

Britta Osthaus · Stephen E. G. Lea · Alan M. Slater

Dogs (*Canis lupus familiaris*) fail to show understanding of means-end connections in a string-pulling task

Received: 12 August 2003 / Revised: 19 July 2004 / Accepted: 19 July 2004 / Published online: 26 August 2004
© Springer-Verlag 2004

Abstract Domestic dogs (*Canis lupus familiaris*) were tested in four experiments for their understanding of means-end connections. In each of the experiments, the dogs attempted to retrieve a food treat that could be seen behind a barrier and which was connected, via string, to a within-reach wooden block. In the experiments, either one or two strings were present, but the treat was attached only to one string. Successful retrieval of the treat required the animals to pull the appropriate string (either by pawing or by grasping the wooden block in their jaws) until the treat emerged from under the barrier. The results showed that the dogs were successful if the treat was in a perpendicular line to the barrier, i.e. straight ahead, but not when the string was at an angle: in the latter condition, the typical response was a proximity error in that the dogs pawed or mouthed at a location closest in line to the treat. When two strings that crossed were present, the dogs tended to pull on the wrong string. The combined results from the experiments show that, although dogs can learn to pull on a string to obtain food, they do not spontaneously understand means-end connections involving strings.

Keywords Dog · Means-end task · Problem-solving · String-pulling

Introduction

Understanding means-end relationships is a key step in cognitive development, whether that development is phylogenetic or ontogenetic. Without such an understanding, it will be impossible for an organism to transform an intention into a plan, and arguably impossible to form an intention at all; discussions of intentions and

planning in mental philosophy rightly take such understanding for granted (e.g. Bratman 1981).

In human cognitive development, a significant transition occurs at around the age of 8 months, when infants move beyond a reliance on what Piaget (1953) called “circular reactions”—in effect, operant conditioning—to an understanding of means-end relationships (Willatts 1984a; Brown 1990; Munakata et al. 1997). Comparative psychologists have long been interested in whether other species of animals make a similar transition. The question is especially important within the approach of using the Piagetian analysis of human cognitive development to structure a comparative understanding of animal cognition (e.g. Doré and Dumas 1987; Pepperberg 2002), but in any assessment of the nature of cognition in a species, its performance in means-end tasks is significant. Several early studies (e.g. Shepherd 1915; Köhler 1925, chapter II) sought to compare different species on their performance in string-pulling tasks, and in particular compared the performance of dogs with that of primates.

In order to study the understanding of means-ends relationships, it is necessary to use problem-solving tasks in which a solution to the problem can in principle be perceived directly, without trial and error or previous experience of similar tasks. Means-end understanding is most clearly demonstrated if the subject shows an “insightful” solution to the problem on the first trial, since correct performance at the end of a period of training may well represent the effects of operant conditioning. The usual way of studying means-end understanding in animals, introduced by Köhler 1925, is by offering a possible physical connection to an out-of-reach object of desire, for example a string that is attached to a piece of food. The food itself is out of reach of the animal, but the near end of the string is accessible. If the animal understands the physical properties of the string it uses it as a means to an end, i.e. pulls the food into reach with the string. Some recent experiments (e.g. Hauser et al. 1999, 2002) have used a variant of this task, derived from the human developmental literature (e.g. Willatts 1984b), in which the object is placed on a cloth which the subject is

B. Osthaus (✉) · S. E. G. Lea · A. M. Slater
School of Psychology, University of Exeter,
Exeter, EX4 4QG, UK
e-mail: b.osthaus@exeter.ac.uk
Tel.: +44-1392-264626
Fax: +44-1392-264623

able to pull towards it. Other tasks, such as tool use, may require an understanding of means-end relationships, but they generally also require other cognitive capacities as well, so they are less appropriate as tests of means-end understanding as such.

String-pulling behaviour has been established successfully with a number of different species. Primates acquire the task readily, as has been shown in a variety of species, including chimpanzees (e.g. Köhler 1925; Spinozzi and Poti 1993) and both old world and new world monkeys and lemurs (Bierens de Haan 1930; Klüver 1933; Hauser et al. 1999, 2002). However, string-pulling has also been demonstrated in cats (Adams 1929), rats (see review by Tolman 1937), and a variety of birds, including corvids (Heinrich 1995), psittacids (Ducker and Rensch 1977; Funk 2002; Pepperberg 2004) and songbirds, though in some songbird species only a minority of individuals are successful (Vince 1956, 1958, 1961). However, although dogs were among the first animal subjects to be tested in formal means-end tasks, there has been little subsequent literature on such tasks with them. Because of their continued importance as working animals, the assessment of cognition in dogs is of interest for practical reasons, but it is also of theoretical interest. Dogs are widely believed to be of high intelligence (Eddy et al. 1993; Nakajima et al. 2002); their ancestors, like early humans, were social hunters, and both the form and the extent of intelligence in dogs is therefore relevant to the “hunting hypothesis” (e.g. Washburn and Lancaster 1968) of the emergence of exceptional intelligence in humans. On the other hand, unlike primates or even some birds, dogs are not well equipped anatomically for manipulating objects, and object manipulation is of no obvious ecological relevance to them, and on these grounds string-pulling might be a task at which they would not do well.

The results that are available on dogs’ string-pulling behaviour present a confused picture. Shepherd (1915) reported dogs’ and cats’ performance as much inferior to that of rhesus monkeys. Köhler (1925) tested one dog, which failed to pull a food-filled basket towards it by means of a string. Fischel (1933) trained two dogs to pull a 25-cm string to reach some meat fixed at the end. On the second day of the experiment the dogs were confronted with two parallel strings, one attached to a biscuit (less preferred food) and one to meat (the preferred food). The dogs did not immediately pull in the string that was attached to the meat, although after a few trials (10 for 1 dog and 16 for the other) they learned to do so reliably. Fischel also reports a subsequent experiment in which a dog sometimes pulled an immediately accessible string that was attached to a non-preferred food item (a biscuit) instead of making a detour that led to a preferred item (a piece of meat). Sarris (1937) found that two out of seven dogs tested were able to learn to pull a piece of meat towards them on a string. Grzimek (1942) tested one dog and found that she could choose a baited over a non-baited plank in two out of three trials. Because of the small sample sizes used in these early experiments, it is impossible to be sure whether they give evidence of

spontaneous understanding of means-end connection, though taken together, they suggest that perhaps dogs do not show such understanding.

