In a recent paper, Kaminski, Call and Fischer report pioneering research on word-learning in a dog. In this commentary we suggest ways of distinguishing referential word use from mere association. We question whether the dog is reasoning by exclusion and, if so, compare three explanations – learned heuristics, default assumptions, and pragmatic reasoning – as they apply to children and might apply to dogs. Kaminski et al.’s work clearly raises important questions about the origins and basis of word learning and social cognition.

Kaminski, Call and Fischer’s [1] fascinating report of word learning in a border collie named Rico has far-reaching implications. Rico’s 200-word vocabulary is impressive, but of more interest to us are the claims about the speed with which new words are acquired and about his reasoning by exclusion to infer the referent of a novel word. In this commentary, we address two fundamental issues that these findings raise: (1) what counts as word learning as opposed to some form of associative learning? and (2) is Rico really reasoning by exclusion and, if so, what are the implications of this finding?

Word versus associative learning in children and in Rico
Over the course of 8 or 9 years, Rico’s owners taught him 200 labels for small toys by rewarding him with food or play after he retrieved the correct toy. In Kaminski et al.’s exceptionally well-designed test of Rico’s comprehension of these words, the experimenter and the owners were in a room out of view when they issued commands, so their behavior could not influence Rico’s choices. Rico was then required to scan an array of 9 or 10 items to find the one requested. He retrieved the correct item in 37 out of 40 such trials. This is convincing evidence that Rico does understand roughly 200 such commands. But is this word learning per se?

Bloom [2] points out that Rico’s performance does not show that he treats words as referring to objects rather than as holophrastic (undifferentiated, phrase-like) commands associated with object-specific desired actions. Bloom notes that children learn words that refer to objects in different contexts, that they can learn through overhearing, and that they learn words for categories and not just proper names. He suggests these as standards to assess Rico’s performance.

We would add that the most striking demonstrations of young children’s understanding that words are used to refer to objects come from experiments that deliberately pit reference against well-known markers of associative learning, such as temporal contiguity. At 20 months children will map a novel word onto an object they are actively attending to only if they confirm that the speaker is attending to the object as well [3]. Eighteen-month-olds will also reject a mapping between a word and an object when the speaker’s emotional reaction (e.g. ‘Whoops!’) suggests that the coincidence between the two is accidental or mistaken [4]. These infants’ treatment of words clearly goes beyond associative learning based on temporal contiguity. By late infancy, humans have an intentional and referential understanding of words, grounded in a lay ‘theory-of-mind’. We do not know from Kaminski et al.’s results whether Rico shares such an understanding. There is evidence from adult dogs as well as newborn puppies to suggest that domestic dogs have an innate preparedness to read human communicative signals, such as pointing [5]. If Rico can be shown to monitor a speaker’s referential intent, that would be a strong demonstration that he treats words as referential rather than merely associative.

Reasoning by exclusion in children and Rico
By far the most striking finding is Rico’s apparent ability to infer the referent of a novel word in the absence of a direct external cue. In the reasoning-by-exclusion test trials in Kaminski et al.’s study, one novel object was placed among familiar objects and Rico was asked to retrieve an object using a novel label. This design closely parallels those used to demonstrate word-learning by exclusion in young preschoolers [6]. For example, when shown a ball and a whisk and asked for the whisk, children infer that ‘whisk’ refers to the whisk, not the ball which they already know is called ‘ball’. Rico too showed a reliable tendency to select the novel object when hearing the novel label. Moreover, in follow-up comprehension tests Rico, like children [7], appeared to retain the knowledge of words learned by exclusion.

These findings are less convincing than they appear at first sight. Although the follow-up comprehension results look impressive, they could be attributable to the reinforcement Rico received after initially retrieving the novel object (J. Fischer, pers. commun.). Would long-term learning result from reasoning via exclusion, with no positive feedback given to Rico? The presence of feedback does not, of course, explain how Rico arrived at the correct mapping in the first place. On this point, however, another potential problem with Kaminski et al.’s exclusion task is that, unlike studies with children, this study did not include a control for baseline novelty preference. If Rico...
tends to pay more attention to novel objects than familiar ones, this tendency could be solving the identification problem for him. Comparing performance on this task with a no-label control where Rico was asked simply to ‘fetch’ from a set that contains a novel object among familiar ones could rule out a novelty bias. In summary, Rico could have initially retrieved the novel object because it was more salient or interesting but he was then reinforced for his selection and so retained the object–word mapping in the follow-up test. At present, the question of whether Rico achieved word learning by exclusion remains open.

Assuming that the exclusion results are not due to a novelty preference, we turn to the explanations that have been proposed for this inference on the part of children, and examine the extent to which they are applicable to Rico.

**Learned heuristics**

One proposal [8] is that children use a heuristic rule they learn from previous experience with words. Although there is a statistical tendency for words to be mutually exclusive, this tendency by itself might not be strong enough to account for children’s performance. Babies whose productive vocabulary is under 50 words expect that a novel word will refer to a novel object rather than a familiar one [9]. Input to young children contains many violations of mutual exclusivity (e.g. hearing the same object referred to as ‘cat’, ‘kitty’ and ‘pet’, or hearing themselves referred to as ‘Elizabeth’, ‘Lizzy’ and ‘sweetie’). By contrast, Rico had been given explicit training and feedback for 200 words, all of which were object labels, all displaying a perfect one-to-one correspondence with the objects they refer to. Thus, previous experience could play a more central role in Rico’s inference than in children’s.

