching between perspectives, starting by being externally induced. Studies by Flavell and colleagues described this ability clearly (Flavell, Everett, Croft, & Flavell, 1981). Children between the age of three and five were shown a picture of a turtle by the experimenter, who sat opposite the child. Children were able to switch between two perspectives when the picture was displayed facing them correctly or upside down. They answered accordingly, whether the turtle was on his feet or on his back. Only the older children were able to express what the experimenter saw. The ability to switch displayed by younger children is externally driven. The experimenter turns the picture and children switch their representations. Older children have control over the ability to switch, they understand perspectival differences. Hence, they can express what the experimenter sees and what they see at the same time.

The perspectival account is therefore able to explain developmental change presented by passing false belief and alternative naming. Children show a difficulty producing two familiar labels for one referent until they understand perspective. A conceptually similar task to the alternative naming task, the disambiguation task looks at children's tendency to avoid two labels for one referent. In the disambiguation task, children are presented with one familiar and one unfamiliar object and asked to choose a referent by a novel word. Children tend to select the novel item, showing a reluctance to consider the novel word as an alternative label for the familiar object. Similarly to the alternative naming task, children avoid two labels for one object. They behave as if nouns were mutually exclusive.

The question arises whether children also present a conceptual change when faced with the disambiguation task. Do younger children pick the novel object more often than older children? Is there a connection to perspectival understanding? There should be at first glance, as both tasks seem to have similar demands. The alternative naming task requires the child to flexibly apply two labels to one object. The disambiguation task would also require the application of two labels to one object, if the child considers the familiar object as referent. However, the studies reviewed in the following do not show evidence of a developmental change around the age of four in disambiguation behaviour.

There are two plausible reasons why this might be the case. On the one hand, this has simply not been looked at in general. On the other hand, choosing the novel item in the disambiguation task as referent is the more sensible choice if no other information is displayed. Adults faced with the same task would choose the same way. Following the perspectival account, mutual exclusivity can be interpreted as metacognitive immaturity in children. Young children select the novel object, as they cannot accept a second label for an already named object, similarly to their behaviour in the alternative naming task. Older children and adults choose the novel object, as they assess the information provided and select the sensible referent. The result is the same. The cognitive path towards it is different. Mutual exclusivity might be a conceptual difficulty expressed by young children during word learning.

Further word learning theories.

Other theoretical accounts aim to explain mutual exclusivity assumptions differently. Supporters of the lexical principles approach describe the mutual exclusivity bias as integral to children's word learning (e.g., Markman & Wachtel, 1988; Merriman & Bowman, 1989). Children's assumption that word meaning cannot overlap is supposed to limit possible referents for a novel word and so benefit word learning. The mutual exclusivity bias has been hypothesized to be one of several lexical principles children employ in word learning (Golinkoff, Mervis, & Hirsh-Pasek, 1994; Merriman & Bowman, 1989).

Markman and Wachtel (1988) presented a series of studies around the mutual exclusivity bias. Children between the age of three and four tended to select the novel item as referent for the novel word at ceiling level in the standard disambiguation task. When children were shown only familiar items and a novel word, they tended to ascribe the label to a part or the substance of the familiar item. When shown a novel item and a novel label, children tended to ascribe it to the whole object. This was again seen as support for the assumption that labels are mutually exclusive.

Merriman and Bowman (1989) described several potential effects relating to mutual exclusivity. The disambiguation effect explained above is the most heavily researched. They further named the correction effect, which describes the removal of a familiar term in favour of a novel term for an item. The correction effect is the only other effect which has gained some attention in the research field (Savage & Au, 1996). A further effect, the rejection effect was defined as refusal to accept a novel term for an item. Merriman and Bowman (1989) stated that if a child is adhering to the mutual exclusivity bias, one of these three effects will be observed when new terms are introduced. The child will also always display a restriction effect: a new term cannot be generalized as additional label for an already named object. An object cannot have two overlapping terms. The authors explained different views on when the mutual exclusivity bias emerges: at the start of word learning, during early childhood or, supported by opponents of the theory, never; Merriman and Bowman (1989) described the bias to develop throughout early childhood as a useful word learning heuristic. They argued, the bias would restrict infants in learning new words quickly if it would emerge with word learning and is of real use after some vocabulary has developed.

Golinkoff et al. (1994) supported the idea of a general lexical development occurring first. They also argued that children are driven by a desire to map a novel term onto a nameless category (N3C principle) rather than the desire not to map a novel term on an already named category. Children might further employ principles more or less strongly depending on their memory processes and present individual differences in preferences for employing certain principles (Merriman & Lipko, 2008).

Supporters of the sociopragmatic approach present an alternative view of children's decision making in the disambiguation task. Here, children determine word meaning by interpreting speaker intention (Bloom, 2000; Diesendruck & Markson, 2001). Speakers around them give enough pragmatic information to enable children's learning of novel words. The following principles are claimed to explain object choice in the disambiguation task:

Clark's *Principle of Contrast*: "whenever there is a difference in form of a language, there is a difference in meaning (Clark, 1987, p. 1) combined with the *Principle of Conventionality*: "For certain meanings, there is a conventional form that speakers expect to be used in the language community" (Clark, 1988, p. 319)

Presented with an apple and a whisk and asked to point to the “hinkle”, a child might reason that if the speaker wanted to refer to the apple, she would have said “apple”, which is the conventional form. Because she said “hinkle”, she must mean something else, and the whisk is a sensible choice. Problematic is the use of term “meaning” instead of reference. Apple and “hinkle” differ in meaning but could still refer to the same object, similar to, for example, apple and fruit. Following the *Principle of Conventionality*, which speaks of meaning, this does not generally prohibit the selection of already labelled objects as referents in the disambiguation situation (Doherty & Perner, 2013).

If conceptual changes can be observed in perspectival understanding around the age of four, changes might be observable around children’s apparent mutually exclusivity assumption. The similarities between alternative naming difficulties and mutual exclusivity would point towards this. The following presents previous findings around apparent mutual exclusivity assumptions across the age range. So far, the literature only contains one study showing an age related change while also implicitly pitting predictions of lexical principles and sociopragmatic cues against each other: Haryu (1991) is addressed in detail.

Disambiguation studies addressing lexical principles explanations.

Studies looking at the emergence of an apparent mutual exclusivity bias in development used a variety of set ups. One of the most frequent methods with infants is the preferential looking technique. Halberda (2003) presented infants between 14 and 17 months with sets of a familiar plus an unfamiliar picture. Children were then told to look at the “dax”. Results showed 14 month old infants to look systematically at a familiar object when hearing a novel word. Infants at 16 months performed at chance level. Only from 17 months did infants systematically look at the novel object above baseline. The author concluded that infants learn words initially without a word learning strategy, as the results pointed towards stable mutual exclusivity bias only in the older age group.

Mather and Plunkett (2010) explained that the 10 month old children in their studies prefer the novelty of objects and labels, which acts as a precursor for the development of mutual exclusivity. They showed infants sets of familiar and novel objects and told them to look either at the familiar item or at a novel referent ("Look at the mido!"). Initially, infants did not show any looking preferences. After the third trial, infants looked significantly longer at the novel object in response to the novel label than in response to a familiar label. The authors concluded that infant's interest in novel items was enhanced throughout the trials because of the novel labels. They mentioned that infants in their study probably did not know the familiar terms for all objects and could therefore not display a strict mutual exclusivity bias, which would exclude the familiar item as it is already named. Infants looked at the novel item in response to the novel word because of a preference for novelty. This was explained as initial step towards the development of mutual exclusivity.

Another study supporting children's preference for novelty when trying to map novel words presented 22 month old infants with sets of three items, one familiar and two novel (Mather & Plunkett, 2012). One of the novel items had previously been familiarized to the child, without labelling it. Children were then asked to look at the "dif". The not familiarized novel item was looked at with higher preference in response to the novel label. The authors suggested that children's disambiguation was driven by the search for complete novelty after hearing a novel label. The preference for looking at the super-novel object upon hearing a novel label was therefore not driven by knowing a label for the other novel object. Also, hearing novel labels increased children's preference for looking at novel-novel labels over the course of trials.

A study with 18, 24 and 30 month olds found that the youngest group did not prefer novel objects as referents for novel words (Bion, Borovsky, & Fernald, 2012). Different to the studies presented before, 18 month old children disambiguated at chance level, the other two age groups displayed clear disambiguation. Children were then tested on retention of the novel word, presenting the familiarized novel next to an unfamiliar novel item. The oldest group performed just above chance, both younger groups looked randomly. It was concluded, that disambiguation is related to word learning, but not essential at the beginning. It develops and increases with growing vocabulary and lexical experience.

In a slightly different set-up, with only a familiar object visible and a bucket as possible location for more objects, children between 15-17 and 18-20 months resisted the familiar object when asked for a novel label and searched for another possible object (Markman, Wasow, & Hansen, 2003). The younger children searched even more for an alternative in a variation of the task where the bucket was removed and only the familiar object was in sight. The authors concluded that infants in their tasks applied mutual exclusivity. Without a novel object, children clearly demonstrated a reluctance to select the familiar item as referent for the novel word and not a tendency to fill lexical gaps, supporting the existence of a mutual exclusivity bias. The authors further mentioned that reduced task demand might have enabled younger children to demonstrate their word learning strategies more clearly.

Overall, most studies using the preferential looking paradigm agreed that infants show disambiguation around the age of 15 to 17 months. Studies with younger children found recognition of novelty but not in connection with word learning, rather displaying novelty-preference. Mutual exclusivity seems to afford some existing knowledge and novelty to bootstrap new meaning to new words (Mather, 2013).

The preference for super-novelty was further investigated with two year olds (Horst, Samuelson, Kucker, & McMurray, 2011). Children were familiarized with novel objects and then presented with set-ups of two now-familiar and one super-novel item and asked to pick by a novel label. Children presented a clear preference for the super-novel item. A variation of the task identified that novelty to the child was more important than novelty to the experimenter. Children even selected the super-novel item more often, if all objects were equally novel to the experimenter asking for a novel label referent. This seems to clearly violate any suggestions of interpreting speaker intentions as presented by the sociopragmatic account.

Suanda and Namy (2013) added a new line of support to the lexical principles account. They presented children aged 18 months with standard disambiguation tasks using novel words and additionally tasks using novel gestures. In the gesture condition, a novel object was introduced accompanied by a novel gesture. Another novel object was presented without a gesture. The child was then asked to pick an object accompanied by a new gesture. Children disambiguated clearly in the novel word task but performed around chance in the gesture disambiguation task. A second experiment ensured that children were mapping gestures onto objects. The child was asked to choose in the exact same situation, only now using the gesture or word introduced initially. Here, children selected the introduced referent on similar levels in both conditions. Therefore, children mapped gestures onto novel objects, but did not treat gestures as mutually exclusive. The authors argue that the two conditions require the same inferences about the other's referential intentions. Thus the lack of disambiguation effect in the gesture condition suggests that the standard disambiguation effect is not a result of inferences about intention, but a lexical phenomenon.

Jaswal (2010) also supported the view that children follow lexical principles to determine word meaning. Interpreting speaker intent is secondary and will overrule lexical principles only if pragmatic information is overwhelming. He presented children aged 2;6 with disambiguation tasks. Experimenters made either a neutral request (asking for "*it*") or a request using a novel word combined with either pointing or gazing towards the familiar item. Children used the pragmatic information following the neutral request but not after the label request. In a follow up, the author combined pragmatic information cues: The experimenter pointed <u>and</u> gazed towards the familiar object while requesting an object using a novel label. Now children were using the pragmatic information to overcome the mutual exclusivity bias following the author's interpretation.

Studies around the lexical principle approach seem overall to agree that disambiguation occurs after some vocabulary has been acquired. It seems to be reliably shown from around 17 to 19 months. This may be artificially reinforced by testing techniques and the problematic of finding the appropriate set-up for an age group.

Disambiguation studies addressing sociopragmatic explanations.

The sociopragmatic approach presented a variety of studies analysing the disambiguation effect. Children's object choice is here promoted to be driven by quite advanced cognitive abilities to interpret speaker intention. Diesendruck and Markson (2001) found the disambiguation effect to occur not only with novel labels, but also with novel idiosyncratic facts. The authors offered three year old children sets of two novel objects and a novel fact about one of them (e.g., "My sister gave this to me"). Then children were asked to select the referent of a different novel fact for puppet ("the one my dog likes to play with"). Children were found to disambiguate in a similar way with fact references (4.4/6 trials) as with labels (4.9/6). The authors used this as argument against the disambiguation effect being the result of a specifically lexical principle, stating similar disambiguation levels should not be observed with facts, if the mutual exclusivity bias is a pure word learning principle.

A further finding from the initial Diesendruck and Markson (2001) paper was the impact of speaker presence and absence. A puppet functioned as second experimenter and was absent when a novel label was introduced for one of two novel objects. The puppet then asked for an object upon return, using a different novel name. Children disambiguated strongly, on average 4.9/6 cases. The same set-up was used again with novel facts. An additional condition was added with puppet present when the experimenter introduced a novel fact to compare the influence of shared and unshared novel knowledge.

Children disambiguated slightly less in the condition with puppet present than in the initial study mentioned before, but on average 3.9/6 times. When puppet was absent throughout the introduction, children performed at chance level. This was seen as support for the sociopragmatic explanation, as children make judgements based on shared and unshared knowledge. Puppet has not heard the fact about the object when he asks for something using a different fact. He could want either of the objects. But puppet is expected to know the proper labels for objects, hence disambiguation in the first case mentioned.

Diesendruck (2005) extended the investigation into presence and absence of the second experimenter. He introduced the comparison of common ("a teega") and proper ("Meloo") nouns. When puppet was present while the experimenter named one picture of an unknown creature, there was no difference between common and proper nouns. Children disambiguated equally (3.0 & 2.7/3). When puppet was absent, children disambiguated to the same level (2.9/3) only when introduced to a common noun, not to a proper noun. Here, they performed at chance. Again, no interaction with age was analysed. Hence, children appeared to take speakers' knowledge into account when disambiguating novel words at three and four years of age.

Children's ability to disambiguate novel terms for parts of an object was also subject to investigation (Saylor, Sabbagh, & Baldwin, 2002). The authors argued that teaching in a more naturalistic way would include referring to the object by its familiar label and then asking for a part by a novel label. This set up proved to be the most successful for three and four year old children in identifying the referent. If the experimenter asked, for example "See this butterfly? What colour is the *thorax*?" Children were able to disambiguate the new label thorax significantly more often than when the butterfly was labelled as “thing” or not at all. In a variation, the authors replaced the initial labelling with gestures or ambiguous pointing. Children disambiguated the novel part label significantly less in these conditions. This showed children using referential gesture to disambiguate novel words equally, when gestures were salient enough to provide reference. The authors concluded that children’s ability to read speaker intention is fundamental to word learning.

The disambiguation paradigm was further employed with younger children in many variations. Grassmann, Stracke and Tomasello (2009) presented two-year-olds with a novel object and an experimenter looking at the object while excitedly announcing a novel word. If children had not seen the experimenter with the object before, they interpreted the novel word as reference for the object through object-directed actions. If they had seen the experimenter play with the object before, children searched for another referent. This was interpreted as children being aware that the experimenter would not show this much excitement towards a familiar object. The object would be “old news” and not provoke an enthusiastic response.

Grassman and Tomasello (2010) investigated how much pragmatic information like pointing would add to children’s object choice. They presented two and four year olds with a standard disambiguation task. The experimenter then asked for a “modi” while pointing to the familiar item. Children followed the pointing over the lexical information, when the experimenter also gazed in the pointed direction. Different to Jaswal (2010), pointing without gaze resulted in choice at chance level.. When the experimenter used familiar lexical information (asked, for example, for the car) but pointed to the novel object, children again followed the pointing. Children chose randomly between following pragmatic or lexical information when the experimenter used a familiar word and pointed to a different familiar object. The authors concluded that in first instance children rely on pragmatic information like pointing to resolve ambiguity.

Brand (2000) employed an experimental set up, using repeated labelling and eye gaze for 12, 19 and 24 month old infants. This included one interesting and one boring toy, with eye gaze either coinciding with the interesting toy or creating conflict by indicating the boring toy. The 12 month old children focused on the interesting toy regardless of any cueing provided. By 19 months, children looked at the boring toy when indicated through eye-gaze, but still focused the majority of time on the more salient toy. The 24 month old children were able to focus their attention in line with eye gaze of the experimenter on the target. In a variation of the task, the two objects swapped places before the test trial. Now only the 24 month old children recognised the target. The author argued, that word learning is best observed following the child's interest. Then, even 12 month old infants might show cue sensitivity and disambiguate accordingly. Children react therefore not strongly to speaker intention at the beginning of word learning, which would constrain their ability to acquire novel words at a young age.

Children were repeatedly found to use some of the referential information provided by speaker and context. Problematic for the sociopragmatic account were findings by Scofield and Behrend (2007) who analysed age effects for label and fact disambiguation. They presented two, three and four-year olds with a set of two novel objects and introduced one of them by either a novel label or a novel fact. They then asked the children to choose one object by either a novel label or another novel fact, hence four conditions: word/word, word/fact, fact/word, fact/fact. Children across the age groups disambiguated at similar levels for word probes (word/word: 41/49 children, fact/word: 35/49). The fact probe was not different from chance (word/fact: 29/49, fact/fact: 25/49) and showed developmental changes. The two year olds performed below chance in these conditions, whereas the three- and four- year olds disambiguated to the same extent as with word probes. This was interpreted as direct conflict for the sociopragmatic account, which does not account for developmental change. Words and facts should be disambiguated to the same extent by two year olds as by three and four year olds, if children interpret speaker intention to determine word meaning.

Autism and the theories.

Children with autism have also been presented with the disambiguation paradigm. Their difficulties with social cues were assessed in relation to sociopragmatic word learning predictions. Children with autism have been shown to find it difficult to read social cues. Hence, reading speaker intention should be measurably harder for them, which puts the sociopragmatic theory to test.

A group of children with autism and a group of toddlers, matched to their vocabulary, were presented with standard disambiguation tasks (Preissler & Carey, 2005). The children with autism performed extremely similarly to the younger toddler group. The authors then assessed children's ability to judge referential content through eye gaze. Experimenter and child were both holding a novel object. The experimenter labelled one object and cued the referent with eye gaze. Children were then asked to find the newly labelled target in a set of four objects. Toddlers performed correctly on the majority of cases. Children with autism were more likely to select the object they had held during the labelling phase. In comparison to their good ability to disambiguate novel words in the first study, children with autism seem to struggle to determine word meaning when reading speaker intention is necessary.

A study with children and adolescents with autism (aged eight and 15) and typically developing matches compared performance on disambiguation situations when novel words and facts were included (de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011). Participants were presented with standard disambiguation tasks, including two novel objects and two novel labels. In a variation, the first object was only described with an idiosyncratic fact ("My uncle gave this to me"). Then a request was made, using a different fact ("Give me the one my dog likes to play with"). Participants disambiguated clearly in both conditions (similar to Scofield and Behrend's (2007) three and four year olds). The effect with facts was less strong, but significantly different from chance and did not differ between participants with ASD and their typically developing counterparts. This would speak against mutual exclusivity being a lexical principle, as disambiguation with facts would not guide word learning. However, the authors were able to present, that word and fact disambiguation were not related. Word disambiguation related strongly to receptive vocabulary size, which supports the idea as a word learning heuristic. Fact disambiguation did relate to communicative skills, which in turn did not relate to word disambiguation. Facts seem therefore to afford pragmatic interpretation abilities and the underlying mechanisms described here could be independent.

Overall, the studies supporting a sociopragmatic approach to word learning focus heavily on reading additional cues like pointing or eye gaze in disambiguation situations. Age related behavioural change was presented, strongly connected to children's general development and their abilities to follow task demands. The sociopragmatic account can not explain why fact and word disambiguation are unrelated. Both clearly exert speaker intention to resolve ambiguity. The lexical principles account has difficulties explaining why developmental tendencies are observed in the application of mutual exclusivity. If the bias is applied to benefit word learning, it should also not be observed in at least similar form with facts and gestures unless this proves to be unrelated.