All these early studies suffer from small sample sizes. Scott and Fuller (1965) conducted related tests with a much larger sample, using puppies of five different breeds. In their experiments, they tested the ability of the animals to pull a food-filled dish out of a wooden box with the help of an attached rope. However, in their experiment, the puppies were unable to see the set-up of the food and the connecting string (which were hidden in the box), and therefore their performance had nothing to do with an understanding of means-end connections, being rather an example of simple instrumental learning. The same procedural issue arises in a study by Frank and Frank (1985), where they compared the manipulation abilities of wolves (*Canis lupus lupus*) and dogs (*C. l. familiaris*) of comparable physique.

None of the early experiments, in which the animals could actually see the set-up, had adequate controls for the spontaneous behaviour of dogs. It is not in dispute that dogs have the ability to manipulate objects with their front paws, and pawing is a frequent response to novel objects. Pulling food into their reach by simply pawing at the string, or in its vicinity, is not as such an indicator of understanding the cause-effect relationships: it might simply be an innate behaviour pattern, which would be rapidly strengthened by operant conditioning or other associative learning mechanisms if it successfully procured food. Tolman (1937) and Vince (1961) argued that the elicitation and subsequent strengthening of such behaviours is the explanation of rats’ and songbirds’ success in string-pulling tasks. In the terms of Piaget (1953), the ability to learn the behaviour of pulling the string emerges at sensori-motor Stage III, but the understanding of the relation between the means and the end not until Stage IV; in these terms, there is little doubt that dogs attain Stage III, but it remains quite unproven whether or not they attain Stage IV. The research question is therefore not whether dogs are able to pull in the string, but whether they understand the means-end properties of the string.

In previous research on string-pulling behaviour two basic designs have been used to test the understanding of means-end tasks: first, the ability to pull a string that is not at a 90° angle to the barrier, i.e. straight ahead (as tested in a lemur and a macaque by Fischel 1930); and second, the ability to choose a baited string over a non-baited one. The experiments described here tested the cognitive abilities of dogs in more complex string-pulling tasks, involving one string laid out diagonally (experiment 1), with two parallel strings perpendicular to the barrier (experiment 2) and with two strings at an acute angle, either parallel (experiment 3) or crossed (experiment 4). Figure 1 summarises the different conditions used.

Expt	Situation	Conditions	dogs
1	One string	a) Short, perpendicular	16
		b) Long, perpendicular	
		c) Long, acute angle (diagonal)	
2	Two strings, perpendicular	a) Near (string separation 10 or 20cm)	24
		a) Far (string separation 50cm)	
3	Two strings, parallel, acute angle	a) Overlap (bait is opposite the accessible end of the non-baited string)	12
		b) Exterior (no overlap between the baited string and the accessible end of the non-baited string)	
4	Two strings, crossed		12

Fig. 1 Summary of the experimental situations and conditions used on dogs (*Canis lupus familiaris*), and the numbers of subjects in each. In experiments 1–3, all subjects experienced all conditions

Methods

Subjects

Dog owners were recruited to allow their pets to take part in the study either by contacting the general public via local media, or through local dog clubs. The dogs were adults (at least 1 year old) of a variety of ages and breeds, and both sexes. The characteristics of all dogs that took part in the experiments were recorded, and steps were taken to ensure that there were no major age, breed or sex differences between conditions within each experiment. The subjects for experiment 3 were all recruited at a dog club and they were of a more uniform, younger age. However, there was no evidence that age, breed or sex had any impact on the results, and they will not be discussed further.

Apparatus

In each experiment, a box with transparent plastic walls and a wire-mesh lid was used (see Fig. 2). The base of the apparatus consisted of a wooden board, 60×80 cm, with a 0.5-cm layer of white plastic on top, to provide a smooth surface and a good contrast for the red-coloured or blue-

coloured strings. The enclosed area on top of the base was 60×60 cm, and the sides were 20 cm high and made of Perspex. Where the base protruded from the box, there was a small gap (2 cm) between the transparent wall and the base, enabling the string to be threaded through. Because of the rather enthusiastic approach of several dogs, this front sheet of plastic was supported by small wooden pegs, so that it could not be bent and broken, or trap the paw of a dog. There were four wooden pegs at regular distances across the front panel connecting the bottom of the transparent front sheet with the base. The wire-mesh lid was fixed on hinges at the back wall and secured with latches at the front. This enabled the experimenter to access the enclosed area of the box. The strings used in these experiments were bright red or blue plastic, with an approximate diameter of 0.7 cm. Each string had longitudinal slits every 2.5 cm. The food rewards were strips of chewy dog treats, approximately 3×1.5×0.3 cm. These were inserted into one of the longitudinal slits to form a sort of T-shape with the string. At the other end of the string, outside the box, a wooden cube (5 cm³) was fixed to prevent the string from sliding into the box and out of reach of the dogs. The white base was marked inside the box with black lines at a distance of 10 cm, parallel to the opening. These provided the guidelines for the positioning of the food rewards.

Procedure

The testing took place indoors, either in a suitable room in the house of the owner or the dog club or in a testing room in the School of Psychology, University of Exeter. The dogs were first allowed to make themselves familiar with the surroundings (if necessary), the experimenter and with the apparatus. The owner of each animal was present whilst the dog was taking part in the experiment, but was located behind the dog as it faced the apparatus, and took

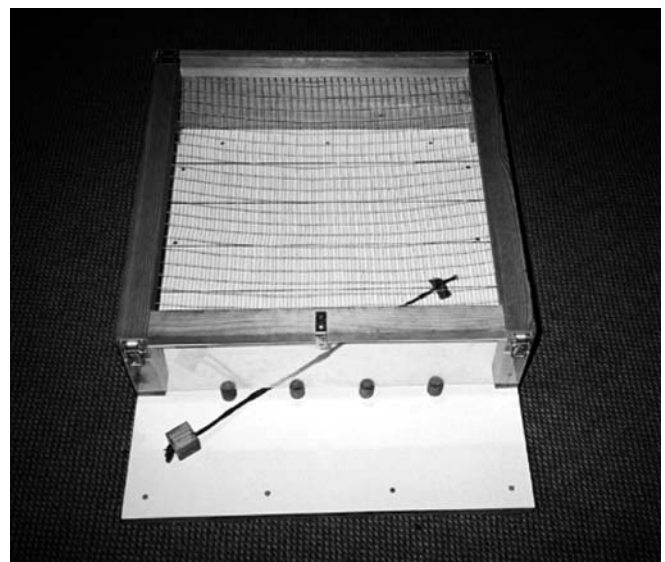


Fig. 2 Apparatus, set up for the diagonal condition of experiment 1

no active part in the procedure. Owners were told at the time of recruitment that the experiment involved problem-solving, but not that it involved string-pulling. At the beginning of the experiment, they were told that there were no right or wrong answers in the tests. The dogs were not restrained during tests.