**Default assumptions**

An alternative explanation for children’s performance involves default assumptions that guide children’s initial guesses about the meaning of novel words. One such default principle is mutual exclusivity, which leads children to expect any given object to have only one correct label unless they are given clear evidence to the contrary [6,10]. Mutual exclusivity is not necessarily specific to language and could well be a result of a general cognitive tendency to seek simple regularities in complex learning tasks, leading learners to exaggerate one-to-one correspondences [11]. This is seen across a wide range of tasks in adults and animals as well as children (e.g. illusory correlations and causal discounting, blocking in classical conditioning). If mutual exclusivity in word learning is an expression of a more general cognitive bias, then it is not so surprising to see it manifested in Rico. Thus, Rico’s word learning may well be influenced by a version of mutual exclusivity.

**Pragmatic reasoning**

Another explanation that has been put forward for children’s learning by exclusion is based on pragmatic reasoning about a speaker’s communicative intent [12,13]. On this view, children expect speakers to use the conventional label for objects and therefore reason, for example, ‘if she meant ball she would have said ‘ball’ so ‘whisk’ must refer to this other thing’. This requires inferring an adult’s communicative intent in the absence of any observable pragmatic cues (e.g. no eye gaze or emotional reactions) and reasoning counterfactually about it (see [9] for a discussion). So far there is no evidence that infants can reason this way so this pragmatic explanation seems unlikely to account for 16-month-olds’ success at learning by exclusion [9]. This skepticism holds for Rico as well. Although dogs might have some rudimentary understanding of humans’ communicative intent [5], using general pragmatic principles alone to disambiguate word reference requires much more than being able to follow a pointing gesture. For neither infants nor dogs is there any independent evidence that they have the capacity to reason from pragmatic principles without the support of overt pragmatic cues. Until such evidence becomes available, there is reason to believe that such inferences are based on mutual exclusivity rather than high-level pragmatic reasoning.

**Conclusion**

There are therefore several potential explanations for Rico’s apparent learning by exclusion, and Rico could still be performing similarly to children for a different reason (e.g. a novelty bias or relying more on learned patterns of word use and less on pragmatic reasoning). It is also possible that, as the authors claim, dogs share some of the cognitive capacities that we use to learn words (e.g. domain-general principles that give rise to mutual exclusivity). Indeed, what makes humans such proficient word-learners might not be one particular species-specific learning principle or a priori competence, but rather our ability to use and coordinate redundant sources of information in word learning (e.g. initial assumptions, generalizations from experience, socio-pragmatic reasoning). Kaminski et al.’s pioneering work has set the stage for a line of inquiry into dogs’ word-learning abilities that can help address important questions about the origins and basis of word learning and social cognition in humans too.

**Acknowledgements**

We thank Mikkel Hansen for his helpful discussions of this work.

**References**

A pluralistic account of word learning

Julia Fischer, Josep Call and Juliane Kaminski

Max-Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

How do language learners come to understand the meaning of words? This is an intriguing question that has given rise to a vast research field [1]. One smaller aspect within this research field is how novel words are mapped onto novel objects. We studied this question in a domestic dog, Rico, who had already learned the names of more than 200 toys. We found that he was able to link a novel word to a novel object on the basis of exclusion [2]. Markman and Abelev have taken a closer look at how Rico's behavior could be accounted for ([3], this issue). Their main concern is that Rico selected the new toy among the familiar ones because of an inherent novelty preference. They also critically discuss the importance of rewarding Rico after fetching the novel toy. Finally, they note that only experiments that pit reference against associative learning (e.g. [4]) can uncover a true understanding of reference. In the following, we will briefly address these concerns.

Firstly, novelty is indeed the discriminating feature of the 'novel object'. It remains an empirical issue whether Rico can link a novel word to an unnamed object that he had been familiarized with before. Markman and Abelev suggested using the command 'fetch' without specifying the object, to see if Rico would show a preference for the novel object. However, we do not think that this test is conclusive because he might still bring the novel object because, unlike all other familiar objects, it did not have a label. In other words, Rico could solve the problem on the basis of recognizing that all objects have labels except one rather than on the basis of preference for novelty. It is also important to note that Rico was able to control his interest in the novel item as he first brought familiar objects upon request (see online video material; [2]). Therefore novelty alone cannot explain our results.

Secondly, we feel the need to clarify that the name of the novel object was never mentioned at the time when Rico was rewarded with food or play. Importantly, within the setting of the fetching game, a lack of reward would in fact constitute a negative reinforcement. That is, without a reward, Rico might still have learnt that he had brought the 'wrong' item.

The final issue concerns the importance of the speaker's intention for word learning. Clearly, children can use the speaker's attention to guide their process of word learning. However, this is not a prerequisite for word learning. In Carey's experiments [5], for example, children selected the 'chromium tray' without being able to use the speaker's attention as a cue. Again, it is an empirical issue whether or not Rico monitors the speaker's attention to disambiguate which object the speaker refers to. In conclusion, we – like Markman and Abelev – propose a pluralistic account of word learning where some mechanisms are shared among children and language-trained animals, and others are uniquely human.

References