Introducing a direct test of the predictions.

Haryu (1991; Haryu & Imai, 1999) presented a direct test of the predictions of the lexical principles and socio-pragmatic accounts. She presented 3.5, 4.5 and 5.5 year old Japanese children with a classic disambiguation paradigm, for example involving one familiar object (e.g., an apple) and one unfamiliar object (a lipstick case). Children were then instructed to hand a puppet one of the objects using a novel term (word only condition). In another condition (word + pragmatic context condition), children were additionally given a strong pragmatic cue indicating that the puppet desired the familiar object: "Mary is hungry. I would like to give Mary (the) heku". The word-only-condition replicated previous findings. Children selected the unfamiliar object as referent for the novel term. No change was observed between the age groups. The other condition presented a very different picture, a clear developmental shift in object choice: 3.5-year-olds picked the novel object and disregarded the pragmatic cue; 5.5-year-olds selected the familiar object as referent for the novel word and did not demonstrate a mutual exclusivity bias.

The sociopragmatic account cannot explain this developmental shift. The speaker clearly vocalises his desire, the intention should be easily read. The puppet is for example “hungry” and there is only one edible item in the set of two. However, children overlooked the pragmatic cue in the second condition until about 4.5 years of age. A control condition ensured that 3.5 year olds were able to understand the pragmatic information in general. Here, children were told the puppet was hungry and asked to pick an object puppet might want. The 3.5 year olds had no problems selecting the familiar object then. The developmental shift in the pragmatic condition was explained with growing experience in word learning. Children were assumed to adhere to the mutual exclusivity bias firstly. With increasing vocabulary and general knowledge they would be able to overcome the bias and consider for example the pragmatic information in this study to determine word meaning. This explanation can be accommodated by the lexical principles approach. It is difficult to see, though, how the acceptance of pragmatic information in Haryu’s study follows such a distinct developmental pattern.

A more likely explanation for the timing of the developmental shift in Haryu’s (1991) study may come from the comparable studies with known words described before. In the alternative naming task, children were shown a picture and asked to provide a different label to a puppet (Doherty & Perner, 1998). The task is therefore similar to the disambiguation task except that all labels were familiar. Children are asked to produce the second familiar label in the one and accept a second unfamiliar label in the other task. Children in the alternative naming task were unable to produce alternatives until the age of about four. This strongly correlated with the ability to pass a false belief task. Doherty and Perner (1998) argued both tasks need children to make a distinction between the object or situation and how it is thought or talked about. Perner, et al. (2002) characterize this in terms of an understanding of perspective. Understanding that two known words can apply to one object is conceptually similarly to understanding that a known and a novel word can both apply to one object. The latter understanding is required to pass Haryu's pragmatic condition.

The child was presented with a familiar item (apple) and an unfamiliar item (lipstick holder). Asked for a "hinkle", the lipstick holder is the sensible choice. Asked to pick an item for the hungry puppet, the apple is the sensible choice. The combination of "hinkle" and "hungry" presents the child with a dilemma. The apple has already a label, the lipstick holder has not. But taking pragmatic information supplied by the speaker into consideration requires the selection of an edible item. This means, the new label has to be attached to the apple. The child needs to be able to accept two labels for one object.

The sociopragmatic account should predict all children to be able to select the apple as referent for the hinkle. Speaker intention is obvious here. The lexical principles account would accept children can override mutual exclusivity in the face of clear pragmatic information but makes no predictions about the developmental pattern Haryu found in her study. The perspectival account predicts young children should be unable to consider the pragmatic cue. Children should be able to take the additional information into account at the same time as they become able to pass false belief and alternative naming. They should be able to pass a new pragmatic cue task as they reach a level of metacognitive maturity, indicted by theory of mind. – These predictions were addressed in the following studies.

Experiment 1

Method

Participants.

Participants were 88 children (44 girls) from predominantly middle-class nurseries in central Scotland: 43 children took the standard disambiguation task (mean age 43 months, range = 31 - 59 m, *SD* = 7 m); 45 children took the pragmatic cue task (mean age 45 months, range = 34 - 60 m, *SD* = 7 m). Children's verbal understanding was measured by the British Picture Vocabulary Scale II (Dunn, Dunn, Whetton, & Burley, 1997), a standard measure of verbal mental age, and did not differ between groups (disambiguation group: mean age 41 months, range = 29 – 59, *SD* = 7; pragmatic cue group: mean age 44 months, range = 20 – 84, *SD* = 12; t = -1.25, df = 86, p = .215, d = -.27).

Design.

The false belief task, the alternative naming task, and either the disambiguation task or the pragmatic cue task were administered over two sessions in randomized order. The BPVS II was administered last.

Procedure and Materials.

Disambiguation task and pragmatic cue task.

The child was introduced to Jimmy the puppet, then presented with a familiar object (e.g., a banana) and an unfamiliar object (e.g., a bottle stopper). Children were asked to choose an object, using a novel word.

Disambiguation condition:

"Jimmy would like a hinkle, please give Jimmy a hinkle."

Pragmatic cue condition:

"Jimmy is hungry and would like a hinkle, please give Jimmy a hinkle."

The objects, novel words, and pragmatic cues for all five trials are listed in Appendix A.

False belief task.

The child was introduced to two Playmobile® figures, a box, a jar and a marble and the following story was acted out:

"Now look, this is Sally and this is Tom. They have a box and a jar. Sally has a green marble. Sally puts her marble in the box and then she goes away. Now, Tom picks up Sally's marble from the box and puts it in the jar. Then Tom goes away. Look, Sally is coming back."

Each child was then asked the following three questions in order:

Belief question: *Where will Sally look first for her marble?*

Reality question: *Where is the marble really?*

Memory question: *Where did Sally put the marble in the beginning?*

Children had to answer all three questions correctly to pass the task.

Alternative naming task.

Vocabulary check.

Four sheets of paper were presented individually, each displaying six pictures. Children had to point to each experimental item twice on different sheets, once under the basic label (e.g., "Show me the cat") and once under the superordinate label ("Show me the animal"). The correct item was pointed out to the child if the child refused to make a choice or pointed to the wrong picture.

Alternative naming phase.

Children were presented with an individual picture and told:

"Now, here are some more pictures. Each picture has two names. I am going to tell you one name for it and you can then tell me another name for it. Let's try that. This is fruit. What else is it?" If the child did not respond, encouragement was given. "We can also call it an apple."

After this practice trial, the procedure continued with four pictures (cat, food, owl, drink), then a second time using the alternative label (animal, burger, bird, milk). Children had to provide both superordinate and basic label to pass a particular trial.

Results 1

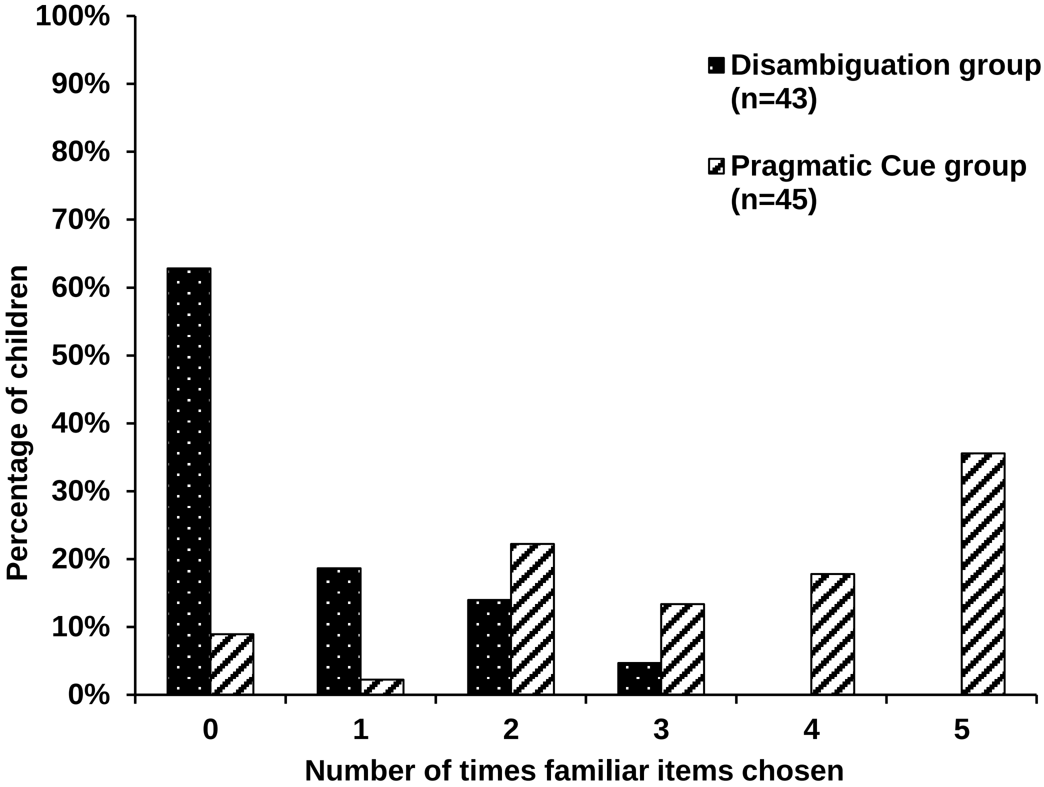


Figure 1. Performance on disambiguation and pragmatic cue task

Disambiguation and pragmatic cue task.

Figure 1 shows the number of times children chose the familiar item. For the disambiguation task, most children chose the unfamiliar object on every trial; the most frequent response for the pragmatic cue task was to choose the *familiar* object on every trial, a highly significant difference: t = 9.87, df = 86, p < .001, d = 2.13. To compare performance, children were counted as passing each task if they chose the novel object (disambiguation task) or familiar object (pragmatic cue task) on 4 or more of 5 trials. Using these criteria, 81% of children passed the disambiguation task but only 53% passed the pragmatic cue task.

False belief task.

About half the children passed the false belief task in each group: disambiguation group 49%, pragmatic cue group 53%, t = 0.79, df = 86, p = .430, d = 0.17.

Alternative naming task.

Mean performance on the vocabulary check was 7.4 out of 8 items (SD = 0.70), indicating that failures on the ANT were not due to lack of relevant vocabulary. Setting a passing criterion of 2 out of 4 pairs, 40% of children in the disambiguation group and 47% of children in the pragmatic cue group passed the alternative naming task, a non-significant difference (t = 1.33, df = 86, p = .189, d = 0.29).

Most failures on the ANT resulted from inability to provide superordinate labels. The 38 children who passed the task produced 3.45 basic labels and 2.89 superordinate labels. The 50 children who failed the task produced a mean 3.04 out of 4 basic labels and 0.2 superordinate labels; 81% of superordinate labels children failed to provide were in their receptive vocabulary.

Twenty-one percent of responses were unanticipated but valid, such as “sandwich” instead of “burger” and “pussycat” instead of “cat”. The analysis presented below leniently scored such responses as correct. A separate analysis conducted using strict criteria produced the same overall pattern of results.

Comparison of tasks.

Associations between task performances were examined separately for the disambiguation and pragmatic cue groups. Performance on the disambiguation task approached ceiling and was not significantly associated with other variables. Age and verbal mental age correlate strongly with performance on the false belief and the alternative naming task for this group (Table 1). No correlation was found between the disambiguation task and age, verbal mental age or FB and AN. The correlation between FB and AN remained stable after the influence of age and verbal mental age was accounted for.

Table 1 Correlations between tasks for disambiguation group (partial correlation after partialling out age and VMA)

	VMA	False belief	Alternative naming	Disambiguation
Age	.53**	.62**	.43*	.26
VMA		.56**	.40*	-.05
False belief			.63** (.47*)	.09 (-.02)
Alternative naming				-.03 (-.11)

* p <. 01,   ** p < .001

Children in the disambiguation group who passed the false belief task were also likely to pass both of the other tasks. Children who failed false belief were also likely to fail AN, but not DT. A binomial test demonstrated that children found the disambiguation task significantly easier than the false belief task ($p = .004$) and the alternative naming task ($p < .001$). The false belief task and the alternative naming task were not significantly different ($p=.344$). Hence there was no indication of a connection between performance on false belief and the disambiguation task.

The disambiguation and pragmatic cue groups performed virtually indistinguishable on all tasks apart from DT and PC, so the pragmatic cue group analysis is presented in greater detail.

Table 2 Correlations between tasks for pragmatic cue group (partial correlation after partialling out age and VMA)

	VMA	**False belief**	**Alternative naming**	**Pragmatic cue**
Age	.75**	.46**	.55**	.41*
VMA		.54**	.60**	.52**
False belief			.68** (.52**)	.65** (.51**)
Alternative naming				.68** (.53**)

*p < .01, **p < .001

Performances on the false belief, alternative naming and pragmatic cue tasks were significantly intercorrelated (Table 2) and remained substantial and significant after age and VMA were partialled out (in parentheses) .

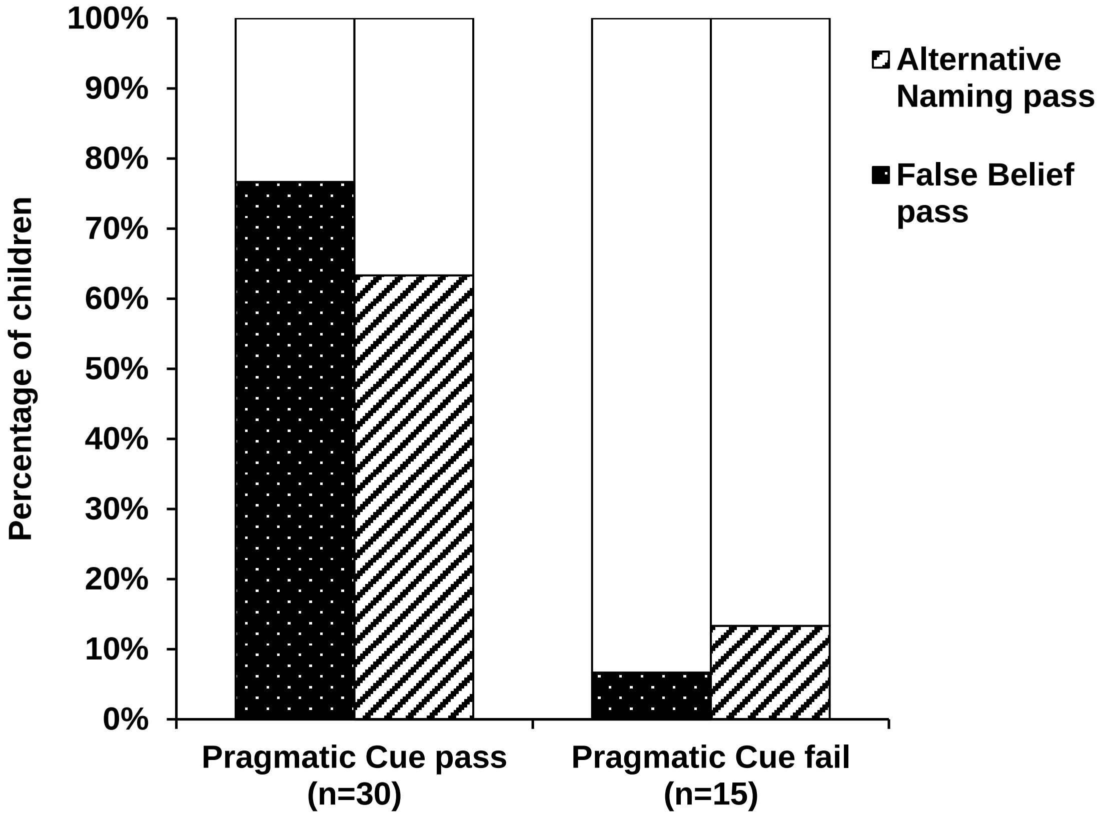


Figure 2 Percentage of children passing the false belief and alternative naming tasks according to performance on the pragmatic cue task

Figure 2 shows the percentage of children who passed the false belief and alternative naming tasks according to whether they passed or failed the pragmatic cue task. Most children who passed the pragmatic cue task passed the false belief and the alternative naming task; few children who failed the pragmatic cue task passed either of the other tasks. There were no significant performance differences between individual tasks (binomial test: pragmatic cue versus false belief task, $p = 1.00$, pragmatic cue versus alternative naming task, $p = .453$, false belief versus alternative naming, $p = .549$).

Discussion 1

The clear difference in performance between the disambiguation task and the pragmatic cue task lead to the following interpretation. The disambiguation group did not receive a pragmatic cue and performed as if following a mutual exclusivity bias. Children's performance on the disambiguation task did not change with age or metarepresentational awareness. No association between disambiguation and alternative naming was apparent. Children made a logical choice by selecting the novel object as referent for the novel label. With no other information provided, this seems reasonable. The pragmatic cue group on the other hand showed a pattern of results depending on the children's ability to pass false belief. Children who were able to use the pragmatic cue to direct object choice were also able to predict behaviour in the false belief task and produce basic and superordinate labels in the alternative naming task. This difference in combination with the observation that all three tasks (PC, FB, AN) correlate well for the second group (Table 2), but only two of the tasks correlate for group 1 (FB, AN; Table 1) clearly indicates a connection between children's ability to use pragmatic information and their metacognitive abilities.

The pragmatic cue task seems therefore to measure metalinguistic abilities in a similar way to the alternative naming task. Where the latter requires children to flexibly apply two familiar labels to one object, the pragmatic cue task requires children to accept a second yet unknown label for a familiar object. Both tasks involve the understanding that more than one object label can be applied in one conversation with one conversational partner. The pragmatic cue task shows a clear change in children's disambiguation of novel words in the presence of additional information, contrary to the disambiguation task, which presents no change in behaviour. Disambiguation in the latter task remains constant as young children cannot conceive of an object having multiple labels in the given task. Older children make the same object choice because there is no additional pragmatic information to behave differently. This points to the mutual exclusivity bias being a developmental phenomenon. Younger children seem to disambiguate slightly less in the pragmatic cue task, though, which was investigated with a later study.

One concern was that in the pragmatic cue task the cue (e.g. “Jimmy is hungry…”) is always stated before the novel word (“… and would like a hinkle”). Younger or more impulsive children might choose a referent based on the cue, without attending to the following word, producing false positives.

A second experiment was therefore conducted with the order of mention of the cue and novel word reversed for half the participants. Additionally, an inhibition task was added to address the possible influence of inhibitory abilities on object selection.

Experiment 2

Method

Participants.

Twenty one children (9 girls; mean age = 47 months; range = 35 - 57; *SD* = 6 m) heard the pragmatic cue followed by the novel word; 23 children (11 girls; mean age = 47 m; range = 40 - 59; *SD* = 6 m) heard the novel word followed by the pragmatic cue.

Design.

The design was as for Experiment 1, with the addition of the Day-Night Stroop task (Gerstadt, Hong, & Diamond, 1994) included either in the first or second session, counterbalanced.

Procedure and Materials.

Pragmatic cue task.

There were two versions of the instructions:

Cue + novel word:

"Jimmy is very hungry and would really like a hinkle. Every time when he is hungry he likes a hinkle. Please give Jimmy a hinkle."

Novel word + cue:

"Jimmy would really like a hinkle, because he is very hungry. He always likes a hinkle when he is hungry. Please give Jimmy a hinkle."

The procedure was the same as in Experiment 1. Pictures of familiar and unfamiliar objects were used instead of real objects (Appendix B) to avoid distractions caused by children manipulating the objects. This has been shown to produce the typical disambiguation effect (Diesendruck, 2005).

Alternative naming task.

To avoid unanticipated responses, cat/animal and burger/food were replaced by dog/animal and vegetable/carrot. Otherwise the procedure was as before.

Day-Night Stroop.

The child was presented with coloured sun pictures and black and white night sky pictures, and told to respond to the sun card by saying “night” and to the night card by saying “day.” After a brief training phase, children were presented one at a time with eight sun and eight moon cards randomly mixed. The criterion for passing the task was to correctly respond to five out of eight in each set.