All the experiments required a training phase, in which the dogs were allowed to familiarise themselves with the apparatus, the string, and the treats. In all the experiments, minor variations in the training procedure between groups were introduced, to ensure that details of training were not influencing the results. Since these details did not in fact have any effect, they are recorded but not discussed further below. In all experiments, the entire procedure, including both training and testing, took place in a single session.

Tests consisted of a number of trials, separated by intertrial intervals of approximately 10 s during which the apparatus was arranged for the next trial while the dog waited at a distance of 2–3 m, either with its owner or in a “down” position.

Experiment 1

In experiment 1, dogs were tested for the ability to use single strings to pull in food. Three conditions were used, in increasing order of their expected difficulty for the dogs: a single short string laid out perpendicularly, a long string laid out perpendicularly, and a long string laid out diagonally. Performance was measured by the dogs’ latency to pull the treat right out of the apparatus so that they could eat it.

It was predicted that:

- (a) In each condition (short perpendicular, long perpendicular and diagonal strings), performance would improve across trials, since previous literature suggests that dogs become more skilful at string-pulling tasks with experience
- (b) Between conditions, performance on the first trial would improve, reflecting transfer of training between conditions
- (c) Performance on the first trial of the later conditions would be worse than on the final trial of the preceding condition if and only if it was to any extent a new task for the dogs
- (d) If a performance asymptote was reached, it would be at a lower level (higher latency) for the later, more complex conditions

Methods

Sixteen dogs were tested (7 female, 9 male, age: mean \pm SD = 5.38 \pm 3.54, range 1–12). The dogs were first tested with a short string (20 cm), laid out centrally in the box at a right angle to the front panel with the opening. The food was placed on the first marking line in the box, 10 cm distance from the opening, so that 10 cm of the string

protruded from the box. While the food was put into position, the owner held the dog, and no attempt was made to prevent the dog seeing what the experimenter was doing. The dog was then released and allowed to explore the situation and to try to reach the food. The short set-up was presented 10 times to each dog and the time between approaching the apparatus and successfully obtaining the treat was measured. Approaching the apparatus was defined as the first contact with the box (either paw or nose) or when the animal put its nose above the protruding area of the box; obtaining the treat was defined as pulling the treat fully out of the apparatus. Next, each dog was tested with a long string (60 cm). Again the string was laid out at a right angle to the front panel, between the two central wooden pegs. Each dog was tested 10 times with this set-up.

Finally, the dogs were tested with a diagonal set-up. The string was the same length as the one used in the previous test (60 cm). It was presented to each dog 10 times, alternating between left–right and right–left set-up trials. Again, the time between approach and successful retrieval was measured. Also, it was recorded whether a dog pawed at the side of the box where the food, and not the end of the string, was to be found. This was described as a proximity error.

All trials of experiment 1 were filmed with a video camera located either at the side or at the back of the box. The recordings were transcribed to digital form and analysed frame-by-frame on a computer. The times of events were established from the frame numbers and the timestamps on the video tape. The video recorded 40 frames per second, so the maximum accuracy of the coding was 0.025 s. In practice, the behaviours recorded, and their times, could be assigned unambiguously within two frames, so durations of behaviour were timed to an accuracy of better than 0.1 s. All coding was carried out by the first author.

Results

All dogs learned to get the food out of the box, although the methods varied. In the short string set-up some dogs were successful by trying to “lick out” the treat. Because the treat was only 10 cm away from the barrier some dogs were able to reach the treat directly with their tongue without touching the string. With the longer string, they changed to using their paws. With the long string only, two dogs applied a technique in which they took the wooden cube at the end of the string in their mouth and pulled until the food was in reach. All other dogs retrieved the food by pawing at the string. In the diagonal condition, most of the dogs were seen to make a characteristic error in which they pawed at the ground at the point on the barrier nearest to the treat, rather than at the end of the string; this was described as a proximity error.

Figure 3 shows the median time taken to retrieve the treat on each trial in each of the three conditions. Within each condition (short and long perpendicular string, and

diagonal string) a clear learning curve was found, showing reduction in the time until retrieval, with an asymptote being reached after about five trials. For all three conditions, the correlation between mean retrieval time and trial number was negative and significant (short string, Spearman $\rho=-0.93$; long perpendicular string: $\rho=-0.88$; long diagonal string, $\rho=-0.70$; $P<0.01$, $P<0.01$ and $P<0.05$, respectively, two-tailed in all cases).