Results 2

Children in the cue + novel word condition chose the familiar object on 3.0 out of 5 trials (*SD* = 1.3) compared to 2.7 out of 5 trials for children in the novel word + cue condition (*SD* = 1.2). This difference was not significant (*t* = .69, *df* = 42, *p* = .496, *d* = .22). The two groups performed virtually the same on all other tasks.

Comparison of tasks.

Correlations between the pragmatic cue, false belief and alternative naming tasks were similar to the first experiment, remaining substantial and significant after partialling out age and verbal mental age, and additionally the day-night Stroop (all partial *r* - values > .35, all *p* - values < .05).

Table 3 Correlations between tasks (partial correlation after partialling out age, VMA and DNS)

	BPVS	**False belief**	**Alternative naming**	**Pragmatic cue**	**DNS**
Age	.59***	.32*	.46**	.07	.35*
BPVS Standardised Score		.51***	.46**	.35*	.38**
False belief			.56*** (.43**)	.52*** (.44**)	.16
Alternative naming				.40** (.35*)	.15
Pragmatic cue					.09

* p < .05, ** p < .01, *** p < .001

Discussion 2

The concern about children's ability to inhibit responses was addressed with Experiment 2 in addition to possible order effects of pragmatic cue and novel word. The order of presenting the cue in the pragmatic cue task was not influential on children's decision following the results. Even though this cannot completely rule out the possibility that children might react only on the cue, it ensures that children were listening to the full task instruction. The addition of the day-night Stroop task was able to address inhibitory abilities further. DNS related to age and verbal-mental age, but not to any of the metacognitive tasks presented. Even more, when the ability to inhibit was partialled out of the correlational analysis, the relationship between false belief and the two linguistic tasks stayed strong and significant. This confirms the results presented in Experiment 1. The pragmatic cue task shows strong connections to alternative naming and false belief.

The first two studies demonstrate that children can select a familiar object as referent for a novel label in the presence of an unfamiliar object given information that this is appropriate around the same time as they become able to pass false belief tasks. Until then, children's tendency to choose the novel object as referent of a novel label is remarkably strong. Even when they are told, before hearing the novel label, that Jimmy is hungry, younger children still subsequently choose the inedible novel object when asked to give Jimmy a hinkle.

The ability to take the pragmatic cue into account is specifically related to the development of understanding of false belief, and of alternative naming. The relationship between these two metacognitive abilities has previously been established, and has been attributed to the fact that both require a clear distinction between the object talked or thought about and how it is mentally or verbally represented (Doherty & Perner, 1998). Perner et al. (2002) refer to this distinction in terms of requiring a common understanding of perspective. This distinction also seems critical to success in the pragmatic cue task. Children need to be able to consider the novel label as new label for the familiar target. This requires the mental flexibility to consider multiple labels for one object, similar to the alternative naming task. The only difference is that the alternative naming task involves two familiar labels for one referent and the pragmatic cue task introduces a second label for one referent. Children who are able to consider the pragmatic cue for object choice present the metalinguistic ability to also produce alternative names for an object. Alternative labels for an object differ in perspective (Clark, 1987; Tomasello, 1999). Understanding perspective is therefore the key to passing the pragmatic cue task.

Findings presenting some alternative naming abilities in the literature possibly challenge the presented alternative naming findings. Waxman and Hatch (1992) presented children with a misnaming task, asking them to teach “teddy” multiple names for a picture (e.g. rose, flower, plant). In order to elicit these responses, “teddy” misnamed the picture on each level (following the previous example: dandelion, tree, animal). They found 75% of the 3 year olds in their study to produce more than one label for a picture on at least 50% of the trials. This could contradict the assumptions of the perspectival account, considering the 3 year olds are most likely failing false belief. If these children are able to give a response and then select a different target label on a different level as next response while still keeping the former label in mind, they are demonstrating perspectival understanding. This understanding has previously not been detected in children before they pass false belief. Experiment 3 a) was set up to address possible connections between the misnaming task, alternative naming and false belief.

A further finding from the literature which was addressed in Experiment 3 is the correction effect, which describes the removal of a familiar label in favour of a novel one. Next to the disambiguation effect, the correction effect is the only other proposed word-learning effect for which there is convincing evidence (Frank & Poulin-Dubois, 2002; Savage & Au, 1996) and should therefore be investigated in relation to the perspectival account.

Savage and Au (1996) reported a study teaching children two new labels for one object (e.g. lemur/primate). One was introduced by the teacher, the other by the experimenter. Children were later asked to identify first all lemurs in a set of toys. After this sorting, they were asked to identify all primates (or all lemurs). Less than half of the 3 to 5 year olds accepted two labels for the object. The majority corrected one of the previously taught labels and persevered with the other one.

Frank and Poulin-Dubois (2002) presented children with two experimenters referring to a novel object by two different novel labels. The authors investigated whether children would accept the overlap or choose an unnamed alternative when asked for the second label. Older children (mean = 35 m) in their study displayed a stronger disambiguation effect. If the pragmatic cue task identifies children's ability to take pragmatic information into account to connect novel words with familiar objects, an additional pragmatic cue in Frank and Poulin-Dubois task should do the same and direct children to a specific target from the same age as they pass false belief. The ability to consider two novel labels for an object would therefore demonstrate the connection to perspectival understanding in another word learning situation. This hypothesis was addressed in Experiment 3a.

Experiment 3a

Method

Participants.

This study involved 38 children (18 girls) from local Nurseries in Salzburg (Austria) and was fully conducted in German. Children were aged between 36 and 79 months (mean = 53 months, SD = 11). Parents signed their children up through a consent form, confirming no history of language or hearing impairment and age-appropriate understanding of the German language.

Design.

Each child was seen individually in a small room next to the playgroup. The tasks were randomly split over the course of two days and included two false belief tasks, one on each day, the alternative naming task, pragmatic cue task, disambiguation task, a cued version of a task around the correction effect (extending a procedure used by Frank & Poulin-Dubois, 2002) and a misnaming task (after Waxman & Hatch, 1992).

Procedure and Materials.

The first false belief task was identical to Experiment 1 and was fully translated into German by a native speaker and reviewed by a second native speaker.

Alternative naming task.

The carrot-vegetable pair was used as teaching item and a new item (water-drink) was added to the set-up to avoid two food pairs in the test set. The procedure was as in Experiment 2.

Pragmatic cue task & disambiguation task.

Both tasks were identical to Experiment 1 and 2 with one alteration to further improve the child's focus on the novel word:

The two pictures were presented to the child face down first. The child was then told that the puppet wanted (e.g.) a "Kulde", and in the pragmatic cue task additionally informed that the puppet was very hungry (thirsty/cold/sleepy/sore). The child was then asked if she knew what a "Kulde" was and told there was a "Kulde" on one of the pictures. The pictures were then turned over and the child was asked to select the "Kulde". The alteration was introduced to prevent children from reacting to the cue only and disregarding the novel label (materials: Appendix K & L). The tasks were also followed by a recall task immediately after.

Recall.

The child was presented with all five pictures she had selected in response during the pragmatic cue task and asked: "Which one is Jimmy's hinkle?", testing every novel word used before. Children were encouraged to point to an object for each of the novel words.

False belief 2.

The child was presented with teddy “Sepperl” and told the following story: “This is Sepperl, he brought his lunchbox in today. Oh, look, he puts his keys in the lunch box to keep them safe. He wants to have a nap just now.” Teddy was then put in the bag, out of sight. The child was told teddy would be asleep by now and asked “Do you think Sepperl is having a nice nap?” in order to ensure the child assumed teddy would not listen to the following. “Now, I’m going to play a trick on Sepperl, I’m going to take his keys and hide them under my sheet.” The experimenter took the keys out of the box and placed them under the score sheet in front of her. She then looked into the bag. “Oh, I think Sepperl might be getting up now!” Teddy appeared again on the table and was placed between the box and the score sheet. The child was asked the following questions in order:

Belief question: *Where will* Sepperl *look first for his keys?*

Reality question: *Where are the keys really?*

Memory question: *Where did* Sepperl *put the keys in the beginning?*

Children had to answer all three questions correctly to pass the task.

Correction task.

The experimenter introduced the puppet “Freddy”, who was able to move his mouth and speak to the child. The puppet was therefore filling the purpose of a second experimenter (experimenter 2 in the following). The task consisted of a training and a test phase. Each training phase introduced four objects to the child, two familiar items such as small toys and two unfamiliar items. A full list of all materials can be found in the appendix C. The task entailed four trials in set order.

Training phase.

Experimenter 1 handed an object to the child, stating “Look at this!”, and allowed the child to explore the object for up to 20s. When the child handed the object back, it was given to experimenter 2 to look at. Then the next object was introduced. The objects were ordered to introduce two familiar objects first, then two unfamiliar. If the child asked about object 3, which was unfamiliar, the experimenter explained she didn’t know what this one was, correcting the child if she tried to name it: “No, this is not an airplane. I don’t know what it’s called.” When the fourth object was presented to the child, the experimenter said: ”Oh, look, this is a Puhne, we have a Puhne here, look at the Puhne.” , labeling the novel object with a novel name three times. After the child explored the “Puhne”, it was handed to experimenter 2, who said: ”Ah, a Grieber, this is a Grieber, I like this Grieber.”, labeling the same object with a second novel name. The experimenter also demonstrated a function of the novel object, for example pretending to try and eat it to cue the object was edible. The object was then handed back to the child to briefly explore again. The function was introduced at the end of each set to have no interference between the two labels children heard for one object.

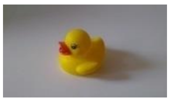
Test phase.

All four objects were then placed on a tray in front of the child. The unfamiliar and double-labeled objects were placed on a different position on the tray for each of the trials. The remaining objects were positioned randomly on the tray. One experimenter asked the child to identify one of the familiar items:

“Show me the car.”, and then to identify the labeled unfamiliar object (target): “Show me the Puhne.” The other experimenter then asked for the second familiar item. He further asked for the target using the label introduced by him, reasoning with the function presented before: “Show me the Grieber, I am hungry.”. The order of experimenter to start the questioning was counterbalanced, the function was always demonstrated at the final demand for an object (familiar object – target object – familiar object – target + function). Each experimenter used the novel word he/she had introduced before. The function demonstrated was always presented by the experimenter who had introduced it.

Example of a set of items:

A car, a toy duck, an item made up of several curlers, a piece of bath salt in a candy-like form.

			 1. Puhne 2. Grieber + function

Misnaming task.

The task used four sheets with nine different pictures of familiar items for the vocabulary check. Each sheet included two target items. The target items were later presented in the test phase on individual sheets of paper. A full list of all items can be found in appendix D. Target items were similar to the ones used in Waxman and Hatch (1992) if the names could be translated into

German in a reasonable way. Each target item had to be easily identified by preschool children on three different levels: subordinate, basic and superordinate labels, for example: parrot, bird, animal.

Vocabulary check.

Each child went through a vocabulary check similarly to the one the alternative naming task introduced. This ensured the child's ability to identify each target item by three different labels. During the first set, the child was for example asked to point to a ladybird. The child was then asked to point to a beetle at the second show of the same sheet, and to an insect in third instance. If the child could not identify the item, the correct one was pointed out.

Training phase.

Puppet "Lisa" was introduced to the child. The experimenter explained that Lisa wanted to learn a lot of new words and the child was going to help teaching her. A picture of a rose was then placed in front of the child with the instruction: "Look at this. We want Lisa to learn many words for this picture. We want her to know, this can be called a rose. It can also be called a flower. And we can call it a plant. See, three names for one picture. Lisa should learn it's called a rose, a flower and a plant."

Test phase.

A set of eight pictures was then presented to the child, one at a time. The puppet looked at the picture and asked a set of three contrasting questions, aiming to draw three different responses from the child. The questions consisted of a contrasting sub-, basic, and superordinate label. Questions were counterbalanced so a child would either receive all in ascending or descending order, starting with the sub- or superordinate label first. For a picture of a police car, the child was asked for example: “Is this a fire engine? - Is this a plane? – Is this furniture?” and the target responses would be “police car – car – vehicle”. If the child responded by shaking her head or with “no”, encouragement was given: “What is it then?”

Results 3a

Task performance.

Children’s performance on the false belief tasks, the alternative naming task, the pragmatic cue task and the disambiguation task was scored by pass or fail. Criteria were the same as in Experiment 1.

Table 4 Task performance

	Fail	Pass	Transition
False belief 1	18	20	
False belief 2	17	21	
False belief combined	14	17	7
Alternative naming	15	23	
Pragmatic cue	20	18	
Disambiguation task	7	31	

Comparing individual tasks.

False belief.

Children had to answer all three questions correctly in order to pass each false belief task. False belief 1 was passed by 53% of the children, false belief 2 was passed by 55% of the children. One of the failing children failed the reality question in addition to the belief question in task 1 and a different child in task 2. All children answered the memory question correctly. Children who passed one of the false belief tasks but not the other were assigned to the

“Transition” group (18% of the children). Data from false belief 1 and false belief 2 was added up to false belief combined.

Alternative naming.

Children were asked to identify target items by pointing in the vocabulary check. An average of .55/8 mistakes were made (SD = .65, range: 0 – 2). The majority of children made no mistakes (52.6%, 20), 39.5 % (15) made 1 mistake, 7.9 % (3) made 2 mistakes.

The alternative naming task was passed by correctly producing basic and superordinate label for two of the pictures (e.g. fruit and apple). The 39 % children who failed the task produced on average 2.13 basic level labels (SD = 1.55) and .60 superordinate labels (SD = .63). The remaining children produced on average 3.78 basic labels (SD = .42) and 2.65 superordinate labels (SD = .71).

Pragmatic cue and disambiguation task.

Children chose the familiar object as referent for the novel word on average 3.11 (SD = 1.82) in the pragmatic cue task, compared to .63 times (SD = 1.32) in the disambiguation task, a highly significant difference: $t = 8.152$, $df = 37$, $p < .001$, $d = 2.68$.

Both tasks were followed by a recall phase. Children correctly identified on average 3.03 (SD = 2.26) of pictures after the pragmatic cue task and 2.97 (SD = 2.39) in the disambiguation task, a non-significant difference ($p = .861$), both significantly different from chance (PC: $t = 7.71$, $df = 37$, $p < .001$; DT: $t = 7.16$, $df = 37$, $p < .001$).

Correction task.

The child was presented with an array of two familiar and two unfamiliar objects. One of the unfamiliar objects was labelled differently by the two experimenters, the other one remained unnamed. Children were asked to identify both familiar objects by pointing ("Show me the car"). All children apart from 1 pointed to the correct familiar items on all occasions. This one child pointed to a toy duck when asked for bread, but then correctly identified the duck when asked for the duck.

Children were scored as avoiding overlap if they picked the unnamed novel item as referent for one of the two newly introduced labels and the target object as referent for the other novel label. No child refused to make a choice. Children avoided overlap on average on 2.87/4 cases (SD = 1.28). Overall, 45% (17/38) of the children avoided overlap on all sets and only 5% (2/38) applied both labels to the target object on all sets. Performance did not differ in relation to whether the first or second experimenter introduced the cued label ($t = -.854$, $df = 36$, $p = .399$) or whether the cue was given pre- or post- novel label ($t = .630$, $df = 36$, $p = .533$).

Following the original task by Frank and Poulin-Dubois (2002), children were divided by age into two approximately equal groups. A main effect of age was observed ($F(1,30) = 10.084$, $p = .003$, partial $\eta^2 = .252$), with younger children avoiding overlap significantly more often (mean = 3.56, SD = .62 vs mean = 2.25, SD = 1.41, Figure 3).

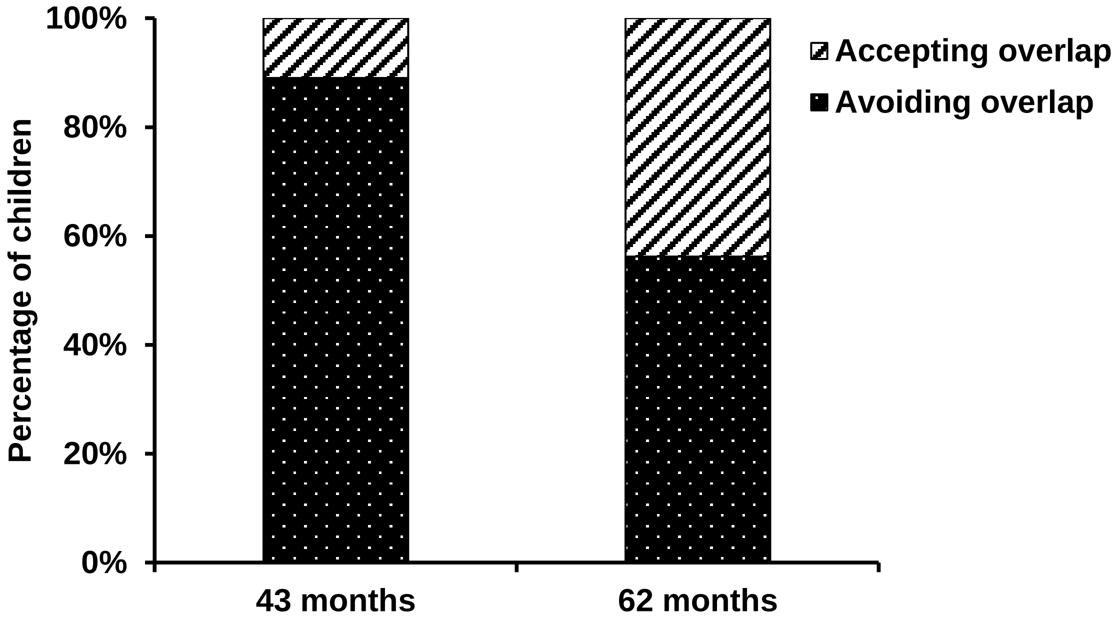


Figure 3 Percentage of children accepting and avoiding label overlap

Age split: N = 18, mean age = 43 m, 36 – 50 m, SD = 4 m, 4 yr olds

N = 20, mean age = 62 m, 53 – 79 m, SD = 7 m, 5 yr olds

Comparing tasks overall.

Performance on the combined false belief task, the alternative naming task, the pragmatic cue and disambiguation task and the correction task were entered in a correlational analysis.

Table 5 Correlation of tasks

	FB combi	AN	PC	PC recall	DT	DT recall	Correction task
Age	.57***	.59***	.68***	.10	-.32	.11	.59***
False belief combined		.61***	.63***	.24	-.16	.09	.46**
Alternative naming			.66***	.23	-.22	.12	.52**
Pragmatic cue				.39*	-.33*	.43**	.53**
PC recall					-.26	.69***	.26
Disambiguation task						-.07	.25
DT recall							.33*

* $p < .05$, ** $p < .01$, *** $p < .001$

Children's ability to produce familiar alternative labels and accept an unfamiliar label for an already named object correlates strongly with false belief. As for the correction task, children's tendency to select the newly labeled target item twice and accept two labels correlates with age, false belief, alternative naming and pragmatic cue. This correlation disappears once age is partialled out.

Table 6 Partial correlations, partialling out age

	Alternative naming	Pragmatic cue	Disambiguation task	Correction task
False belief combi	.42*	.40*	.03	-.18
Alternative naming		.43**	-.04	-.25
Pragmatic cue			-.16	-.21
Disambiguation task				.08

* p < .05,   ** p < .01

Children failing false belief did not disambiguate to the same level in the pragmatic cue as in the disambiguation task (Table 7).

Table 7 Novel object chosen by false belief performance for PC and DT task

	FB fail (n = 14)	FB transition (n = 7)	FB pass (n = 17)
PC novel object, mean (SD)	3.36*/**(1.55)	1.57 ***(1.27)	.82**** (1.43)
DT novel object, mean (SD)	4.79 (.58)	3.71 (1.89)	4.29 (1.45)

* significantly different to FB transition (p = .035)
and FB pass (p < .001; post-hoc Bonferroni); (F (2,35) = 11.964, p < .001)

** significantly difference to DT FB fail (t = -3.98, df = 13, p = .002)

*** significantly different to DT FB transition (t = -2.68, df = 6, p = .037)

**** significantly different to DT FB pass (t = -8.42, df = 16, p < .001)

The false belief failers were split by age and their selection of the novel object was compared for both tasks. Unfortunately, the number of young children in this experiment was extremely limited and the age range of FB failers was large. Table 8 illustrates, that the youngest FB failers disambiguate to similar levels and start to change decision with age increase. The sample size in each age range is very small, though.

Table 8 PC and DT for FB failers: 14 children

Mean age = 44 m, SD = 7 m	**PC novel object** **Mean = 3.36, SD = 1.55**	**DT novel object** **Mean = 4.79, SD = .58**
Age 36-38, (n = 4)	Mean = 4.25	Mean = 5.00
Age 40-44, (n = 5)	Mean = 3.40	Mean = 4.80
Age 49-56, (n = 5)	Mean = 2.60	Mean = 4.60

Misnaming task

Children misidentified an average of 1.76 (SD = 1.50) target words during the initial vocabulary check. Most words were superordinate labels (1.63, SD = 1.34) Only .03 basic labels were misidentified (SD = .16) and .11 subordinate labels (SD = .31). Older children identified significantly more target words than younger children (t = 3.822, df = 36, p = .001, d = 1.274; 4 yr olds: mean = 2.61, SD = 1.50; 5 yr olds: mean = 1.0, SD = 1.03).

Each target could elicit up to three different labels; children persevering with one label received a score of 1, 2 different labels were scored as 2, and 3 different labels as 3. The younger group produced an average of 1.14 (SD = .23) labels per target, compared to an average of 1.31 (SD = .42) in the older group, a non-significant difference (t = -1.536, df = 36, p = .135, d = -0.512). There was no effect of order (ascending mean = 1.18, SD = .25; descending mean = 1.28, SD = .43; t = -.869, df = 36, p = .391).

Children produced significantly more subordinate answers (68% of all answers), than basic (23%) or superordinate (3%) answers, $F(2,72) = 159.049$, $p < .001$, partial $\eta^2 = .815$ (all pairwise comparisons $p < .001$, Figure 4).

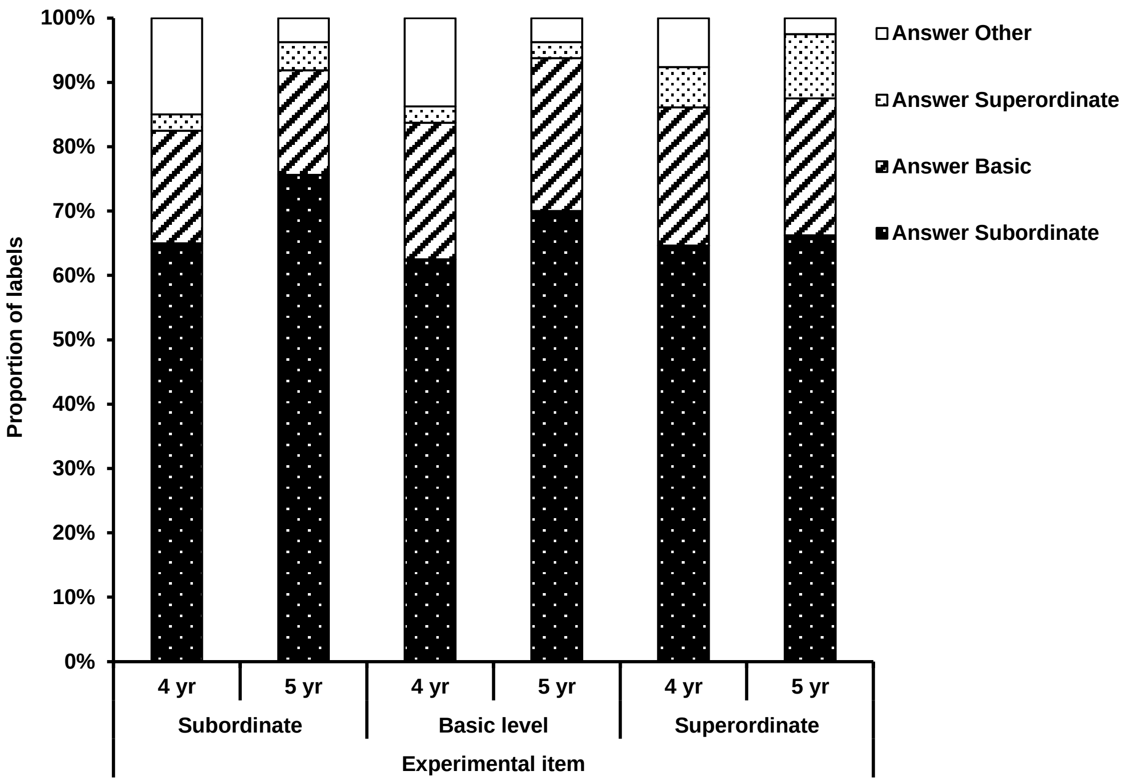


Figure 4 Proportion of labels production per target

Comparing the response pattern for the different age groups, 56% of the four year olds produced no alternative label for any target and persevered with the first answer given. No child in this age group was able to name any alternative for any target. In the older age group, 40% of children produced no alternative label, two children were able to produce one alternative for each target.

Discussion 3a

Similarly to Experiment 1 and 2, children showed distinctly different behaviour in the disambiguation and pragmatic cue task. While the novel object was picked across age ranges in the majority of cases in the disambiguation task, the pragmatic cue task displayed a different picture of responses. Children passing the pragmatic cue task also passed the other metacognitive tasks. The false belief task, alternative naming task and the pragmatic cue task correlated strongly. The relation remained stable after age was taken into account. The connection between the tasks was therefore again demonstrated and with children from a different native language background. Word learning was also presented as children performed a recall task after pragmatic cue and disambiguation. Word learning did not show a relation to metacognitive abilities. As children learn a large amount of language before they understand the more abstract use, this is not surprising.

In the correction task, younger children displayed a strong tendency to avoid two labels overlapping. The older age group (mean age = 35 months) in Frank and Poulin-Dubois' (2002) study avoided overlap on average 2.26/4 times, significantly more often than the younger group (mean age = 27 months, avoid overlap 1.24/4). In Experiment 3a, the younger group was closer in age (mean age = 43 months) to Frank and Poulin-Dubois' older group and avoided overlap on 3.56/4. This would match Frank and Poulin-Dubois' explanation that children display stronger mutual exclusivity when getting older. But the older children in the new study (mean age = 62 months) avoided overlap significantly less, only on 2.25/4 cases. The trend was therefore reversed. These differences can be explained by the added pragmatic cue in the new task.

Young children in the new task cannot consider the additional pragmatic information and need to select two different objects for two different novel labels in the test situation. These children performed similarly to the children in Frank and Poulin-Dubois study, although additional pragmatic information was given. The older children in the new study started to use the pragmatic information and select the target twice, accepting overlap. The ability to accept both labels was displayed in relation to false belief understanding, as the perspectival account would predict.

Waxman and Hatch's (1992) findings could not be repeated. The younger children in this study mostly persevered with their first answer and did not produce more labels in line with the different categorical questions. Even older children did not produce similar results as in the original study. The majority of children persevered with the first label they produced. One concern was the repeated questioning children were exposed to. This might have seemed odd and unnatural and encouraged Children to stick with the first label they had given. This was addressed in the following study.

Experiment 3b

Method

Participants.

Of the 38 children in Experiment 3a, 24 (11 girls) were available to perform an additional task on a separate day (mean age = 55 m , SD = 9 m, range: 40 – 70m).

Design.

The additional task was performed on a different day to the original misnaming task, with approximately two weeks between tasks.

Procedure and materials.

The procedure for the vocabulary check was identical, using four new target items only (Appendix E). The test phase differed from the original. First, a new puppet was introduced. Then, the pictures were presented one at a time and the puppet asked only one question before moving on to the next picture, starting either at the sub- or superordinate mislabelling level. Arriving at the first picture shown again, the puppet continued to mislabel, now at the basic level. Finally, the remaining set of questions was asked going through the pictures a third time. This manipulation was introduced to see if changes between pictures would elicit a different answer from the child and reduce the possibility of perseverance. This would give clearer evidence of whether children would respond on the hierarchical level they were exposed to. The repeated questioning introduced in Experiment 3a might also have seemed strange to some children.

Results 3b

Children made on average .46/12 mistakes during the vocabulary check (SD = .59). The younger group made .50 (SD = .67) mistakes compared to .42 (SD = .52) in the older group, a non-significant difference ($t = .340$, $df = 22$, $p = .737$, $d = .144$). Mistakes were made mostly by misidentifying superordinate labels (mean = .42, SD = .50). No basic labels were misidentified and only a mean of .04 (SD = .20) subordinate labels.

Age split: N = 12, mean age = 47 m, 40 – 53 m, SD = 4 m, 4 yr olds

N = 12, mean age = 62 m, 55 – 70 m, SD = 5 m, 5 yr olds

The four year olds produced on average 1.38 labels per target (SD = .27), compared to 1.46 (SD = .53) produced by five year olds ($t = -.484$, $df = 22$, $p = .635$, $d = -0.206$). Children produced significantly more subordinate (45% of all answers) and basic answers (45%), than superordinate (4%) answers, $F(2,44) = 49.383$, $p < .001$, partial $\eta^2 = .692$ (Figure 5; pairwise comparisons: basic vs superordinate, $p < .001$; sub vs superordinate, $p < .001$; subordinate vs. basic, $p = .367$).

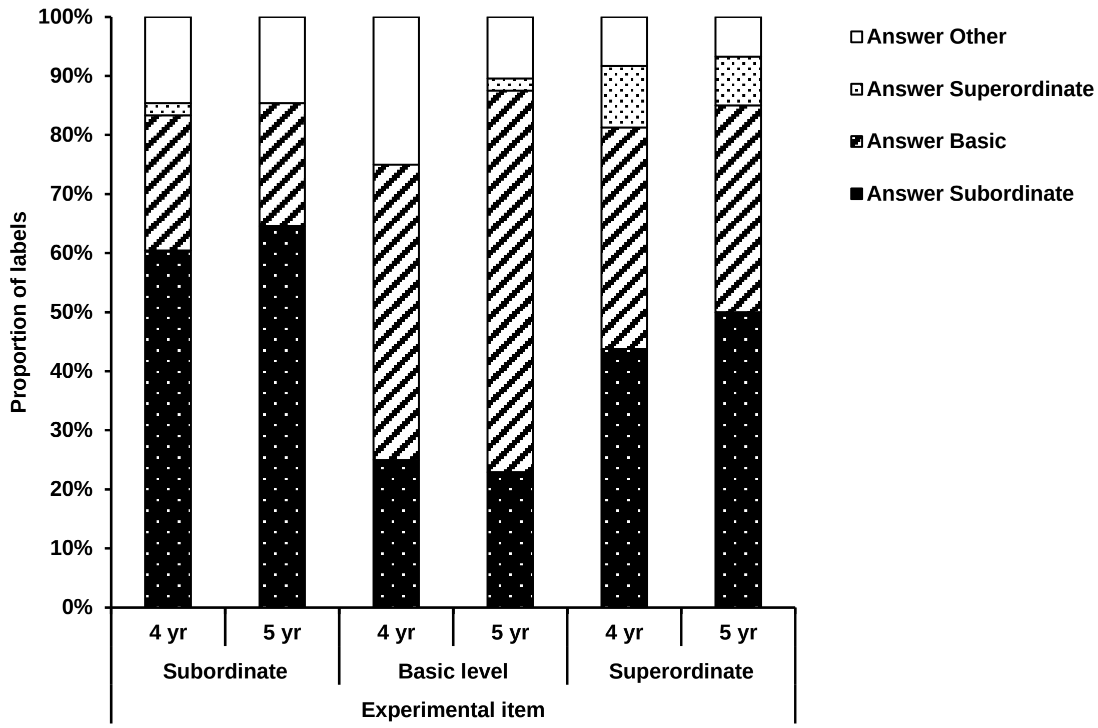


Figure 5 Proportion of labels produced per target

Comparison Experiment 3a and b.

It was of interest whether children would be able to give an answer on a matching level and find it easier to resist perseverance than in Experiment 3a. Similarly to Experiment 3a, each target could elicit up to three different labels. Children persevering with one label received a score of 1, 2 different labels were scored as 2, and 3 different labels as 3. The proportion of mean perseverance was calculated for each child and task and compared for both tasks. The 24 children who undertook both studies produced an average of 1.31 (SD = .38) in Experiment 3a and an average of 1.52 (SD = .65) in Experiment 3b. The difference was not significant (t = 1.645, df = 23, p = .114). Children were therefore not producing significantly more labels in the new condition.

Children gave a subordinate answer in response to a subordinate misnaming on average 2.5/4 (SD = 1.02) times, a basic in response to a basic misnaming on average 2.42/4 (SD = .88) times and a matching superordinate answer to a superordinate misnaming .42/4 (SD = .83) times. Pairwise comparison revealed significant differences between superordinate and basic ($t = 8.880$, $df = 23$, $p < .001$) and superordinate and subordinate answers ($t = -8.961$, $df = 23$, $p < .001$). Comparing this to the matching responses children gave in Experiment 3a, a significant difference was found on the basic level ($t = 5.816$, $df = 23$, $p < .001$), with children in Experiment 3b being more likely to respond on a basic level when presented with a basic term (Table 9).

Table 9 Proportion of matching responses per child and task

	3 a)	**3 b)**
Subordinate misnaming, subordinate response, mean (SD)	.69 (.22)	.63 (.26)
Basic misnaming, basic response, mean (SD)	.26 (.15)	.57 (.21)
Superordinate misnaming, superordinate response, mean (SD)	.10 (.16)	.09 (.19)

Comparison alternative naming and misnaming tasks.

The alternative naming task described in Experiment 3a required children to produce an alternative label for a picture in response to the experimenter labeling it. Both misnaming tasks also required children to produce a label. The difference was that the experimenter **mis**labeled the original picture. Task demand might be comparable, though, as children are asked to switch to another label in the misnaming task. A correlational analysis was performed for alternative naming and both misnaming tasks. The misnaming tasks included 3 possible labels per item, different to the alternative naming task. Hence, children were scored to produce either none or any alternatives per item, comparing proportions. There was no correlation between children's ability to produce correct alternative names in the alternative naming task and their ability to switch from one correct label to another one (all $r < .358$, all $p > .086$). Both misnaming tasks further did not present correlations to the combined false belief data (all $r < .165$, all $p > .440$), whereas alternative naming correlated strongly ($r = .454$, $p = .026$).

Discussion 3b

As the children in Experiment 3a showed very different results in a repeat of Waxman and Hatch's (1992) procedure, a variation of the task was conducted. This was expected to be much easier for children, as it did not include repeat questioning when faced with one picture. The variation produced slightly different results, but children were not able to produce significantly more labels. An increase of basic level labels was observed. Whereas children in Experiment 3a most often responded with a subordinate label, children in Experiment 3b produced equal amounts of subordinate and basic level labels. Children behaved according to the perspectival account's predictions. In Experiment 3a, they assign a label in response to the experimenter's misnaming. They would be expected to stick with the label as the conversation is still on-going with the same counterpart. In Experiment 3b children could show more flexibility, as every time the picture is presented again, the possibility to produce a different label is given. Indeed, children produce more basic level labels than in Experiment 3a, but not overall significantly more labels.

A further question was to address whether the misnaming task touched on similar abilities as the alternative naming task. If so, a relation to false belief would be expected. But no relationship was found in either version of the misnaming task. The misnaming task neither related to false belief, nor to alternative naming. It cannot be presented as another metalinguistic task.

Discussion 1,2,3

The first 3 studies presented were able to establish the new pragmatic cue task as a novel task measuring metalinguistic awareness. The task proved to be replicable and reliable. Children's ability to consider a pragmatic cue when disambiguating a novel word strongly relates to their ability to predict belief based behaviour. Perner et al's (2002) perspectival understanding of the task is going to be applied in the following in contrast to other claims around the disambiguation effect.

In the literature, there have been two general suggestions. The first is that children assume that object kinds have only one name, the mutual exclusivity bias (Markman & Wachtel, 1988). This bias would limit the number of hypotheses that need to be considered for the referent of a novel word, and is plausibly a useful word-learning heuristic. This theory can accommodate some of the present findings. The developmental change in selection of the familiar object in the pragmatic cue task could reflect a greater reliance on pragmatic information. Supporters of lexical principles accounts acknowledge that children use numerous cues to determine word meaning, and that the ME bias is gradually relaxed over time.

However, the lexical principles account is unable to explain why the change occurs around the age of four years, nor why it is associated with other metacognitive developments. In particular, the first two studies show that children's difficulties accepting a novel name for a familiar object are very similar to their difficulties using coreferential names that have already been learned. The ME bias therefore does not seem to be specific to word learning. The second general explanation of the disambiguation effect is that it results from inferences about speaker intentions: had the speaker intended to refer to the familiar object, she would have used the familiar name (Bloom, 2000; Diesendruck, 2005; Diesendruck & Markson, 2001). Children almost certainly use cues to speakers' intentions to guide their word learning. However, in the pragmatic cue disambiguation task, younger children did not use a clear pragmatic cue to the speaker's intended meaning. This suggests that selection of the novel object in the standard disambiguation task is not due to sophisticated inferences about the speaker's mental states. Instead, the role of theory of mind developments in the disambiguation effect is to allow children to suspend it when not appropriate, as is the case in the PC task.

An alternative explanation of the disambiguation task was suggested by Perner et al. (2002), following suggestions by Flavell (1988) and Markman (1989). Perner et al. hypothesise that the ME bias is related to understanding of perspective, which they also claim is the link between the false belief and alternative naming tasks. The current findings support the general claim and allow more specific conclusions. The data suggest that younger children are unable to choose the familiar object for a novel word. This is not because the word is novel, but is because children cannot conceive of an object having more than one name. This is why they show comparable difficulties with the alternative naming task, and why they become able to pass both tasks at the same age.

It may seem counterintuitive to attribute the disambiguation effect to a metacognitive limitation. After all, choosing a novel object as referent of a novel name is typically the pragmatically most sensible response. Adults also show the effect (e.g. Au & Glusman, 1990). The difference is that adults and older children can respond differently. The effect of this limitation may simply be a fortuitous coincidence. Alternatively, it may be that sensible pragmatic considerations are built into the basic apparatus of naming. For example, Doherty and Perner (2013) suggest that a basic constraint prevents the attachment of more than one defining label to objects for the duration of a conversation. Such a constraint would prevent potential confusion generated by having multiple labels, and therefore multiple perspectives, for the same object. Adults certainly rapidly converge on a common set of referents in a conversation (conceptual pacts; Brennan & Clark, 1996) and children show similar phenomena (Matthews, Lieven, & Tomasello, 2008). Thus the presented studies are consistent with the original idea that the ME bias is a useful constraint, although on language use rather than on word learning as such. They are also consistent with the idea that the bias is pragmatic in nature. However, this is unlikely to result from explicit pragmatic reasoning, but from automatic constraints built into the basic mechanics of discourse.

Findings from the misnaming task (Waxman & Hatch, 1992) could not be repeated, even with simplifying changes.

Findings from the correction task (after Frank & Poulin-Dubois, 2002) were addressed in a modified version, applying the pragmatic cue paradigm. The ability to accept two previously presented labels was strong when suggested by pragmatic information for children passing false belief. Young children did not accept two labels for one object, they corrected one of the original ones.

Introduction II - Monolingual and bilingual children

The previous experiments presented the pragmatic cue task as new addition to the range of metacognitive tasks demonstrating developmental change in behaviour in normal developing children around the age of four, indicated by children's theory of mind. As the pragmatic cue task is concerned with children's word learning, a potentially challenging field should be addressed: bilingual children's word learning. Bilingual children are constantly exposed to multiple labels for referents within and between languages surrounding them. Their metalinguistic awareness might be enhanced and be subject to a different developmental trajectory.

Bilingual's metarepresentational abilities on the other hand might also develop earlier, which would show whether both abilities are expressions of an underlying concept. Hence, if bilingual children's metalinguistic awareness and theory of mind development are not as connected as in monolinguals, the perspectival account will be challenged.