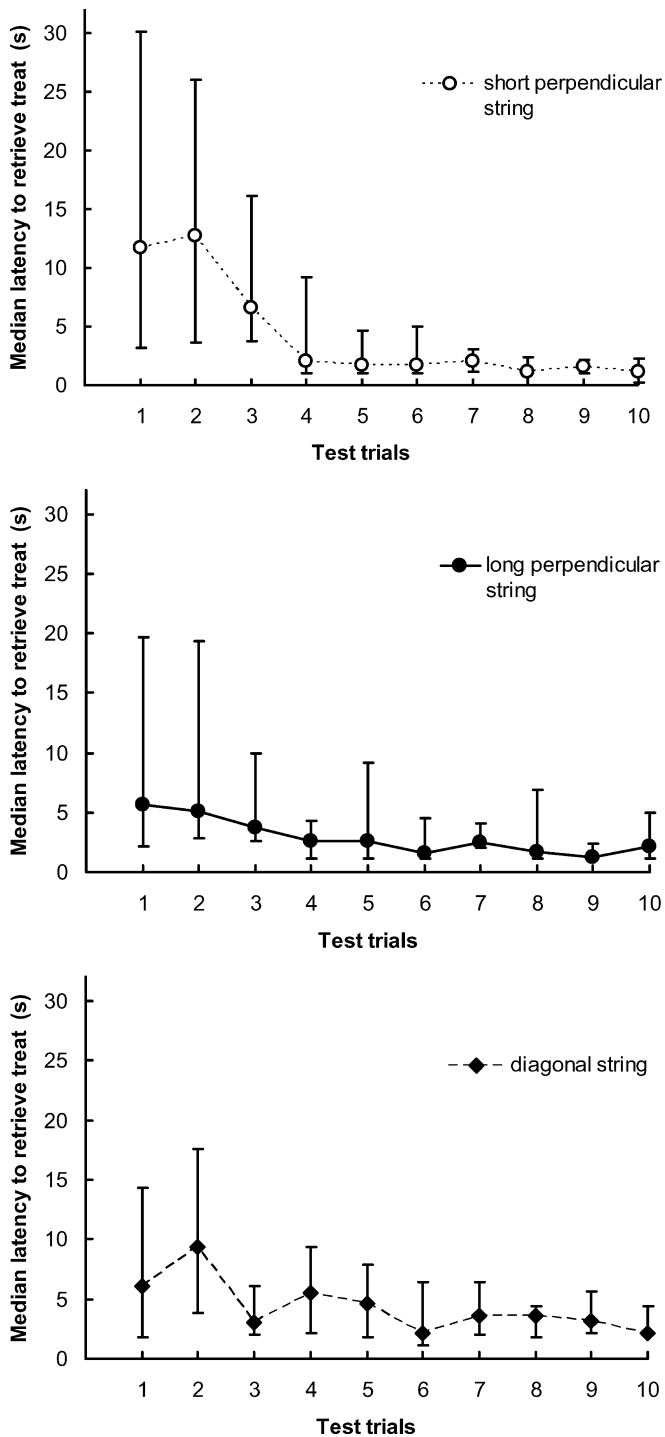


Fig. 3 Median and interquartile ranges of times to retrieve the treat in each test trial under each condition of experiment 1

To examine whether there was transfer between the three tasks, retrieval times from the first trials of the three conditions were tested to see whether they showed a decreasing trend across conditions. A significant trend was found (Page 1963 trend test: $L=204$, $P<0.05$). It is clear from Fig. 3 that the difference is largely due to the first condition, which led to higher initial retrieval trials than the other two conditions. The retrieval times in the first trials of the long perpendicular and diagonal conditions were compared with the final trials of the short perpendicular and long perpendicular conditions, respectively; both increases were significant (one-tailed Wilcoxon-tests, $T=8$, $P=0.0004$; $T=31$, $P=0.029$), confirming that each change in conditions led to slower retrieval.

To examine whether there were differences in the final performance under the three conditions, the median latency for each dog in the last five trials of each condition was calculated, and these medians were compared. Median asymptotic retrieval times in the three conditions were 1.2, 2.1 and 3.1 s, respectively. The increasing trend was significant (Page test: $L=204$, $P<0.05$).

As noted above, with the diagonal string, a characteristic proximity error appeared: several of the dogs pawed near the food and not near the end of the string. Whenever a dog showed this behaviour of digging or pawing close to food out of reach it was recorded, although as usual the trial continued until the dog retrieved the treat. Thirteen out of the 16 dogs showed this error at least once during their ten trials with the diagonal string (mean \pm SD = 3.75 ± 2.67 , range 0–8). There was a significant negative relationship between trial number and number of proximity errors (Spearman's $\rho=-0.68$, $P<0.05$).

Discussion

All of the 16 dogs tested were able to retrieve the food by means of the string. The animals showed significant improvement in each of the three conditions, reaching asymptote after five trials. The change from the short string to the long string did not cause an increase in the time for retrieval up to the level of the very first trials with the short string. It can therefore be concluded that the change from a short to a long string did not constitute an entirely new and different problem to the animals. Their previous problem-solving strategy of pawing the string provided an equally successful approach for this second task.

On changing to the diagonal set-up, the initial time to retrieval increased to the same level as during the first trials with the long string. Again this means that the set-up was a different, but not entirely unknown, problem to the animals. However, most of the dogs initially showed what we have called the proximity error, i.e. pawing near the food, not at the accessible end of the string. This error could not arise with the perpendicular strings, since the point nearest the food was at the end of the string. Over time the frequency of this behaviour decreased signifi-

cantly and retrieval times became fast, though not as fast as with the perpendicular string. The proximity error can be regarded as a form of “goal-tracking” (Boakes 1977), which often obstructs successful performance of instrumental tasks. Similar errors have been reported in related tasks, e.g. by Adams (1929), for cats in a two-string task, and by Santos et al. (1999) in the support task in tamarins; the gravity error shown by tamarins (Hood et al. 1999) and dogs (Osthaus et al. 2003) is essentially the same phenomenon in the vertical rather than the horizontal plane.

Although the dogs became skilful at pulling the string under all conditions, the fact that changing to the diagonal string returned their performance to its initial levels with a long perpendicular string suggests that they did not understand the means-end connection between the string and the treat. The motor responses required to pull in the string did not differ between perpendicular and diagonal strings, so the increase in retrieval latencies must mean that the dogs initially did not know what they had to do in the diagonal condition. Initial performance is critical for demonstrating means-end understanding, since even in the absence of such understanding we would expect performance to improve over trials as a result of reinforcement. Further experiments were conducted to give a more sensitive test of means-end understanding.

Experiment 2

To test whether the dogs were simply learning to paw at the location where the string protruded from the barrier, experiments 2, 3 and 4 employed two strings, one with food attached to the end and one without. This set-up tested whether the dogs were able to distinguish between baited and non-baited strings. Simply pawing at any accessible string would provide a random success rate, whereas, if the animals understand the connection via the string, they should be able to pull in only the baited string. In experiment 2, two parallel, perpendicular strings were used, and the distance between the two parallel strings was varied in order to test whether proximity of the strings made discrimination of the correct string more difficult.

Methods

Twenty-four experimentally naïve dogs were tested (11 female, 13 male, age: mean \pm SD = 5.25 \pm 3.33 years, range 1–12). First, all dogs were made familiar with the box and the task by presenting to them either one short string (20 cm) with food in different locations (8 dogs) or two short strings simultaneously, one baited, one not (16 dogs). Each dog was trained for 20 trials with the short string(s), except that when two strings were used, training was stopped earlier if a dog pulled the correct string on five consecutive correct trials. After training was complete the dogs proceeded immediately to the test phase.

Both conditions involved two long strings (60 cm), one baited, one not, placed parallel and at a 90° angle with the barrier. In the “far” condition, the distance between the strings was 50 cm. In the “near” condition, the distance between the strings was 10 cm on half the trials for each dog and 20 cm on the remaining trials, in pseudo-random sequence. The strings were always equidistant from the centre of the barrier. Each dog had 20 trials in the “far” condition and 20 in the “near” condition, with the position of the treat changing pseudo-randomly between each trial to avoid location learning. The eight dogs initially trained with one string all had their first test trial in the “far” condition; 8 of the 16 dogs initially trained with two strings were tested with the “far” condition first, and the remaining eight with the “near” condition first.

A correct response was recorded if the baited string was the first one that the dog pulled out completely: touches and ineffective pulls were not scored, though they were noted. Regardless of errors, each trial continued until the dog pulled out the baited string. In both conditions the strings were frequently interchanged between trials, to avoid preference learning for one of the two strings.

Results

As in experiment 1, all the dogs learned to retrieve the treats. However, they were not reliably successful at pulling the baited string first, and their success at doing so varied between conditions. Figure 4 shows the median proportion of trials in which the correct treat was pulled out first on test trials in both “near” and “far” condition. In the “near” condition, 22 out of the 24 dogs pulled the correct string on more than half the trials and only 2 did so on fewer than half the trials. In the “far” condition, these numbers were 16 and 3, respectively (the remaining dogs pulled the two strings equally often). Both these proportions are significantly different from 50% (two-tailed binomial tests, $P=0.00004$ and $P=0.0044$, respectively). Performance was significantly better in the “near” than in the “far” condition (Wilcoxon $T=45$, $P=0.0016$, two-tailed). The order in which the dogs experienced the two conditions did not affect these trends.

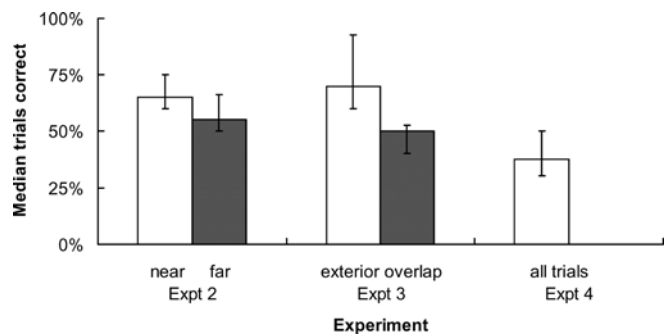


Fig. 4 Median and interquartile ranges of the percentage of test trials on which the baited string was the first to be pulled completely out of the apparatus, in each condition of experiments 2–4

Neither the type of initial training (1 or 2 short strings) nor the order of testing (first “far” then “near” or vice versa) had any influence on the number of correct responses in each condition (Mann–Whitney test: training, $U_{8,16}=58.5$, $P>0.05$; order: $U_{8,16}=43$, $P>0.05$).

Discussion

These results indicate that dogs are able to choose a baited string over a non-baited one if these two strings are close to each other (10 cm or 20 cm distance). The performance of the animals deteriorated when the strings were presented at a distance of 50 cm. This difference is surprising given the proximity errors observed in experiment 1. However, none of the findings suggest that the dogs understand the means-end properties of the strings. In the “far” condition, the dogs apparently pawed at either string at random, and as experiment 1 with the diagonal string demonstrated, pawing to bring food into reach does not mean that the animals grasped the principle of the connection via the string. In the “near” set-up, the dogs showed a better performance but again it cannot be concluded that they actually understood the physical relationship between the string and the treat. It seems that close grouping of the two strings and treats may have made it easier for the dogs to see that a choice had to be made. But given the physical set-up of experiment 2, where the accessible end of the correct string was positioned on the same side of the configuration as the food, the dogs could have chosen to pull the string on the same side of the box where they perceived the treat to be. Experiment 3 was designed to test whether dogs could learn to choose the string actually connected to the food.

Experiment 3

In this experiment, dogs were tested with a set-up which combined the tasks of the two previous experiments: they had to choose from two strings, one with food attached, one without, and both were laid in parallel at an acute angle to the barrier of the box. This resulted in a set-up where the baited string could “overlap” with the accessible end of the non-baited string. If both strings were laid out with a tilt to the right, and the food attached to the left one of the strings, then this was called “overlap” as the end of the non-baited string was in line with the food. With the food attached to the right string, no overlap between the baited string and the accessible end of the non-baited string would occur, and this is referred to as the “exterior” condition (see Figure 1). If dogs were able to learn the means-end properties of the strings they would always pull out the string attached to the food. However, if the animals applied the problem-solving strategy of pulling the string closest to food they would choose the wrong string in the overlap condition while performing well when the food was attached to the exterior string.

Methods

Twelve dogs, who had not taken part in any previous experiments, were tested in this experiment (7 female, 5 male, age: mean \pm SD = 1.83 \pm 1.11 years, range 1–4). The testing took place in a quiet room at a dog training club. The strings used were 30 cm long and the same material as in the previous experiments, but blue. The shorter training strings were the same as those used in experiment 2. Half of the dogs were trained with one 30 cm diagonal string, with the set-up changing pseudo-randomly between left and right tilt, but with the string always placed in the centre gap. The other six dogs were trained with two short strings (20 cm), also laid out diagonally, with the tilt and the position of the food changing randomly between each trial. However, because of the short length of the strings they did not overlap and the dogs were merely trained to choose the baited string out of the two available ones. Each dog had ten trials with its training set-up.