Previous research into bilingual children's word learning and theory of mind development presents very inconsistent results at best. Some studies suggest less or no disambiguation effect, which would suggest young bilingual children do not show the same metalinguistic immaturity caused by the effect. Other studies suggest earlier theory of mind development.

Investigations into early differences around the disambiguation effect often apply the preferential looking paradigm, presented by Halberda (2003). Monolingual 17 month old children were shown to increase looking at a novel object in the presence of a familiar object when listening to a novel name. Younger children did not show the same results and it was concluded monolingual infants start to disambiguate reliably around this age in development.

Byers-Heinlein and Werker (2009) used the Halberda's study design to investigate differences between monolingual, bilingual and trilingual children between the ages of 17-18 months. Monolinguals significantly increased their attention to the novel object when presented with a novel word. Bilinguals did so only marginally. Trilingual children did not increase their attention towards the novel object upon hearing a novel word at all. The authors concluded that disambiguation is a result of knowledge about the familiar object, not a result of absent knowledge about the novel object. Children with more than one language might show less disambiguation, because they don't know a label for a familiar object in all languages available to them yet and might therefore consider the familiar object as referent.

Another study following Halberda's procedure compared a group of monolingual children aged 17 to 22 months to bilingual children, matched on age and vocabulary size (Houston-Price, Caloghiris, & Raviglione, 2010). The authors found a clear disambiguation effect in monolingual children, while bilingual children at the same age showed no evidence of this. This was explained with monolingual's tendency to avoid selecting familiar objects as referents for unfamiliar labels. Bilinguals encounter this situation often and do not immediately look for novel objects when hearing novel words. Hence, this study applied the same set-up as Byers-Heinlein and Werker (2009) with the same age range and returned slightly different results for the bilingual children.

Bilingual preschool children between the age of four and six were first compared to their monolingual peers on the mutual exclusivity bias by Au and

Glusman (1990). They found bilinguals and monolinguals to be equally willing to accept two novel labels from two different languages for one object. Monolingual children disambiguated within their language and avoided overlap of novel terms. Being informed that a novel word would be in a foreign language lead them to assign labels randomly. But being asked whether a previously (English) labelled object could have a second name in Spanish made 89% of monolinguals accept overlap. The authors suggested that explicit task instructions lead to the same presentation of metalinguistic abilities in monolinguals and bilinguals at this age.

In a comparison between monolingual and bilingual three to four, and five to six year olds, no significant differences were found within the younger age groups (Davidson, Jergovic, Imami, & Theodos, 1997). The younger monolinguals and half of the younger bilinguals disambiguated significantly above chance level (69% and 65% of trials), the other half of the younger bilinguals were close to this (60%). The older monolinguals selected the novel target in response to the novel word at ceiling level (92%) and significantly more often than the older bilingual group (69%). Disambiguation was therefore detected in bilingual children in this study from the age of three and not increasing at the same rate as in monolingual children.

Davidson and Tell (2005) found further differences between monolingual and bilingual children also only in an older age group. They used a task introduced by Markman and Wachtel (1988), presenting children with the decision whether a novel word referred to a whole object or a salient part. This was done to show, that children will look for a substance or part of the object as referent for a second label rather than accepting the new label as an alternative name. Davidson and Tell (2005) found three and four year old monolingual to select a part of a familiar object as referent for a novel word in 90% and bilingual in 82% of all cases. The older bilingual children accepted two names for the whole object significantly more often than the monolinguals (35% BL accepting overlap compared to 3% ML). Monolingual's disambiguation increased significantly with age, but not bilinguals, similarly to the first author's study presented before.

Diesendruck (2005) found bilingual children to display a disambiguation effect if they interact with a bilingual puppet and the puppet was present while the bilingual experimenter labelled one of two figures (70% of all cases). If the puppet was absent or monolingual (not understanding the test language), children picked at chance level between two figures at puppet's demand for a newly introduced label (Bl+absent 53%, ML+present: 45%, ML+absent: 52% disambiguation). The author explained that bilingual children use their knowledge about speaker knowledge to interpret the situation. If a speaker has more than one language, bilingual children tend to disambiguate less.

Merriman and Kutlesic (1993) tested monolingual and bilingual children between the age of five to six and seven to eight on the correction effect. They first taught children a novel name for a novel object and asked children to help a puppet pick more referents of this label from an array of objects which had varying similarities and either shared a special feature or not (1st sorting). The child was then introduced to another novel name for one of the objects by a new puppet speaking either the child's or a foreign language. Children were then asked to pick all referents of the second novel label from the array (2nd sorting). After objects were placed back, they were reminded of the first label and asked to pick the corresponding objects(3^{rd} sorting). In the same language condition, older children were significantly more likely to correct the first label and avoid objects with corrected labels in the third sorting. No differences between bilingual and monolingual children were observed. But in the different language condition, bilinguals were more likely to accept overlap compared to monolinguals. This was again explained with their constant exposure to multiple labels across languages.

Research with children around the age of two found monolinguals and bilinguals to be equally likely to display a correction effect (Frank & Poulin-Dubois, 2002). The study presented a slightly different version of a correction task, introduced by Savage and Au (1996). In Savage and Au's task, children were taught two novel words for an object by two experimenters. About half of the three to five year olds accepted and used both labels. The other half persevered with the label which was first tested. No age effects were observed. Frank and Poulin-Dubois (2002) presented monolingual and bilingual children, aged 26-28 months and 34-36 months with a correction task. Children were taught a new label for a novel object by one experimenter ("Mido"). Then the other experimenter named the same object with a different novel name ("Gavi"). Children were next presented with a set of four objects, two familiar, an unfamiliar and the previously labelled novel target. Each experimenter then asked for the target, using the label introduced by her, requesting the familiar objects in between as filler items. Monolingual and bilingual children avoided two labels for one object at similar level and significantly more in the older group (younger group: 31 & 39%, older group: 57 & 58%). Bilingual children additionally performed the task in two languages, with one experimenter addressing them in English, the other in French. The same pattern of results occurred across languages (younger group: 38%, older group: 57%). The authors concluded that this was probably the critical time in development when monolingual children start to disambiguate clearly within, and bilingual children additionally between languages.

The sensitivity to cues in regard to word meaning was examined in an experiment including 30 month old children (Brodje, Ahmed, & Colunga, 2012). Children were introduced to a novel object and taught a novel name for it ("zuly"). They were then presented with further novel objects and asked to find more "zuly" on the table. Pragmatic cues (eye gaze) were given towards shape matches or colour/texture matches. Children have been previously found to prefer shape matches to classify object groups (Jones, Smith, & Landau, 1991). Bilingual, but not monolingual children considered the experimenter's eye gaze in situations, when pragmatic cues and object property did not match. Both groups were able to attend to cues in control conditions. The authors suggested, that bilingual children need to examine the speaker's cues to resolve the constant conflict of knowing more than one label for a single object. Young bilinguals might therefore rely on referential cues from adults in slightly ambiguous situations.

The ability to differentiate between the two languages was examined in a study by Tare and Gelman (2010). Bilingual children around the age of three and four, speaking English and Marathi, were tested on a variety of language-based and metacognitive tasks. Children interacted with one experimenter speaking English and one experimenter speaking Marathi. They were presented with objects and asked to name them, once in each language. A free play session with each experimenter also recorded the spontaneous use of the individual language. Both age groups showed differentiation in the free play task, with children using the speaker's language the majority of the time and switching between both experimenters. Children's ability to name objects was stronger in English than Marathi. Children who did not provide the label in the correct language were given up to three prompts to switch. Responsiveness to prompts was significantly correlated to children's metalinguistic awareness (measured by a language check which asked children about the experimenter's languages and how they would name certain objects). Responsiveness to prompts was further related to theory of mind measures (3 tasks from Wellman & Liu, 2004: diverse desire, diverse belief, knowledge access), even after controlling for age effects. The authors suggested that better theory of mind scores enable children to reflect closely on the experimenter's language abilities. They further concluded that metalinguistic awareness related closely to theory of mind, as metalinguistic abilities predicted children's responsiveness to a language only in connection with metacognitive ability.

General cognitive differences between monolingual and bilingual children were the focus of a series of studies. A meta-analysis looking at studies concerned with the cognitive outcome of bilingualism reliably identified increased attentional control and better working memory abilities. Higher metalinguistic and metarepresentational skills were further pointed out (Adesope, Lavin, Thompson, & Ungerleider, 2010).

Increased attentional control was also found in bilingual children (Bialystok & Majumder, 1998), as well as better abilities to inhibit in dimensional card sorting tasks (Bialystok & Martin, 2004). Bilingual children performed comparably to monolinguals on several other cognitive measures, as for example working memory, though.

Further investigations of word mapping and executive inhibition found a higher tendency in monolingual children to disambiguate novel words. Bilinguals performed again better on inhibition tasks. No connection between the tendency to disambiguate and inhibition was found for any of the groups. (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010).

Developmental differences have therefore been found for disambiguation, with bilinguals showing the effect later and less pronounced. This could question the perspectival accounts claim that disambiguation in young children is the result of cognitive immaturity (and in older children and adults the result of a sensible decision). The account would predict bilingual children to disambiguate similarly to monolinguals and suspend the effect when presented with pragmatic cues at the same time as they pass false belief.

It is also of interest whether bilinguals show superior alternative naming performance. They are constantly exposed to multiple labels for one object between languages; they might demonstrate more flexibility within as well. The perspectival account would not predict any differences, as alternative labels specify different perspectives on objects. And the understanding of perspective as indicated by the ability to pass false belief is necessary to flexibly apply alternative labels in a conversation, if not externally induced.

Experiment 4

Method

Participants.

Participants were 46 children (23 girls), recruited from local nurseries and child-minders in Scotland and England. Parents signed a consent form and gave information regarding their children's language background. They also confirmed no history of language or hearing impairment and no exposure of their children to German language through language classes or other teaching. Children were assigned to one of 3 groups: Monolingual children (n = 16, 8 girls, mean age = 48 months, range = 39 – 64, SD = 8) spoke English as their first and only language; Bilingual children (n = 18, 8 girls, mean age = 47 months, range = 31 – 69, SD = 12) spoke English at an age appropriate level and one other language with either one or both parents at home (see Appendix F for list). Parents classed their children as fluent in the home language. Exposed children (n = 12, 7 girls, mean age = 49 months, range = 42 – 62, SD = 6) spoke English as their first and only language and went to regular language classes in their nurseries (Spanish/French). Language teaching was performed once a week for 30 minutes and children had attended classes for a minimum of 3 months. Bilinguals scored significantly lower on the BPVS standardized scores (BPVS standardized scores, monolingual: 101, SD = 14; bilingual: 88, SD = 15; exposed: 100, SD = 8; F (2,43) = .5.261, p = .009; Post-hoc Bonferroni: bilingual vs monolingual, p = .015, bilingual vs exposed, p = .052, no significant difference between monolingual and exposed).

Design.

Children performed ten tasks in randomized order over two sessions. These included a false belief task, an alternative naming task, a forward digit and a backward object span task and the bear–dragon task. Two disambiguation and pragmatic cue tasks presented novel words implicitly in English in one set and explicitly as words in a foreign language in the other. The BPVS III (Dunn, Dunn, Sewell, & Styles, 2009) was administered last.

Procedure and Materials.

False belief task.

The false belief task was as for Experiment 1 and 2.

Alternative naming task.

The materials were the same as in Experiment 3. The procedure was as in Experiment 2.

Disambiguation task and pragmatic cue task, original (implicit English).

Both tasks were identical to Experiment 1 and 2 (materials: Appendix B & G). The pragmatic cue task was followed by a recall task immediately after. (The disambiguation task was not followed by a recall task here, as this was first introduced in experiment 5. Experiment 3 took place after 4 and 5, but was presented before to keep topical continuity.)

Recall.

The child was presented with all five pictures she had selected in response during the pragmatic cue task and asked: “Which one is Jimmy’s hinkle?", testing every novel word used before. Children were encouraged to point to an object for each of the novel words.

Disambiguation task and pragmatic cue task, German.

Children were introduced to the puppet Lisa and told "This is Lisa, she is from a different country, from Germany. Lisa doesn't speak any English, only German. She has some pictures with her to show you." The German disambiguation and pragmatic cue tasks followed the same procedure as the original tasks. Novel words were replaced by German words. ("Lisa is thirsty and would like Saft, please.").

All pictures, novel words, and pragmatic cues for the different trials are listed in Appendix H and I. Children had to select the familiar item on 4 occasions to pass a pragmatic cue task and the novel item on 4 occasions to pass a disambiguation task.

Forward digit span task.

Instructions were as follows: "I'm going to tell you a few numbers now and I would like you to say them after me, ok? For example, I say '5,9' then you say '5,9'. Let's try that: I say 3,8, now you!" If the child responded correctly, a set of 3 numbers was offered next to ensure understanding. If the child did not replicate the two digits, the instruction was repeated and encouragement was given until the child responded. The testing started with a set of two digits, another set of 2 followed. The number of digits was increased by one after each set of 2 (2 sets of 2 digits, 2 sets of 3 digits, and so on) until children responded wrongly to both sets.

Backward object span task.

Training phase.

A set of two pictures was put in front of the child with the instruction: “This is a horse and this is a sheep. Look, I put this horse down first and then this sheep. But I’m going to say the names of these two in a backwards order, so I say sheep (point) then horse (point). Now you say them in a backwards order (pointing at each picture to help).” The next set followed with the experimenter labelling each picture, but only pointing when it was the child’s turn to repeat backwards. This was practiced with another set of 3. Then the experimenter omitted pointing, only asked the child to repeat backwards (after Slade & Ruffman, 2005).

Test phase.

“Now I’m not going to put the pictures down, I’m just going to say the words. I want you to say what I say in a backwards order.” Pictures were provided again if the child failed on the first set of 2. Three sets of 3 and three sets of 4 were administered until the child failed on two consecutive trials.

Bear – Dragon task.

Children were introduced to Nice Teddy and Naughty Dragon. In a training phase they were instructed to perform any action Nice Teddy asked them to do (“Teddy says: Clap your hands!”) but to ignore Naughty Dragons commands. Each command was prefaced by “Teddy says” or “Dragon says”.

This was practiced with 5 commands from each puppet. Children were reminded of the instructions every time they failed. After 5 failures on the dragon trials they were instructed to sit on their hands before responding. The test trial followed immediately with 10 commands, 5 per puppet, in random order (after Reed, Pien, & Rothbart, 1984).

Results 4

Task performance.

Groups were compared on mean performances for the forward digit and backward object span task and the bear dragon task. The remaining tasks had set pass/fail criteria. No significant differences were found on task performance between monolinguals, bilinguals and children exposed to another language (Table 10).

Table 10 Task performance by language group in percentage of children passing/ means

	Monolingual (n = 16)	**Bilingual (n = 18)**	**Exposed (n = 12)**	**Comparison**
FB pass	56	50	75	X^2 (2) = 1.917, p = .384, d = .204
AN pass (≥3)	50	39	42	X^2 (2) = .447, p = .800, d = .099
PC O pass (≥4)	56	50	75	X^2 (2) = 1.917, p = .384, d = .204
PC G pass (≥4)	56	56	58	X^2 (2) = .023, p = .988, d = .023
DT O pass (≥4)	94	61	83	X^2 (2) = 5.549, p = .062, d = .347
DT G pass (≥4)	75	72	75	X^2 (2) = .044, p = .978, d = .031
Forward digit	Mean = 3.375, SD = 1.190	Mean = 2.972, SD = 1.170	Mean = 3.667, SD = .5365	F (2,43) = 1.054, p = .357
Backward object	Mean = 2.94, SD = 4.20	Mean = 3.50, SD = 4.70	Mean =2.75, SD = 2.53	F (2,43) = .144, p = .866
Bear-dragon (≥8)	Mean = 7.75, SD = 2.82	Mean = 7.50, SD = 2.98	Mean = 9.00, SD = 1.21	F (2,43) = 1.32, p = .279

Comparing individual tasks.

Pragmatic cue and disambiguation tasks.

The performance of the three different language groups (number of correct responses 0-5) was analysed using a 3*2*2mixed ANOVA with between participants factor “language background” (monolingual, bilingual, exposed) and within participants factor “cue” (pragmatic/no cue) and “test language” (English/German version). The main effect of “cue” was significant: $F(1,43) = 99.828$, $p = .001$, partial $\eta^2 = .699$. Children were more likely to select the familiar object as referent for the novel word in both pragmatic tasks (PC O: mean = 3.39, SD = 1.56; PC G: mean = 3.43, SD = 1.63) than in the disambiguation tasks (DT O: mean = .83, SD = 1.08; DT G: mean = 1.00, SD = 1.25).

The main effect of test language, the tree-way interaction and none of the two-way interactions were significant (all $F < .89$, all $p > .418$).

The main effect of language background was not significant: $F(2,43) = .839$, $p = .439$, partial $\eta^2 = .038$, children in all 3 groups performed similarly. A planned comparison was performed between monolingual and bilingual children on the original disambiguation task, as the literature suggests differences on the general disambiguation effect (e.g. Davidson et al., 1997; Davidson & Tell, 2005): Bilingual children displayed significantly less disambiguation than their monolingual peers ($t = 3.215$, $df = 32$, $p = .003$, $d = 1.137$; BL: mean 3.83/5, SD = 1.043; ML = 4.75/5, SD = .577).

Recall.

Both pragmatic cue tasks were followed by a recall task. A paired sample t-test for each language background and both recall tasks presented significant differences for bilingual and exposed children (Table 12):

Table 12 Recall after PC tasks (mean ± SD, max. = 5)

	PC Original	**PC German**	**Comparison**
Monolingual	.44 ± .82	.56 ± .63	t = .565, df = 15, p = .580
Bilingual	.44 ± .86	1.67 ± 1.75	t = 2.572, df = 17, p = .020
Exposed	.42 ± .67	1.58 ± 1.83	t = 2.461, df = 11, p = .032

In order to specifically test the assumption that speaking or starting to learn another language might influence the ability to recall novel words, bilingual and exposed children were combined for an analysis. No difference was detected between groups for the ability to correctly identify referents in the implicit English task (t = .024, df = 32, d = .008). The explicit German task revealed a significant group difference, with bilingual and exposed children identifying significantly more German words than monolingual children after only listening to them twice (ML mean = .56, SD = .629; BL+EP mean = 1.67, SD = 1.749; t = 2.503, df = 32, d = .8849).

Alternative naming.

The initial vocabulary check demonstrated children's knowledge about the task labels. Children made on average .64 (SD = .900, range: 0 – 3) mistakes. 26 (56.5%) children identified all pictures correctly, 13 (28.3%) mistook 1 picture, 4 (8.7%) children mistook 2 and 3 (6.5%) children mistook 3 pictures.

The 26 children failing the main task produced on average 2.69 (SD = 1.289) basic and 1.15 (SD = .967) superordinate labels. The 20 children passing produced 3.80 (SD = .410) basic and 3.65 (SD = .489) superordinate labels.

False belief.

The false belief task was passed by answering all 3 questions correctly (belief, reality, memory). 58.7% (27) passed this task, all children answered the reality question correctly, 1 child failed the memory question as well as the belief question.

Forward digit span task.

Children were able to repeat on average 3.29 digits (SD = 1.31, range: 0 – 6). All but 1 child passed the two training trials (1 set of 2, 1 set of 3). Four children did not repeat any more correct numbers after.

Backward object span task.

The training phase was counted as 4 sets, 40 children passed through this phase without mistakes. 6 (13%) children failed, 2 by repeating the labels in forward order, ignoring all instruction. 4 children did not produce any labels. The data was included, as these children responded in the remaining tasks. Children scored 2 points for reversing a set of two words, 3 points for reversing a set of three. The maximum score possible was 27 (6 for three sets of two, 9 for three sets of three, 12 for three sets of 4). The majority of children (52.2%, 24) were able to reverse the first two sets of two words, 36.9% (17) mastered 2x3 words. Three children (6.5%) were able to repeat 3x3 words backwards.

Bear –dragon.