The testing consisted of 20 trials with two parallel strings, arranged at an approximately 45° angle with the barrier. The accessible end of one string was always at the middle of the barrier, and the accessible end of the other string was 20 cm from it, in the direction towards which the first string tilted. The strings could either be tilted to the left or to the right, and the treat could either be positioned on the string that overlapped with the other one, or on the exterior string. In the overlap position, the end of the string without the food was actually closer and in a straight line to the treat, so that application of the problem-solving strategy of pulling the string closest to the food would lead to an incorrect response. Each dog was presented with five trials of each possible combination of the string set-up in a pseudo-randomised order: strings tilted left, food on the left string (exterior); tilted left, food on the right string (overlap) and the same for the two strings tilted to the right.

The position of the treat and the direction of the tilt of the strings varied according to a pseudo-random schedule: the baited string was never in the same position (overlap/exterior) more than twice in a row, and the tilt of the strings followed the same rule. The same trial sequence was used for all dogs. Each trial continued until the dog retrieved the treat. As in experiment 2, the trial was scored as correct if the baited string was the first one that was pulled out completely, though touches and incomplete pulls at the strings were noted.

Results

As in previous experiments, all dogs learned to retrieve the treats, and they showed some success in choosing which string to pull. Ten out of the 12 dogs scored significantly above the chance level of ten trials correct, and this proportion is significantly more than 50% (two-tailed binomial test, $P=0.032$). There was no significant correlation between trial number and number of dogs correct

(Spearman's $\rho = -0.11$, $P = 0.64$ one-tailed). This suggests that no learning took place over the 20 test trials.

Figure 4 shows the results broken down into “overlap” and “exterior” trials. Out of the 12 dogs tested 9 did better in the exterior condition than in the overlap one, 2 dogs performed equally well in both conditions and 1 dog did better when the strings overlapped. The performance of the dogs was significantly better in the exterior condition, when the treat did not overlap with the non-baited string (Wilcoxon $T = 3$, $P = 0.002$, two-tailed). Out of the ten trials in the overlap condition, 5 dogs chose the correct string on more than five trials and 5 dogs chose it on fewer than five; this proportion is not significantly different from chance (two-tailed binomial test, $P = 0.73$). When the treat was fixed to the exterior string, 11 out of the 12 dogs chose that string on more than half the trials, significantly more than 50% (two-tailed binomial test, $P = 0.01$).

For all dogs, the strings were equally often tilted to the left and to the right. The difference in number of correct responses between the two orientations of the strings was not significant (Wilcoxon $T = 24$, $P = 0.86$, two-tailed; left tilt: mean = 6.00, right tilt: mean = 5.83). Nor was there any significant difference in the performance of the dogs that were initially trained with one string (mean trials correct = 12.00) and those that encountered two strings during their training (mean = 11.67) (Mann–Whitney $U_{6,6} = 17.50$, $P = 0.93$).

Discussion

The dogs in this experiment were all relatively young, whereas the earlier experiments used dogs of more mixed ages. However, in the previous experiments age had no discernible effect on behaviour, so this should not be a confounding factor.

This experiment was designed so that the dogs would make errors if they used their default problem-solving strategy when encountering food out of their reach, i.e. by simply pulling a string that was close to the food. Since the dogs did make such errors, the results give no evidence that they understood the means-end properties of strings. When the food was attached to the outer string their performance level was comparable to that found in experiment 2, as would be expected if they were using the same strategy (pull the string that is nearest to food) in both cases. Adams (1929) reports very similar errors by cats in a two-string task when the string that had to be pulled was not the one whose accessible end was closest to the food.

The difficulty of the overlap set-up seems to result from the same kind of proximity error as was observed in experiment 1. This suggests that dogs used a hierarchy of problem-solving approaches in the present means-end situation. First the dogs tended to dig in a position closest to the food. If this was not successful, for example if the string was presented at an angle as in experiment 1 and in the present experiment, they were able to learn to paw a string that protruded from the box. If two strings were

accessible, and only one had food attached to it, the animals could learn to bring these two strategies together, and paw the string that was closest to the food, i.e. on the same side of the apparatus as they perceived the food. This method was successful in experiment 2, and in experiment 3 enabled the dogs to perform significantly above chance level when the food was positioned on the exterior string. However in the overlap condition it would not solve the task, and in fact should lead to incorrect responding. The fact that the performance of the dogs was not significantly below chance level in the overlap set-up is probably due to the same processes that influenced the results of experiment 2. There, the dogs chose the baited string significantly more often than the one without food, but they never reached more than around 70% correct responses.

Because experiment 3 involved both overlap and exterior conditions, the animals were frequently rewarded for applying the strategy of pulling the string closest to food. This may have hindered their development and application of a new problem-solving approach that was required for success in the overlap condition, namely taking into account the means-end connection provided by the string.

The next experiment was designed in such a way as never to reward the pulling of the string closest to the food, to see whether dogs were able to learn to utilise means-end connections if the conflicting strategy was not encouraged.

Experiment 4

The results of experiment 3 indicated that, at least within the 20 test trials, the dogs persisted in pulling the string closest to the food and were not able to learn a more successful strategy for retrieving the treat. Although suboptimal, this strategy was frequently reinforced, since whenever the dogs were confronted with the exterior condition, pulling the nearest string led to the correct choice. Experiment 4 tested whether dogs are able to learn a different problem-solving strategy when the suboptimal strategy was never reinforced. Dogs were again tested with two strings at an acute angle with the barrier, but this time the strings were crossed. With this set-up, choosing the string closest to the treat would never be successful.

Methods

Twelve dogs, which had not taken part in any previous string-pulling experiments, were tested in this experiment (4 female, 8 male, age: mean \pm SD = 4.00 \pm 2.45 years). The testing again took place in the dog owners' homes or in a quiet room at a dog training club. The same apparatus and strings as in experiment 3 were used. To avoid the crossed strings getting tangled a Perspex bridge was constructed and positioned at the cross-point of the strings

in the box. One string went on top of the bridge, the other one underneath it.

The training procedure was the same as for experiment 3, with half the dogs trained with two short strings, laid out diagonally but not overlapping, and the other half trained with one string the same length as in the testing phase of the experiment (30 cm). All 12 dogs were then tested with the crossed set-up only for 20 trials. The position of the treat varied according to a pseudo-random schedule: the treat was never in the same position more than 3 times in a row. As in experiments 2 and 3, the trial was scored as correct if the baited string was the first one pulled out completely.