All children were able to pass the initial training phase. The bear's commands seemed easier for most children, 89.1% (41) performed all actions as demanded compared to 50% (23) of children correctly ignoring all dragon's demands. The difference was highly significant: $t = 4.635$, $df = 45$, $p \leq .001$.

Correlation of tasks.

All tasks were entered in a correlational analysis in addition to age and the standardized score of the BPVS III (Table 13).

Table 13 Correlation between task performance, age and verbal-mental age

	VMA	FB	AN	PC O	PC O R	DT O	PC G	PC G R	DT G	FDS	BOS	Bear Dragon
Age	.44**	.52***	.67***	.54***	.54***	.33*	.43**	.20	.20	.68***	.50***	.69***
VMA			.48**	.22	-.06	.36**	.10	.10	.28	.51***	.46***	.51***
False belief			.62***	.68***	.04	-.00	.52***	.22	-.10	.52***	.36*	.48**
Alternative Naming				.63***	.22	.42**	.40**	.08	.18	.59***	.47***	.51***
PC Original					.35*	.09	.62***	.23	-.06	.62***	.35*	.45**
PC O Recall						.08	.13	.11	.11	.12	-.10	.12
DT Original							-.02	-.10	.26	.18	.26	.22
PC German								.42**	-.08	.45**	.22	.23
PC G Recall									-.09	.12	.45**	.03
DT German										.18	.19	.19
Forward DS											.43**	.65***
Backward OS												.36*

* $p < .05$, ** $p < .01$, *** $p < .001$

Further analysing the strong correlation between the pragmatic cue tasks, the alternative naming task and the false belief task, a partial correlation was performed (Table 14). The influence of age, verbal-mental age, memory (forward digit), working memory (backward object) and inhibition (bear dragon) were taken into account.

Table 14 Partial correlations controlling for age, vma, executive functioning

	Alternative naming	**Pragmatic cue, original**	**Pragmatic cue, German**
False belief	.47**	.67***	.45**
Alternative naming		.39*	.14
Pragmatic cue, original			.45**

* p < .05, ** p < .01, *** p < .001

Comparison of tasks by groups.

Children's task performance was further compared for the individual language backgrounds. Monolingual children presented similar correlations to the complete group (Table 15, 16)

Table 15 Correlation of tasks for monolingual children (n = 16)

	VMA	**FB**	**AN**	**PC O**	**PC O R**	**DT O**	**PC G**	**PC G R**	**DT G**	**FDS**	**BOS**	**BD**
Age	.15	.67**	.68**	.36	.28	.05	.32	.18	.27	.55*	.41	.71**
VMA		.30	.15	.03	-.22	-.16	-.20	.06	.35	.46	.27	.47
False Belief			.74***	.69**	.01	-.17	.58*	.19	.04	.62*	.35	.80***
AN				.54*	.22	-.18	.37	.04	.39	.66**	.43	.61*
PC Original					.35	-.06	.73**	.40	.20	.42	.18	.42
PC O Recall						.25	-.10	.27	.31	.06	-.15	.11
DT Original							-.12	.23	-.27	-.15	-.31	-.08
PC German								.19	-.10	.27	-.04	.25
PC G Recall									.03	.28	.29	.09
DT German										.42	.31	.26
Forward DS											.29	.82***
Backward OS												.26

* p < .05, ** p < .01, *** p < .001

Table 16 Partial correlations for monolingual children, controlling for age and vma

	Alternative naming	Pragmatic cue, original	Pragmatic cue, German
False belief	.53 (p = .051)	.69**	.63*
Alternative naming		.44	.25
Pragmatic cue, original			.71**

* p < .05,   ** p < .01,   *** p < .001

Table 17 Correlation of tasks for bilingual children (n = 18)

	VMA	FB	AN	PC O	PC O R	DT O	PC G	PC G R	DT G	FDS	BOS	BD
Age	.68**	.58*	.82***	.70**	.25	.64**	.65**	.32	.10	.74***	.63**	.71**
VMA		.45	.59*	.38	.01	.69**	.54*	.48*	.45	.52*	.65*	.54*
False Belief			.72**	.77***	.13	.27	.59*	.52*	.00	.62**	.49*	.25
AN				.73**	.30	.49*	.64**	.44	.28	.67**	.74***	.54*
PC Original					.35	.42	.56*	.39	-.04	.78***	.59*	.49*
PC O Recall						.29	.17	-.09	.05	.13	.05	.12
DT Original							.39	.29	.44	.46	.57*	.56*
PC German								.48*	.02	.66**	.53*	.37
PC G Recall									-.12	.17	.61**	.14
DT German										.10	.18	.22
Forward DS											.56*	.54*
Backward OS												.44

* p < .05,   ** p < .01,   *** p < .001

Table 18 Partial correlations for bilingual children, controlling for age and vma

	Alternative naming	Pragmatic cue, original	Pragmatic cue, German
False belief	.52*	.66**	.34
Alternative naming		.41	.22
Pragmatic cue, original			.22

* p < .05, ** p < .01

Table 19 Correlation of tasks for children exposed to another language (n = 12)

	VMA	FB	AN	PC O	PC O R	DT O	PC G	PC G R	DT G	FDS	BOS	BD
Age	.37	.18	.23	.27	.66*	.07	-.06	.05	.49	.71**	.13	.71*
VMA		-.18	.25	-.14	.11	.13	-.60*	-.16	-.20	.43	.09	.63*
False Belief			.62*	.94***	.34	-.14	.58*	-.14	-.31	.19	-.47	.17
AN				.55	.06	.56	.01	-.27	-.27	.10	.01	.33
PC Original					.42	-.27	.61*	-.05	-.31	.37	-.52	.26
PC O Recall						-.29	.37	.45	.04	.42	-.15	.23
DT Original							-.45	-.28	.22	-.30	.44	.11
PC German								.56	-.20	.07	-.26	-.39
PC G Recall									-.01	.03	.21	-.49
DT German										.14	.27	.16
Forward DS											-.21	.77**
Backward OS												-.19

* p < .05, ** p < .01, *** p < .001

Table 20 Partial correlations for exposed children, controlling for age and vma

	Alternative naming	**Pragmatic cue, original**	**Pragmatic cue, German**
False belief	.69*	.94***	.58
Alternative naming		.60	.19
Pragmatic cue, original			.64*

* $p < .05$, *** $p < .001$

Overall, the individual groups show slight differences in their task correlation (Tables 17 - 20). Children exposed to another language show strongest correlations between false belief and the pragmatic cue tasks. The correlation between alternative naming and PC falls below significance when age influences are taken into account. Bilingual children show additional correlations between age and the disambiguation task, which have not been observed with the other groups.

Discussion 4

Bilingual children have been proposed to show different metalinguistic awareness. Especially in relation to mutual exclusivity, bilinguals were found to display less of a bias from around the age of five (Davidson et al., 1997; Davidson & Tell, 2005). Bilinguals were also suggested to have better inhibitory abilities around the age of five (Bialystok & Martin, 2004) and outperform monolinguals on IQ subtests (Lauchlan, Parisi, & Fadda, 2012). It was therefore of special interest whether bilingual children performed differently on the pragmatic cue task. If their ability to use pragmatic information was better and unrelated to the understanding of false belief, the perspectival accounts predictions would be incorrect.

Overall, the groups presented strong and similar results to previous experiments. The pragmatic cue task, which was presented implicitly as English and additionally explicitly in a foreign language correlated well with the other metarepresentational tasks. No language effects were found comparing the three different groups on metalinguistic awareness. Children between the age of three and five were therefore similar in their ability to pass false belief, alternative naming and pragmatic cue. Previous research which pointed towards differences for the false belief task and children from multilingual backgrounds could not be confirmed (Farhadian et al., 2010; Goetz, 2003). Kovács (2009) presented bilingual children nearly 10 months younger than children in this study to outperform monolinguals on false belief performance. The bilinguals in her study passed false belief around the age of three to nearly 60%, compared to 25% of the monolingual matched controls. Children in Kovác's study were exposed to a much richer bilingual field, though. The children visited bilingual nurseries and had a parent speaking each of the two languages at home. Having two parents speaking a different language might have a bigger impact on early theory of mind development, than growing up with one language at home and speaking another language at nursery and school. The bilinguals in Experiment 4 were mostly from one-language parents and spoke English as second language. This might have contributed to their similar behaviour on metarepresentational tasks.

Bilinguals in Experiment 4 displayed significantly less disambiguation. Previous studies presented bilinguals mostly displaying less disambiguation effect slightly later (~the age of 5, for example Davidson et al., 1997). This seemed not to relate to their ability to use pragmatic information for object choice in the presented study, though. The different performance of bilinguals was further only observed in the original task, not in the German disambiguation task. No order effects of task presentation were observed. Bilinguals' tendency to select familiar objects more often in the implicit English than the explicit German disambiguation task is therefore slightly puzzling.

Looking at task correlations on the individual group levels, all relationships were observed for the original three tasks. The new, explicit German pragmatic cue task was slightly less stable. Children exposed to another language also presented no significant correlations between pragmatic cue and alternative naming. The number of children in this group was quite small and their performance was very unstable in general. Their performance gives an indication, though, that even small foreign language influences as weekly classes might contribute to different behaviour in metarepresentational tasks.

None of the working memory or inhibition tasks presented differences for the three groups, contrary to previous research which had shown bilingual children to perform better on inhibition tasks, relating this to better conceptual and attentional control (Bialystok & Martin, 2004).

An interesting finding occurred around foreign language recall. Recall after the original pragmatic cue task was very low and did not differ between groups. But given the task instruction of facing a foreign language, bilingual children were able to outperform their monolingual peers. Children exposed to another language performed nearly as well as bilinguals and significantly better than in the original task. One possibility was the use of “real” German words. Children might be able to relate German words to their own second language, recognizing familiarities across other indo-european languages. This possibility could not be addressed in the presented studies and might be worthwhile to follow up on.

The differences in recall asked for replication of the study and the inclusion of a recall task after the disambiguation task. This would show whether recall was dependent on task instruction, the combination of task instruction and pragmatic cue or a simple one time finding.

Experiment 5

Method

Participants.

A total of 50 children (21 girls, mean age = 47 months, range 36 – 69, SD = 8) participated in this study from nurseries and child minders in Scotland and England. The study took place between 4 and 7 months after the previous and involved 19 children (6 monolingual, 6 bilingual, 7 exposed) of the earlier sample. Preliminary analyses indicated no differences between participants who were and were not included in Experiment 4. As all materials were also different to the previous study, their data was fully included.

Children were divided into three groups depending on their language abilities, which parents confirmed through consent forms: Monolingual children (n = 18, 9 girls; mean age = 44 m, range = 37 – 59, SD = 6) spoke English as their mother tongue and had no history of foreign language learning. Bilingual children (n = 15, 5 girls; mean age = 50 m, range = 36 – 69, SD = 10) spoke English at age appropriate level and another language at home (see Appendix J for list of languages). Exposed children were monolingual English speakers and attended regular language classes (Spanish/French) at nursery once a week for 30 minutes for a minimum of 3 months (n = 17, 7 girls; mean age = 47m, range = 38 – 56, SD = 6). The groups did not differ significantly on age ($F(2,47) = 1.146$, $p = .327$) or receptive English vocabulary as measured by the BPVS ($F(2,46) = 3.044$, $p = .057$).

Design.

Children participated in the following tasks over the course of two settings in random order: a false belief task, two disambiguation and pragmatic cue tasks, one set implicit in English and one set explicitly using words in a foreign language. Children further performed four subtests of the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 1967), including two verbal tests (arithmetic & sentences) and two performance tests (maze & block design). The BPVS III (Dunn et al., 2009) was administered last.

Procedure and Materials.

False Belief task.

The false belief task was as for Experiment 1, 2 and 4.

Disambiguation task and pragmatic cue task, original and German.

The disambiguation and pragmatic cue tasks were as in Experiment 4. New novel and German words were used as well as new pictures (Appendix K – N). The disambiguation tasks were additionally followed by a recall task, in the same way as the pragmatic cue tasks.

WPPSI, verbal and performance tests.

Materials and procedure for the WPPSI were applied as outlined in the WPPSI manual.

Arithmetic subtest.

Materials consisted of a booklet with four cards, displaying a set of objects, nine flat square wooden blocks, painted red and a list of arithmetic questions. Children were presented with the booklet first and asked to point to one of the objects displayed (e.g. “Here are some balls. Which is the biggest? Point to it.”). After all cards were displayed, two wooden blocks were placed in front of the child and the child was asked: “How many blocks are there?”. This was repeated with four and nine blocks. Leaving the nine blocks on display, the child was told: “Now give me all of the blocks except four. Leave four of the blocks here.” . The last part of the task comprised a set of arithmetic questions which were read out to the child slowly (e.g., “If I cut an apple in half, how many pieces will I have?”). The task was discontinued after four consecutive failures. A full list of all questions can be found in Appendix O. Children received one point for each correct response, the maximum score was 20.

Sentences subtest.

A list of sentences was read out to the child one at a time and the child was required to repeat the sentence precisely. The child was told: “I’m going to say something, and I want you to say it after me, just the way I say it. Ready? Listen.”. Following the WPPSI, each sentence was only read out once, if children asked for it again, they were told to guess. The task was continued until children failed on three consecutive sentences. Errors on each response were counted as follows: each word omitted was counted as one error, the transposition of a word or a phrase was counted as one error, the addition of one or more words was counted as one error, the substitution of words were counted at the rate of one error for each word omitted. A full list of sentences and score table can be found in Appendix P.

Maze subtest.

Materials included ten sheets of paper, displaying a maze each. Mazes 1-3 presented a chick on one side of the maze, mother hen on the other. Maze 4 and 5 displayed a stickman in the center of the maze. In mazes 6-10 the stickman was replaced by a cross. The experimenter used a black pen, the child was given a red one. Instructions followed the WPPSI: “See this little chick? It wants to get to its mother on the other side (point), and it must get there by keeping inside the road and not going into the blocked road like this. (Point to the first blind alley.) Watch me. The chick starts here and then goes this way.” The experimenter drew a line from the chick, following the maze, then stopped in front of the first blind alley. “No, you wouldn’t go in here; it’s wrong.” (Wechsler, 1967, p.64). The experimenter continued and asked the child to finish the second half of the maze by herself. The second maze was identical to the first and the child was encouraged to complete this one from the start. Maze 2 and 3 followed, the child received feedback when mistakes as drawing over the line or ending in a blind alley occurred and a new version of the maze was used to try again once more. Maze 3 was only presented if the child mastered maze 2 without errors on first or second trial, maze 4 only if maze 3 was mastered the second time without mistakes. The experimenter then demonstrated how to show the stickman the way out of the maze for the next picture, the child was encouraged to do the same with maze 4 and 5, and to start at the cross for maze 6 and following. Mazes 4 to 6 allowed 1 error to precede to the next, maze 7 and 8 allowed two errors and maze 9 and 10 three errors. An example of mazes can be found in Appendix Q as well as a score table.

Block design subtest.

A set of six flat wooden blocks, one side painted red, the other side white and 8 blocks red on one side, half red – half white on the other side was set aside for this task. A booklet with three printed cards was used for the last part of the task. The experimenter used a score sheet, containing designs of models for two to four blocks. The first set of three blocks was set up behind a screen and then presented to the child. The experimenter then demonstrated with a second set of three blocks how to replicate the design. The child was then given three blocks and asked to do the same, receiving feedback and also being allowed to try again. Design 2 was presented next. The task was discontinued if the child failed to replicate the design 1 and 2 on both trials. Otherwise, design 3 and 4 were administered. The task was stopped if the child failed on both trails of 2 consecutive designs after design 3. Designs 8 to 10 were presented as pictures on cards. The child was asked to arrange the blocks to make them look like the picture. Each design scored 2 points if replicated on first trial, one point if mastered on second trial and zero points if failed. A full set of all designs can be found in Appendix R.

Results 5

Task performance

Children's performance was compared between the three language backgrounds. The false belief task, both pragmatic cue and both disambiguation tasks had set pass/fail criteria. Mean performance was compared for the subtests of the WPPSI. No significant differences were found between monolingual and bilingual children and children exposed to another language (Table 21).

Table 21 Task performance by language background in percentage/means

	Monolingual	Bilingual	Exposed	Comparison
FB pass	50	33	65	X^2 (2) = 3.137, p = .208, d = .250
PC O pass (≥4)	67	53	71	X^2 (2) = 1.117, p = .572, d = .149
PC G pass (≥4)	61	53	77	X^2 (2) = 1.953, p = .377, d = .198
DT O pass (≥4)	67	53	77	X^2 (2) = 1.907, p = .385, d = .195
DT G pass (≥4)	67	60	41	X^2 (2) = 2.445, p = .295, d = .221
WPPSI Arithmetic	Mean = 5.78, SD = 2.80	Mean = 6.33, SD = 4.05	Mean = 6.71, SD = 2.31	F (2,47) = .401, p = .672
Sentences	Mean = 7.94, SD = 3.95	Mean = 8.40, SD = 6.07	Mean = 8.35, SD = 3.61	F (2,47) = .051, p = .950
Maze	Mean = 2.50, SD = 2.66	Mean = 3.67, SD = 3.90	Mean = 4.35, SD = 3.66	F (2,47) = 1.320, p = .277
Block design	Mean = 4.50, SD = 4.53	Mean = 5.60, SD = 5.24	Mean = 4.41, SD = 2.55	F (2,47) = .387, p = .681
WPPSI combi, proportion correct	Mean = 21.09%, SD = 12.42	Mean = 24.62%, SD = 17.08	Mean = 24.22%, SD = 8.49	F (2,47) = .366, p = .695

Comparing individual tasks.

Pragmatic cue and disambiguation task.

A 3*2*2 mixed ANOVA with between participants factor “language background” (monolingual, bilingual, exposed) and within participants factor “cue” (pragmatic/no cue) and “test language” (English/German version) was performed on task performance. The main effect of “cue” was significant: F (1,47) = 60.98, p < .001, partial η^2 = .565. The familiar object was more likely to be selected as referent for the novel word in both pragmatic cue tasks (PC O: mean = 3.62, SD = 1.52; PC G: mean = 3.50, SD = 1.43) than in the disambiguation tasks (DT O: mean = 1.14, SD = 1.34; DT G: mean = 1.82, SD = 1.80).

There was a significant interaction between cue and test language (F (1,47) = 7.23, p = .010, partial η^2 = .133), with children showing less disambiguation effect in the German task than in the original version, but no clear difference for the pragmatic cue tasks. The main effect of test language was not significant: F (1,47) = 3.51, p = .067, partial η^2 = .069.

The interaction between test language and language background was not significant, neither was the three-way interaction or the main effect of language background (all F < 2.49, all p > . 094).

Experiment 4 had demonstrated less disambiguation effect in bilinguals compared to monolinguals (t = 3.215, df = 32, p = .003, d = 1.137; BL: mean 3.83/5, SD = 1.04; ML = 4.75/5, SD = .58). These findings could not be replicated (t = .587, df = 31, p = .561, d = .21; BL: mean 3.53/5, SD = 1.25; ML = 3.83/5, SD = 1.62).

All tasks were followed by a recall task (Table 22). Analysis discovered a significant three-way-interaction between “cue”, “test language” and “language background” for recall: F (2,47) = 4.69, p = .014, partial η^2 = .166. A follow up revealed that bilingual children recalled significantly more German words after the pragmatic cue task than both other groups: F (2,47) = 4.09, p = .023; post-hoc Bonferroni: bilingual/monolingual, p = .055; bilingual/exposed, p = .040.

Table 22 Recall after PC and DT tasks (mean ± SD, max. 5)

	PC original	**PC German**	**DT original**	**DT German**
Monolingual	1.28 ± 1.27	.78 ± .81	1.06 ± 1.11	1.28 ± 1.32
Bilingual	1.33 ± 1.35	2.00 ± 1.85	1.73 ± 1.58	1.27 ± 1.03
Exposed	1.24 ± 1.30	.71 ± .92	1.41 ± 1.28	1.18 ± 1.07

False belief.