Results

Once again all dogs succeeded in retrieving the treats. However, they did not do so by the efficient means of pulling first at the string connected to the treat. Figure 4 shows the median percentage of trials on which the dogs pulled the correct string out first. Out of the 12 dogs tested 8 scored below the chance level of ten correct trials; the other 4 were correct on 10/20 trials. The proportion scoring below 10/20 correct was significantly greater than chance ($P=0.008$, two-tailed binomial test). There was no significant correlation between trial number and number of dogs correct (Spearman's $\rho=-0.29$, $P=0.24$). Thus no learning was apparent over the 20 test trials; if anything, the dogs' performance deteriorated over trials.

Other factors had no significant effect on the dogs' performance. Each dog had 8 trials in which the food was in the same position as in the previous trial, and 12 trials where it was in a different position. There was no difference in the performance of the dogs between those trials where the food was in the same position as in the previous trial (mean percentage correct \pm SD: 38.5 \pm 20.3%) and those where the food position had changed (mean \pm SD: 36.8 \pm 10.3%) ($T=30$, $P=0.73$, two-tailed). There was also no significant difference in the performance of the dogs between those who were trained with one string (mean = 7.17) and those who encountered two strings during their training (mean = 7.83) (Mann-Whitney $U_{6,6}=15.50$, $P=0.68$). There was some suggestion of a spatial bias, since the number of correct responses on trials when the treat was attached to the left string (mean \pm SD: 3.17 \pm 2.41) was less than when it was attached to the right string (mean \pm SD: 4.33 \pm 2.23). However, the difference was not significant (Wilcoxon $T=14$, $P=0.07$, two-tailed), and in any case, because the experimental design was fully counterbalanced, such a bias could not influence the main result.

Discussion

Experiment 4 was the first of the string-pulling experiments where the dogs performed below chance level. The application of their usual strategy of pulling the string

closest to the food did not lead to success in this set-up with the crossed strings. There was no improvement in the probability of choosing the correct string over the 20 trials of the study. This means that the dogs showed no evidence of being able to learn how to pull out the correct string first, at least within the limited training given in the present experiment. It is of course possible that they knew which string to pull, but were unable to inhibit the prepotent response of pulling the string closest to the food, though there was nothing in their behaviour to suggest that they were in a state of conflict. There was no evidence that the dogs applied the method of pulling in the previously reinforced string, which would have been another way of approaching this problem.

The results of experiment 4 indicate that dogs did not spontaneously understand the means-end properties of the strings, and nor did they come to do so over a number of trials.

General discussion

All the present experiments confirmed the observations of Fischel (1933) and others that dogs are able to learn to retrieve food that is out of their reach by means of an attached string. However, the results showed that, unlike the chimpanzees studied by Köhler (1925) and others since, dogs do not solve this task by virtue of understanding the means-end relationship inherent in the physical connection provided by the string. Rather, they apply one of two problem-solving strategies. First, they paw close to the food, even if there is no string there, as in the diagonal condition of experiment 1. If that is unsuccessful, they paw at the string whose proximal end is closest to the food, even when, as in experiment 4, this strategy is not being reinforced. If either of these strategies is successful, its performance is rapidly perfected, leading to the decline in retrieval latencies shown in Fig. 3. Like the results of the string-pulling experiments reported by Scott and Fuller (1965) and Frank and Frank (1985), these present data are fully explicable in terms of associative learning. In Piagetian terms, the dogs displayed sensorimotor intelligence at Stage III (circular reactions), but not at Stage IV (means-end understanding). The results suggest that dogs are performing at much the same level as the cats studied by Adams (1929): when first faced with a string problem, the cats might pay some attention to the string, but did not attack it systematically. However, when their apparently random responses to the string caused the food to move, they suddenly focused on the string and started pulling on it deliberately, usually recovering the food quite quickly.

Since the experimenter and the owner were potentially within the dog's sight and hearing during the tests, it is necessary to consider the possibility of Clever Hans effects (Pfungst 1911). However, the fact that the dogs did not solve the problem immediately means that such effects, even if they existed, were ineffective at least initially. It is possible that the improving performance shown in exper-

iment 1 involved improved use of cues from the humans present, though in this case it is hard to explain why performance deteriorated sharply when each new task was introduced.

These results imply that dogs do not solve a string-pulling problem by understanding the situation and planning a solution. Of course, they may do so in other tasks, but that has yet to be demonstrated. The present results lend further support to the position, suggested by Hare et al. (2002), that dogs owe their reputation for intelligence and their success at problem-solving tasks, in some degree at least, to an acute interspecific social sensitivity. This sensitivity enables them to interpret human wishes and cues better than other species, and in particular better than their wild progenitors. Such sensitivity could be selected for during the history of a domesticated species, or established during the training of an individual, or both. Since Hare et al. (2002) found that wolves reared in domestication do not succeed in tasks that require understanding of human signals, it seems that in dogs deliberate or accidental selective breeding for interspecific social sensitivity has played an important role. On the other hand, such sensitivity does not always require an animal to have been bred for domestication, because it can be seen in trained individuals of non-domestic species such as seals (Scheumann and Call 2004).

Although domestication and training to be responsive to humans may facilitate the performance of some tasks, they may be a hindrance in others. Perhaps dogs have lost their ability to solve problems like string-pulling because in their co-operation with humans, it is always the human who carries out such tasks. In this context, it is interesting to note that in an experiment on string-pulling in African grey parrots, Pepperberg (2004) found that two language-trained parrots demonstrated no means-end understanding, but simply asked their human trainers to give them the treat, whereas parrots that had had no language training solved the problem easily. It appears that the availability of human-aided solutions to problems can sometimes inhibit the expression of animals' cognitive capacities.

Acknowledgements The authors would like to thank Sally Jones from Dogs "R" Dogs and Taunton Dog Training Club for their help with the data collection, David Taylor for constructing and repairing the apparatus, and Dr Lucy Jacobs for facilities provided to Stephen Lea as a Visiting Scholar at the University of California, Berkeley, where the writing of the paper was completed. All experiments were carried out in accordance with laws governing animal welfare in the United Kingdom.