To pass the false belief task all 3 questions had to be answered correctly (belief, reality, memory). 50% (25) of the children passed this task. 2 children failed the reality question and 3 the memory question as well as the belief question.

WPPSI.

Children were scored in accordance to the WPPSI manual (Wechsler, 1967). No significant differences on mean performance scores were observed between the three language groups (Table 21).

Correlation of tasks.

A correlational analysis was performed between all tasks, age and the standardized scores of the BPVS III (Table 23). The subtests of the WPPSI correlated strongly, as well as to age and verbal-mental age. Following the correlation between the pragmatic cue tasks and false belief, a partial correlation was performed, partialling out age, verbal-mental age and WPPSI combi (Table 24), presenting stable and highly significant correlations.

Correlations for the individual language groups were not as strong (Tables 25 – 30).

Table 23 Correlation between tasks

	VMA	FB	PC O	PC O R	DT O	DT O R	PC G	PC G R	DT G	DT G R	Arithmetic	Sentences	Maze	Block Design	WPPSI combi
Age	.42**	.31*	.36**	.17	.19	.37**	.29*	.25	-.04	.21	.68***	.70***	.62***	.55***	.75***
VMA		.28*	.36*	.21	.21	.35*	.34*	.06	.01	-.06	.55***	.53***	.37**	.32*	.51***
False Belief			.65***	-.03	.26	.05	.55***	.12	-.19	.25	.33*	.34*	.42**	.31*	.41**
PC Original				-.08	.18	.08	.68***	.02	-.14	.23	.37**	.39**	.32*	.25	.39**
PC O Recall					.19	.27	-.03	.20	.08	.14	.33*	.26	.07	.09	.21
DT Original						.09	.10	.09	.42**	.24	.30*	.40**	.21	.25	-.34*
DT O Recall							.20	-.17	.13	.26	.45**	.29*	.19	.15	.31*
PC German								.04	-.24	.19	.31*	.23	.34*	.19	.31*
PC G Recall									.10	.09	.19	.34*	.17	.26	.29*
DT German										.07	.16	.13	-.22	-.03	-.01
DT G Recall											.15	.18	.03	.10	.14
Arithmetic												.72***	.57***	.60***	.86***
Sentences													.54***	.59***	.84***
Maze														.54***	.77***
Block Design															.86***

*p < .05, ** p < .01, ***p < .001

Table 24 Partial correlations, controlling for age, verbal-mental age and IQ

	Pragmatic cue, original	**Pragmatic cue, German**
False belief	.58***	.48**
Pragmatic cue, original		.62***

** p < .01,   ***p < .001

Table 25 Partial correlations, monolingual group

	Pragmatic cue, original	**Pragmatic cue, German**
False belief	.69**	.75**
Pragmatic cue, original		.71**

** p < .01

Table 26 Partial correlations, bilingual group

	Pragmatic cue, original	**Pragmatic cue, German**
False belief	.17	.40
Pragmatic cue, original		.66*

*p < .05

Table 27 Partial correlations, exposed group

	Pragmatic cue, original	**Pragmatic cue, German**
False belief	.82**	.32
Pragmatic cue, original		.46

** p < .01

Table 28 Correlation between tasks for monolingual group

	VMA	FB	PC O	PC O R	DT O	DT O R	PC G	PC G R	DT G	DT G R	Arithmetic	Sentences	Maze	Block Design	WPPSI combi
Age	.64**	.44	.29	.25	-.02	.68**	.18	.10	.11	.18	.63**	.57*	.73**	.56*	.72*
VMA		.28	.31	.39	.03	.53*	.18	.22	.09	-.13	.52*	.52*	.51*	.32	.52*
False Belief			.71**	.22	.04	.36	.73**	.28	.42	.48*	.16	.33	.58*	.24	.35
PC Original				.11	.00	.26	.72**	.11	.25	.40	.33	.37	.40	.05	.28
PC O Recall					-.14	.61**	.02	.18	-.07	.37	.41	.52*	-.01	.27	.36
DT Original						-.01	.29	-.38	.69**	-.44	-.03	-.45	.39	-.18	-.12
DT O Recall							.15	.15	.18	.19	.55*	.61**	.39	.45	.58*
PC German								-.04	.45	.18	.10	.05	.49*	.11	.19
PC G Recall									-.03	.45	.37	.57*	.11	.45	.46
DT German										-.13	.02	-.19	.48*	-.05	.04
DT G Recall											.35	.49*	.13	.41	.43
Arithmetic												.73**	.57*	.75***	.91***
Sentences													.29	.59*	.77***
Maze														.58*	.70**
Block Design															.92***

*p < .05, ** p < .01, ***p < .001

Table 29 Correlation between tasks for bilingual group

	VMA	FB	PC O	PC O R	DT O	DT O R	PC G	PC G R	DT G	DT G R	Arithmetic	Sentences	Maze	Block Design	WPPSI combi
Age	.43	.46	.59*	.04	-.42	.30	.51	.24	-.16	.26	.78**	.86***	.75**	.67**	.84***
VMA		.32	.22	.36	-.37	.44	.41	.19	-.46	.03	.66*	.62*	.50	.52	.64
False Belief			.39	-.18	-.39	-.06	.53*	.36	.22	.24	.41	.56*	.48	.39	.50
PC Original				-.28	-.26	.04	.75**	.08	-.06	.24	.44	.51	.36	.30	.44
PC O Recall					-.40	.45	-.08	.05	-.28	-.07	.28	-.07	.17	-.10	.06
DT Original						-.37	-.28	-.15	.40	-.22	-.53*	-.43	-.63*	-.45	-.56*
DT O Recall							.39	-.34	-.57*	.27	.61*	.24	.25	.08	.32
PC German								.07	-.08	.34	.54*	.50	.35	.21	.43
PC G Recall									.01	-.09	.08	.34	.43	.21	.28
DT German										-.30	-.36	-.26	-.17	-.02	-.21
DT G Recall											-.01	.03	-.05	-.07	-.03
Arithmetic												.85***	.80***	.62*	.89***
Sentences													.85***	.73**	.94***
Maze														.69**	.90***
Block Design															.87***

*p < .05, ** p < .01, ***p < .001

Table 30 Correlation between tasks for exposed group

	VMA	FB	PC O	PC O R	DT O	DT O R	PC G	PC G R	DT G	DT G R	Arithmetic	Sentences	Maze	Block Design	WPPSI combi
Age	.29	.10	.16	.30	-.19	.11	.21	.22	.23	.34	.50*	.49*	.32	.18	.53*
VMA		.09	.41	-.07	-.23	.15	.50*	.13	.12	-.02	.45	.55*	.06	.30	.46
False Belief			.83***	-.16	-.48	-.05	.37	.03	-.19	.01	.45	.11	.28	.57*	.52*
PC Original				-.10	-.38	-.01	.59*	.02	.13	-.03	.35	.24	.27	.77***	.59*
PC O Recall					-.10	-.21	-.06	.58*	.08	.01	.34	.48	.05	.12	.33
DT Original						.06	-.64**	.01	.12	.08	-.45	-.35	-.46	-.27	-.56*
DT O Recall							.07	-.42	.01	.40	.04	.06	-.07	-.27	-.10
PC German								.12	.15	.12	.33	.18	.19	.41	.40
PC G Recall									-.34	.06	.43	.20	-.25	-.03	.10
DT German										.24	-.23	.11	.31	.33	.21
DT G Recall											.12	-.02	.05	-.30	-.06
Arithmetic												.41	.20	.31	.67**
Sentences													.30	.21	.66**
Maze														.40	.73**
Block Design															.71**

*p < .05, ** p < .01, ***p < .00

Discussion 5

Questions addressed with this study were whether differences in ability to recall novel words explicitly presented in a foreign language could be found again in children from different language backgrounds. Children additionally matched on their performance on the WPPIS and can so reasonably be compared on other measures.. General findings of children being more likely to select the familiar object in the pragmatic cue tasks than the disambiguation tasks were repeated without any differences between the three language backgrounds.

The bilingual children in Experiment 5 also disambiguated to the same level as monolingual children. This was due to monolingual children in this study displaying less disambiguation than in study 4. No effects of order (PC/DT) or age differences between the two studies could be observed. Bilingual disambiguation behaviour observed in the last two studies fits previous research therefore as disambiguation was detected from around the age of 3 (Davidson et al., 1997; Frank & Poulin-Dubois, 2002).

The groups performed overall similarly on the pragmatic and disambiguation tasks. The new German pragmatic cue items seemed more stable than the previous and correlated well with false belief and the original task, staying strong after the influence of age and verbal-mental age were accounted for. Looking at the correlations for the individual groups, monolingual children show all correlations between PC tasks and false belief after age and verbal-mental age is accounted for. Bilingual children displayed weaker correlation between the pragmatic cue tasks and false belief, slightly less children passed false belief in this group. Exposed children showed again correlations between false belief and the original pragmatic cue task. The German PC task is therefore a further metarepresentational task, comparable to the implicit English task especially for children with no foreign language background. Foreign language teaching might interact with the reaction to an explicit foreign language task. Group sizes were quite small, though.

Recall was overall better than in Experiment 4 for the original PC task. Recall also did not differ between groups for PC original and both disambiguation tasks. The German word recall presented differences again. Bilingual children recalled more German words than their monolingual peers and now also more than children exposed to another language. It needs to be pointed out, though, that the two groups recalling fewer German words had lower means in this task than in the other three recalls. A novel word explicitly from a foreign language might seem too much effort or not attract enough attention for monolinguals to memorize. Bilinguals recalled more German words than implicit English in the PC and DT task and German words in the DT task. The cue in the German task might therefore to be of use for Bilinguals to initiate memorizing novel words.

Why the two other groups found it harder to memorize German words in the PC task can only be speculated about. Differences in language recall have been studied in adult monolingual and bilingual samples (Kaushanskaya & Marian, 2009; Kaushanskaya & Rechtzigel, 2012). The authors found bilingual adults to perform better in novel word recall. It was suggested that this was due to more tolerant and efficient phonological encoding in bilinguals combined with better inhibition. Especially concrete terms were better and faster recalled by bilinguals. This was also connected to higher activation due to the dual language activation. Bilingual children might display superior recall out of the same reasons. They might additionally need the task instruction to be faced with a foreign language, as they did not display better recall in the implicit English PC task. The pragmatic cue also added certainty that the familiar object was the correct choice, there is no certainty in the disambiguation task, just a sensible choice to make. Hence, no better recall in the German disambiguation task. The adults in the study mentioned above found it harder to memorize abstract words, which might also be similar to children memorizing names for unfamiliar and more abstract objects.

The third group, children exposed to other languages by language classes, behaved slightly differently in this experiment than in the previous study. Their performance was more similarly to bilinguals on recall in Experiment 4. In this study, they behaved more similarly to monolingual children. The amount of time spend learning another language might have influenced this, although all children had been exposed to language classes for at least three months and not more than 15 months. This did not differ to the previous study. Individual differences might account for mean differences on recall. More exact measures of the amount of classes attended might give further insight here.

Overall, Experiment 4 and 5 did not reveal major differences between the language backgrounds. Previous research findings like enhance inhibitory control or other memory advantages in bilinguals could not be repeated. This might be due to the mixed amount of bilingualism presented in the sample, where most children spoke English as their second language and their mother tongue at home with both parents.

Children attending regular language teaching displayed a spread of performance which was more similar to bilinguals on some tasks and more similar to monolinguals at other times on the same tasks. This asks for thorough screening of language exposure for studies like this. Many parents would not consider their children to be experiencing foreign language influences due to the small amount of teaching in nursery. But differences were observed.

The major finding around novel label recall points to a superior ability of bilinguals to learn new words if they are explicitly presented as a foreign language. This ability has been observed in adults before. This is the first study to show similarities in children between the age of three and five.

General discussion

The metacognitive and metalinguistic development of monolingual and bilingual children was investigated. Findings in order of experiments were as follows:

1) Children can suspend the disambiguation effect when given pragmatic information that this is sensible at the same time as they pass other metacognitive and metalinguistic tasks. The new pragmatic cue task correlates well with false belief and alternative naming.
2) Children's ability to use the pragmatic cue is not influenced by impulsive reactions to hearing the cue first. Executive functioning abilities like inhibition did not relate to children's object choice.
3)
    a) The Pragmatic cue task was further refined by providing the cue before showing the pictures. Recall tasks demonstrated actual word learning in the pragmatic cue and disambiguation task. The pragmatic cue paradigm was extended to the correction effect. Children tended to correct one of two novel labels supplied by the experimenter even when faced with strong pragmatic information to allow overlap. The correction effect correlated with other metacognitive tasks but persisted even with older children.
    b) The misnaming task was included. Previous findings by Waxman and Hatch (1992)showed preschool children producing multiple labels in one discourse. These findings were not replicated. Children were not able to produce labels on different categorical levels, even after the task was simplified.

4) Bilinguals:

   a) The pragmatic cue findings were replicated with bilingual children and children exposed to another language. Both new groups performed similarly to their monolingual peers. Bilinguals showed a slight advantage in recalling novel foreign language labels.

   b) Bilinguals and children exposed to another language performed similarly to monolinguals on additional tasks concerning inhibition and general intelligence. There was some support for bilingual's advantage in foreign language learning.

The perspectival account can account for all data presented. The account predicts a connection between metalinguistic abilities and theory of mind development. Perner (2000) explained the understanding of perspective to be observed in the understanding of beliefs. The same perspectival understanding is necessary to pass false belief and alternative naming (Perner et al., 2002). The pragmatic cue task extends this into the area of word learning showing strong correlations to the former two tasks. The connection proved repeatable and stable, taking the influence of age, verbal mental age and executive functioning into account.

Concerns regarding young children's possible impulsive reaction towards cues were thoroughly addressed. Variations of the PC task, presenting the cue before or after the novel word (Experiment 2), turning the pictures over after the novel words were introduced and asking whether the child knew the novel word (Experiment 3) did not produce any differences in children's performance. The ability to inhibit a response was not a determining outcome in the pragmatic cue task.

Although disambiguation tasks are used to investigate word learning, none of the studies cited tested retention of novel words in preschool children (e.g.Diesendruck & Markson, 2001; Haryu, 1991; Markman & Wachtel, 1988). Recall was therefore examined in Experiment 3a and b. Children identified more than half of all novel words in both disambiguation tasks, the original and the cued version. Brief exposure is therefore sufficient for children at this age to identify referents of novel words again. Hence, disambiguation and pragmatic cue task were shown to assess word learning.

The misnaming task addressed previous findings that children can produce multiple labels (Waxman & Hatch, 1992). In Experiment 3a, children mostly persevered with their initial response and did not take the experimenter's misnaming as a cue to alter their response and produce a different categorical label. No relationship between the misnaming task and the alternative naming task was detected. The task does not appear to reliably produce alternative naming in this age range, even after simplifying the task in Experiment 3b.

The perspectival account predicts that children display a different word learning behaviour in regard to their understanding of perspective (Doherty & Perner, 2013; Perner et al., 2002). Investigating bilingual's word learning was therefore essential, as they were often found to show less disambiguation effect (Davidson et al., 1997; Davidson & Tell, 2005; Houston-Price et al., 2010). Bilingual children and children exposed to another language were shown to pass false belief, alternative naming and pragmatic cue at the same time. Their ability to use pragmatic information to change disambiguation behaviour was similar to monolinguals. The perspectival accounts predictions were therefore further supported.

Further, this was the first presentation of the alternative naming task (after Perner et al., 2002) in bilingual children. Bilingual children displayed the same difficulties producing alternative labels for familiar object as their monolingual peers until they passed false belief. Hence, within one language, bilingual children show the same metalinguistic awareness and metacognitive abilities as monolingual children.

No significant performance differences were detected for the different language groups apart from slightly better foreign language learning in the bilingual group. Research into adult language learning pointed to bilingual's superior encoding, storage and retrieval of novel words (Kaushanskaya & Marian, 2009; Kaushanskaya & Rechtzigel, 2012). Children in the conducted studies show early indications of this later advantage.

Theoretical considerations.

The lexical principles approach claims the mutual exclusivity bias to be a useful word learning heuristic (Markman & Wachtel, 1988; Merriman & Bowman, 1989). In Experiment 1 to 5, children tended to behave as if nouns were mutually exclusive in all the disambiguation tasks, but they displayed different behaviour in the pragmatic cue tasks. Here, their referent selection was associated with their ability to correctly identify false belief based judgements. A possible explanation along the lines of the lexical principles account could be that with development children become able to take more additional information into account. This allows them to decide, in the case of the PC task, that the information is sufficient to abandon the mutual exclusivity bias for the familiar object, like on a case-by-case scenario (Markman, 1989). But the lexical principles account does not predict when children can take this information into account. Especially, no association with metacognitive or metalinguistic development is predicted. The clear developmental change in children's behaviour demonstrated by the pragmatic cue task can therefore not be explained by the lexical principles approach, as well as the correlation of false belief understanding and ability to use the cue for referent selection.

The explanations of the sociopragmatic account (Bloom, 2000; Diesendruck & Markson, 2001) also cannot account for the data presented. Children's behaviour in the pragmatic cue task cannot be explained by their ability to interpret speaker intention If children could determine word meaning by "reading the speaker's mind" as suggested by supporters of the socio-pragmatic account, they would easily select the familiar object as referent. The cue given is obvious. Children reacting only on the cue should always select the familiar object as it is the only logical choice (e.g. Susanne Grassmann & Tomasello, 2010). But only children passing false belief consistently do. The sociopragmatic account cannot explain why children disregard very clear pragmatic information and, similarly to the lexical principles account, cannot explain the developmental change in referent selection in relation to passing false belief.

The perspectival account predicts a connection between metalinguistic abilities and theory of mind development (Perner et al., 2002). The pragmatic cue task tested throughout the studies shows a strong and stable relationship to alternative naming and false belief. The perspectival account's predictions can therefore confidently be extended into the area of word learning.

Limitations.

An issue of the novel pragmatic cue task was that younger children seemed to disambiguate less in the pragmatic cue task than in standard disambiguation tasks. They seemed to start using pragmatic information, but unreliably. An attempt to look closer at this development was made in Experiment 3 (which took place after Experiment 4 and 5). But the lack of sufficient young participants did not result in enough data to analyse the problem thoroughly. It seems as if the transition from false-belief-fail to false-belief-pass is paralleled in the pragmatic cue task with a more gradual consideration of pragmatic information. This needs to be addressed in future studies.

Experiment 4 and 5 considered bilingual children from a variety of language backgrounds. All children were able to express themselves in two languages, but it might be important to consider more thorough screening in the future. First, children with two parents speaking two different languages and children with one foreign language spoken at home should be considered separately. It is unknown whether children who speak one language at home and another in nursery/school and children who grow up with two languages at home show the same linguistic developmental pattern.

Also, children taking part in the presented studies came from many different language combinations. A more concise study design would focus on a bilingual combination of English and one other language. This was not possible in the timeframe give for this thesis, but has to be considered for future work.

Conclusion.

The disambiguation effect in young children was demonstrated to be the result of a cognitive limitation. The relationship between theory of mind, alternative naming and the pragmatic cue task suggests that this limitation is a lack of understanding of perspective.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism. *Review of Educational Research, 80*(2), 207-245.

Au, T. K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development, 61*(5), 1474-1490.

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word Mapping and Executive Functioning in Young Monolingual and Bilingual Children. *Journal of Cognition and Development, 11*(4), 485-508.

Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics, 19*(01), 69-85.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325-339.

Bion, R. A. H., Borovsky, A., & Fernald, A. (2012). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition, 126*, 39-53.

Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

Brand, R. J. (2000). IV. Learning Novel Nouns: Children use Multiple Cues. *Monographs of the Society for Research in Child Development, 65*(3), 41-61.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology, 22*(6), 1482-1493.

Brodje, C. L., Ahmed, S., & Colunga, E. (2012). Bilingual and monolingual children attend to different cues when learning new words. *Frontiers in Psychology, 3*(155), 1-11.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science, 12*(5), 815-823.

Clark, E. V. (1978). Awareness of language: Some evidence from what children say and do. In A. Sinclair, R. J. Jarvella & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 17-43). Berlin: Springer-Verlag.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1-33). Hillsdale, NJ: Erlbaum.

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language, 15*(2), 317-335.

Davidson, D., Jergovic, D., Imami, Z., & Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of Child Language, 24*(1), 3-24.

Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology, 92*, 25-45.

de Marchena, A., Eigsti, I.-M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition, 119*(1), 96-113.

Diesendruck, G. (2005). The principle of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology, 41*(3), 451-463.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology, 37*(5), 630-641.

Doherty, M. J. (2009). *Theory of mind: How children understand others' thoughts and feelings*. Hove: Psychology Press.

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

Doherty, M. J., & Perner, J. (2013). *By any other name: Theoretical integration of dual naming problems in development*. Manuscript submitted for publication.

Dunn, L. M., Dunn, D. M., Sewell, J., & Styles, B. (2009). *British Picture Vocabulary Scale: 3rd Edition*. Windsor: NFER-Nelson.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale: 2nd Edition*. Windsor: NFER-Nelson.

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M., Gazanizadand, N., & Kumar, V. (2010). Theory of mind in bilingual and monolingual preschool children. *The Journal of Psychology, 1*(1), 39-46.

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* (pp. 244-267). Cambridge: Cambridge University Press.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*(1), 99-103.

Frank, I., & Poulin-Dubois, D. (2002). Young monolingual and bilingual children's responses to violation of the Mutual Exclusivity Principle. *The International Journal of Bilingualism, 6*(2), 125-146.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children 3 1/2 - 7 years old on a stroop-like day-night test. *Cognition, 53*(2), 129-153.

Goetz, P. G. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and Cognition, 6*, 1-15. doi: 10.1017/S1366728903001007

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21*, 122-155.

Gombert, J. E. (1990/1992). *Metalinguistic development*. London: Harvester Wheatsheaf.

Grassmann, S., Stracke, M., & Tomasello, M. (2009). Two-year-olds exclude novel objects as potential referents of novel words based on pragmatics. *Cognition, 112*, 488-493.

Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science, 13*(1), 252-263.

Halberda, J. (2003). The development of a word-learning strategie. *Cognition, 87*, B23-34.

Haryu, E. (1991). A developmental study of children's use of "mutual exclusivity" and context to interpret novel words. *Japanese Journal of Educational Psychology, 39*(1), 11-20.

Haryu, E., & Imai, M. (1999). Controlling the application of the mutual exclusivity assumption in the acquisition of lexical hierarchies. *Japanese Psychological Research, 41*(1), 21-34.

Horst, J. S., Samuelson, L. K., Kucker, S. C., & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition, 118*(2), 234-244.

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy, 15*(2), 125-150.

Jaswal, V. K. (2010). Explaining the disambiguation effect: Don't exclude mutual exclusivity. *Journal of Child Language, 37*, 95-113.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*(3).

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review, 16*(4), 705-710.

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review, 19*(5), 935-941.

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science, 12*, 48-54.

Lauchlan, F., Parisi, M., & Fadda, R. (2012). Bilingualism in Sardinia and Scotland: Exploring the cognitive benefits of speaking a 'minority' language. *International Journal of Bilingualism*.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology, 20*, 121-157.

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*(241-275).

Mather, E. (2013). Bootstrapping the early lexicon: how do children use old knowledge to create new meanings? *Frontiers in Psychology, 4*(96), 1-8.

Mather, E., & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology, 105*(3), 232-242.

Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive Science, 36*(7), 1157-1177.

Matthews, D., Lieven, E., & Tomasello, M. (2008). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology, 46*, 749-760.

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3-4).

Merriman, W. E., & Kutlesic, V. (1993). Bilingual and monolingual children's use of two lexical acquisition heuristics. *Applied Psycholinguistics, 14*, 229-249.

Merriman, W. E., & Lipko, A. R. (2008). A dual criterion account of the development of linguistic judgement in early childhood. *Journal of Memory and Language, 58*, 1012-1031.

Perner, J. (1991). *Understanding the representational mind. Learning, development, and conceptual change.* Cambridge, MA, US: The MIT Press.

Perner, J. (2000). RUM, PUM, and the Perspectival Relativity of Sortals. In J. W. Astington (Ed.), *Minds in the Making: Essays in Honour of David R. Olson*. Oxford, UK: Blackwell Publishers Ltd.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development, 17*, 1451-1472.

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition, 97*(1), B13-B23.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(04), 515-526.

Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? [P&W]. *Behavioral and Brain Sciences, 1*(04), 592-593.

Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory Self-Control in Preschool-Children. [Article]. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 30*(2), 131-147.

Savage, S. L., & Au, T. K. (1996). What word learners do when input contradicts the mutual exclusivity assumption. *Child Development, 67*, 3120-3134.

Saylor, M. M., Sabbagh, M. A., & Baldwin, D. A. (2002). Children use whole-part juxtaposition as a pragmatic cue to word meaning. *Developmental Psychology, 38*(6), 993-1003.

Scofield, J., & Behrend, D. A. (2007). Two-year-olds differentially disambiguate novel words and facts. *Journal of Child Language, 34*(04), 875–889.

Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. *British Journal of Developmental Psychology, 23*, 117-141.

Sperber, D. (2001). Metarepresentation. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences (MITECS)* Cambridge, Massachusetts: Bradford Books.

Suanda, S. H., & Namy, L. L. (2013). Young word learners' interpretations of words and symbolic gestures within the context of ambiguous reference. *Child Development, 84*(1), 143-153.

Tare, M., & Gelman, S. A. (2010). Can you say it another way? Cognitive factors in bilingual children's pragamtic language skills. *Journal of Cognition and Development, 11*(2), 137-158.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.

Waxman, S. R., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language, 19*, 153-166.

Wechsler, D. (1967). *Wechsler preschool and primary scale of intelligence*. New York: The Psychological Corporation.

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development, 75*(2), 523-541.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.

Appendix A

Novel words, pragmatic cues and objects for pragmatic cue task, experiment 1

Novel word	Pragmatic cue	Familiar object	Novel object
Hinkel	hungry	Banana	Bottle stopper
Flinder	sleepy	Bed	Bracket
Budit	cold	Jumper	Drill slack adjuster
Jintoff	thirsty	Juice	Hose end connector
Lozee	bored	Book	Bicycle trouser clip

(Disambiguation condition used same materials apart from cues)

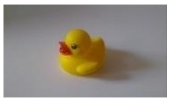
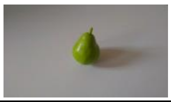
Appendix B

Novel words, pragmatic cues and pictures for pragmatic cue task, experiment 2 & 4

Novel word	Pragmatic cue	Familiar picture	Novel picture
Hinkel	hungry		
Flinder	sleepy		
Budit	cold		
Jintoff	thirsty		
Momtik	sore		

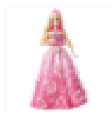
Appendix C

Sets of novel words, familiar and novel items, correction task, experiment 3a

Novel word 1	Novel word 2	Target novel item	Unnamed novel item	Familiar item 1	Familiar item 2	Cue
Puhne	Mieber					hungry
Welne	Nische					thirsty
Sahle	Focher					tired
Grieber	Tachte					cold

Appendix D

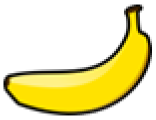
Stimuli and questions for misnaming task, experiment 3a

Picture presented		Subordinate	Basic	Superordinate
	"Is this a …?"	Owl	Dog	Human
	Expected category	Parrot	Bird	Animal
 *	"Is this a …?"	High chair	Table	Vehicle
	Expected category	Rocking chair	Chair	Furniture
 *	"Is this a …?"	Bee	Bird	Plant
	Expected category	Ladybird	Beetle	Insect
	"Is this a …?"	Ken	Child	Person
	Expected category	Barbie	Doll	Toy
	"Is this a …?"	Fir tree	Flower	Insect
	Expected category	Palm	Tree	Plant
 *	"Is this a …?"	Sandal	Boot	Toy
	Expected category	Trainer	Shoe	Clothes
	"Is this a …?"	Boy	Grown-up	Animal
	Expected category	Girl	Child	Human
 *	"Is this a …?"	Fire engine	Aeroplane	Furniture
	Expected category	Police car	Car	Vehicle

* Items are combined nouns in the subordinate category in German (e.g. police car : Polizeiauto)

Appendix E

Stimuli and questions for misnaming task, experiment 3b

Picture presented		Subordinate	Basic	Superordinate
	"Is this a …?"	Dachshund	Cat	Human
	Expected category	Dalmatian	Dog	Animal
	"Is this a …?"	Apple	Vegetable	Drink
	Expected category	Banana	Fruit	Food
	"Is this a …?"	Tulip	Tree	Animal
	Expected category	Rose	Flower	Plant
	"Is this a …?"	Fireman	Woman	Plant
	Expected category	Doctor	Man	Human

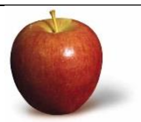
Appendix F

Bilingual children's additional language to English, experiment 4 (n = 18)

Number of children	Second language
4	Spanish
3	Arabic
3	Polish
2	French
2	Hindi
1	Chinese
1	Finnish
1	Chewa
1	Italian

Appendix G

Novel words and pictures for disambiguation task, experiment 4

Novel word	Familiar picture	Novel picture
Lozee		
Kuble		
Pimsulf		
Silder		
Delsy		

Appendix H

Novel words, pragmatic cues and pictures for German pragmatic cue task, experiment 4

Novel word	Pragmatic cue	Familiar picture	Novel picture
Birne	hungry		
Decke	sleepy		
Muetze	cold		
Saft	thirsty		
Salbe	sore		

Appendix I

Novel words and pictures for German disambiguation task, experiment 4

Novel word	Familiar picture	Novel picture
Spiegel		
Brille		
Pilz		
Ente		
Fahrrad		

Appendix J

Bilingual children's additional language to English, experiment 5 (n = 15)

Number of children	Second language
4	Polish
3	Spanish
3	Arabic
2	French
1	Hindi
1	Swazi
1	Russian

Appendix K

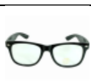
Novel words, pragmatic cues and pictures for pragmatic cue task, experiment 5

Novel word	Pragmatic cue	Familiar picture	Novel picture
Wukti	hungry		
Bamshy	sleepy		
Slintoff	cold		
Gamtik	thirsty		
Dolpho	sore		

Experiment 3a used the same materials, novel words were replaced by more German sounding words: Nohle, Tahne, Doffe, Nehbe.

Appendix L

Novel words and pictures for disambiguation task, experiment 5

Novel word	Familiar picture	Novel picture
Jicky		
Lunti		
Moltin		
Pitshu		
Ranglee		

Experiment 3a used the same materials, novel words were replaced by more German sounding words: Kulde, Fende, Albe, Mehfe, Losse.

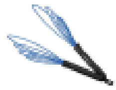
Appendix M

Novel words, pragmatic cues and pictures for German pragmatic cue task, experiment 5

Novel word	Pragmatic cue	Familiar picture	Novel picture
Kuchen	hungry		
Schlafsack	sleepy		
Stiefel	cold		
Kaba	thirsty		
Verband	sore		

Appendix N

Novel words and pictures for German disambiguation task, experiment 5

Novel word	Familiar picture	Novel picture
Tasche		
Muetze		
Stifte		
Loeffel		
Dreirad		

Appendix O

WPPSI- Arithmetic subtest, list of problems

For problems 1 through 4, cards in a booklet were used.	
1. Card 1 - Balls	Here are some balls. Which is the biggest? Point to it.
2. Card 2 - Sticks	Here are some sticks. Which is the longest? Point to it.
3. Card 3 - Stars	Here are some boxes with stars in them. Which box has the most stars? Point to it.
4. Card 4 - Cherries	These bowls have some cherries in them. Which bowls have the same number of cherries? Point to them.
For problems 5 through 8, red wooden blocks were used.	
5. Counting to 2	How many blocks are there?
6. Counting to 4	How many are there? Count them with your finger.
7. Counting to 9	Count these blocks with your finger.
8. (cont. with 9 blocks)	Now give me all of the blocks except four. Leave four of the blocks here.
Problems 9 through 20 were read to the child	
9.	If I cut an apple in half, how many pieces will I have?
10.	Harry had 2 pennies and his daddy gave him 1 more. How many did he have altogether?
11.	Johnny had 3 marbles and lost 1. How many did he have left?
12.	Mary had 5 dolls. She lost 2. How many did she have left?
13.	John had 4 pennies and his mother gave him 2 more. How many did he have altogether?
14.	How many are 2 books and 3 books?
15.	Bob ate 1 piece of candy. Sue ate 2 pieces, and Jack ate 2 pieces. Altogether, how many pieces of candy did they eat?
16.	If 1 apple costs 2 cents, how much will 2 apples cost?
17.	Jane, Alice, and Ann each have 2 crayons. How many crayons do they have altogether?
18.	If 1 orange costs 4 cents, how much will 2 oranges cost?
19.	A boy had 12 newspapers and sold 5. How many did he have left?
20.	James had 8 marbles and he bought 6 more. How many marbles did he have altogether?

Appendix P

WPPSI- Sentences subtest

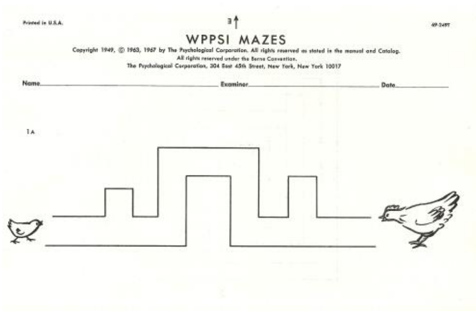
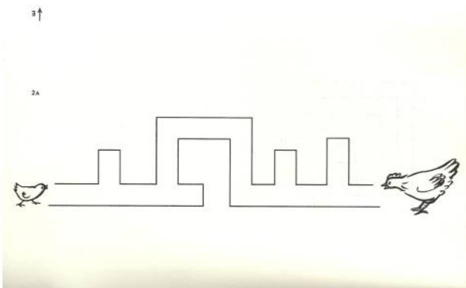
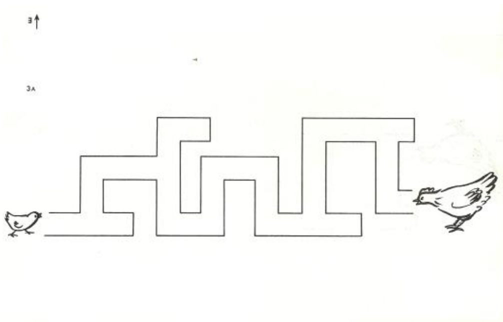
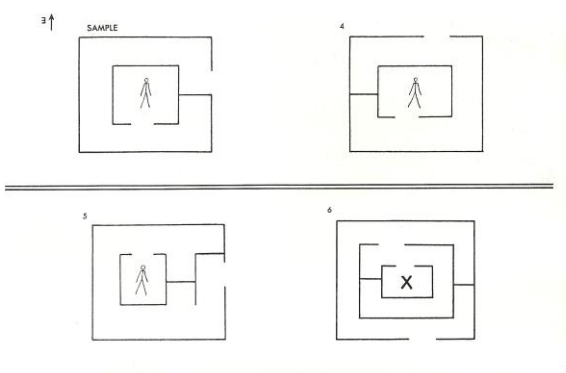
A.	My house.
B.	Cows are big.
C.	We sleep at night.
A, B, and C were only administered if sentence 1 was failed.	
1.	Mary has a red coat.
2.	The bad dog ran after the cat.
3.	Tom found three blue eggs in his birdhouse.
4.	Susie has two dolls and a brown teddy bear.
5.	It is very nice to go to a camp in the summertime.
6.	Peter would like to have new boots and a cowboy suit.
7.	Eating too much candy and ice-cream can give you a stomachache.
8.	The heavy rain which fell last night made many buses late for school.
9.	The price of shoes and winter clothing is not as high as it was last year.
10.	Next Monday our class will visit the zoo. Bring your lunch and be sure to be on time.

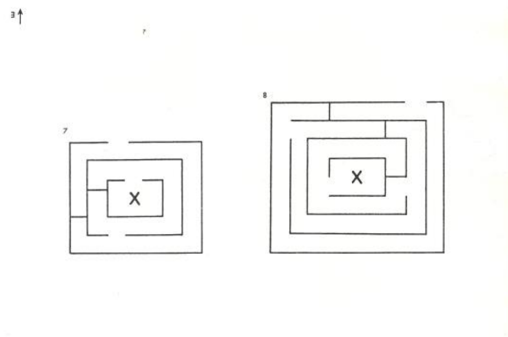
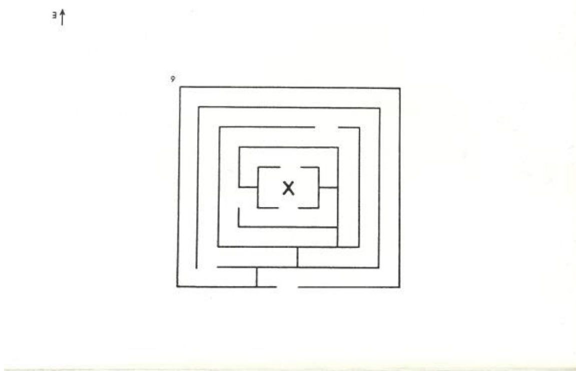
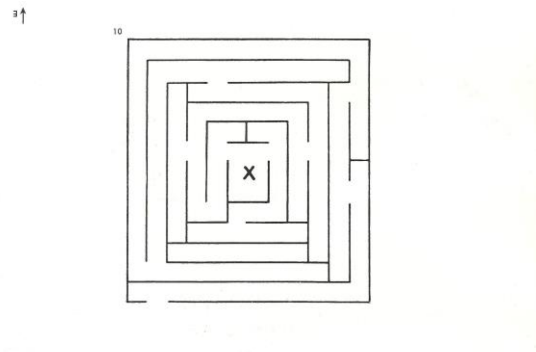
Score table

Sentences	Number of errors				
	0	1	2	3	4+
A - B	1	0	0	0	0
C	2	1	0	0	0
1 – 4	2	1	0	0	0
5 – 6	3	2	1	0	0
7 - 10	4	3	2	1	0

Appendix Q

WPPSI- Maze subtest

	Presentation to the child per sheet
Maze 1 A + B	
Maze 2 A + B	
Maze 3 A+B	
Sample, Maze 4 - 6	

Maze 7 + 8	
Maze 9	
Maze 10	

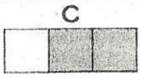
Score table

Maze	Errors allowed	Performance	Score
1 A	*Not counted*	No errors	1
1 B	0	No errors	1
2	0	2A no errors 2 B no errors	2 1
3	1 on Maze 3 A 0 on Maze 3 B	3 A no errors 3 A one error 3 B no errors	2 1 1

4, 5, 6	1	No errors One error	2 1
7, 8	2	No errors One error Two errors	3 2 1
9, 10	3	No errors One error Two errors Three errors	4 3 2 1

WPPSI- Block design subtest

9. BLOCK DESIGN Discontinue: 2 consecutive failures, starting with Design 3

Design		Trial	Time	*	Pass-Fail	Score
1.	C  E	1	30"	D		0 1 2
		2	30"	D		
2.		1	30"	ND		0 1 2
		2	30"	D		
Discontinue: if Designs 1 and 2 are failed						
Age 6 and over: begin here						
3.		1	30"	D		0 1 2
		2	30"	D		
4.		1	30"	D		0 1 2
		2	30"	D		
5.		1	45"	D		0 1 2
		2	45"	D		

Design		Trial	Time	*	Pass-Fail	Score
6.		1	45"	ND		0 1 2
		2	45"	D		
7.		1	60"	ND		0 1 2
		2	60"	D		
8.	See card	1	60"	D		0 1 2
		2	60"	D		
9.	See card	1	75"	ND		0 1 2
		2	75"	D		
10.	See card	1	75"	ND		0 1 2
		2	75"	D		
					Total	

* "D" means examiner demonstrates trial; "ND" means examiner does not demonstrate trial. See Manual.

Card 8

Card 9

Card 10