References

- Adams DK (1929) Experimental studies of adaptive behavior in cats. *Comp Psychol Monogr* 6(1):1–128
- Bierens de Haan JA (1930) Über das Suchen nach verstecktem Futter bei Affen und Halbaffen. Zugleich ein Beitrag zu der Frage nach dem konkreten Verständnis dieser Tiere. *Z Vergl Physiol* 11:630–655
- Boakes RA (1977) Performance on learning to associate a stimulus with positive reinforcement. In: Davis H, Hurwitz HMB (eds) *Operant-Pavlovian interactions*. Erlbaum, Hillsdale, pp 67–101
- Bratman M (1981) Intention and means-end reasoning. *Philos Rev* 90:252–265
- Brown AL (1990) Domain-specific principles affect learning and transfer in children. *Cogn Sci* 14:107–133
- Doré FY, Dumas C (1987) Psychology of animal cognition: Piagetian studies. *Psychol Bull* 102:219–233
- Ducker G, Rensch B (1977) Solution of patterned string problems by birds. *Behaviour* 62:164–173
- Eddy TJ, Gallup GG, Povinelli DJ (1993) Attribution of cognitive states to animals: anthropomorphism in comparative perspective. *J Soc Issues* 49:87–101
- Fischel W (1930) Weitere Untersuchungen der Ziele der tierischen Handlung. *Z Vergl Physiol* 11:523–548
- Fischel W (1933) Das Verhalten von Hunden bei doppelter Zielsetzung und doppelter Handlungsmöglichkeit. *Z Vergl Physiol* 19:170–182
- Frank H, Frank MG (1985) Comparative manipulation-test performance in ten-week-old wolves (*Canis lupus*) and Alaskan malamutes (*Canis familiaris*): a Piagetian interpretation. *J Comp Psychol* 99:266–274
- Funk MS (2002) Problem solving skills in young yellow-crowned parakeets (*Cyanoramphus auriceps*). *Anim Cogn* 5:167–176
- Grzimek B (1942) Weitere Vergleichsversuche mit Wolf und Hund. *Z Vergl Physiol* 5:59–73
- Hare B, Brown M, Williamson C, Tomasello M (2002) The domestication of social cognition in dogs. *Science* 298:1634–1636
- Hauser MD, Kralik J, Botto-Mahan C (1999) Problem solving and functional design features: experiments on cotton-top tamarins, *Saguinus oedipus oedipus*. *Anim Behav* 57:565–582
- Hauser MD, Santos LR, Spaepen GM, Pearson HE (2002) Problem solving, inhibition and domain-specific experience: experiments on cottontop tamarins, *Saguinus oedipus*. *Anim Behav* 64:387–396
- Heinrich B (1995) An experimental investigation of insight in common ravens (*Corvus corax*). *Auk* 112:994–1003
- Hood BM, Hauser MD, Anderson L, Santos L (1999) Gravity biases in a non-human primate? *Dev Sci* 2:35–41
- Klüver H (1933) *Behavior mechanisms in monkeys*. University of Chicago Press, Chicago
- Köhler W (1925) *The mentality of apes*. Harcourt Brace, New York
- Munakata Y, McClelland JL, Johnson MH, Siegler RS (1997) Rethinking infant knowledge: toward an adaptive process account of successes and failures in object permanence tasks. *Psychol Rev* 104:686–713
- Nakajima S, Arimitsu K, Lattal KM (2002) Estimation of animal intelligence by university students in Japan and the United States. *Anthrozoös* 15:194–205
- Osthaus B, Slater AM, Lea SEG (2003) Can dogs defy gravity? A comparison with the human infant and a non-human primate. *Dev Sci* 6:489–497
- Page EB (1963) Ordered hypotheses for multiple treatments: a significance test for linear ranks. *J Am Stat Assoc* 58:216–230
- Pepperberg IM (2002) The value of the Piagetian framework for comparative cognitive studies. *Anim Cogn* 5:177–182
- Pepperberg IM (2004) "Insightful" string-pulling in grey parrots (*Psittacus erithacus*) is affected by vocal competence. *Anim Cogn*. DOI 10.1007/s10071-004-0218-y
- Pfungst O (1911) *Clever Hans (the horse of Mr Von Osten): a contribution to experimental animal and human psychology*. Henry Holt, New York
- Piaget J (1953) *The origin of intelligence in the child*. Routledge and Kegan Paul, London
- Santos LR, Ericson BN, Hauser MD (1999) Constraints on problem solving and inhibition: object retrieval in cotton-top tamarins (*Saguinus oedipus oedipus*). *J Comp Psychol* 113:186–193
- Sarris EG (1937) Die individuellen Unterschiede bei Hunden. *Z Angew Psychol Charakterk* 52:5–6

- Scheumann M, Call J (2004) The use of experimenter-given cues by South African fur seals (*Arctocephalus pusillus*). *Anim Cogn*. DOI [10.1007/s10071-004-0216-0](https://doi.org/10.1007/s10071-004-0216-0)
- Scott JP, Fuller JL (1965) Genetics and the social behavior of the dog. University of Chicago Press, Chicago
- Shepherd WT (1915) Tests on adaptive intelligence in dogs and cats, as compared with adaptive intelligence in Rhesus monkeys. *Am J Psychol* 26:211–216
- Spinuzzi G, Poti P (1993) Piagetian stage 5 in 2 infant chimpanzees (*Pan troglodytes*)—the development of permanence of objects and the spatialisation of causality. *Int J Primatol* 14:905–917
- Tolman EC (1937) The acquisition of string-pulling by rats—conditioned response or sign-gestalt? *Psychol Rev* 44:195–211
- Vince MA (1956) ‘String pulling’ in birds. I. Individual differences in wild adult great tits. *Br J Anim Behav* 4:111–116
- Vince MA (1958) “String-pulling” in birds. II. Differences related to age in greenfinches, chaffinches, and canaries. *Behav* 6:53–59
- Vince MA (1961) “String-pulling” in birds. III. The successful response in greenfinches and canaries. *Behaviour* 17:103–129
- Washburn SL, Lancaster J (1968) The evolution of hunting. In: Lee RB, De Vore I (eds) *Man the hunter*. Aldine, Chicago, pp 293–303
- Willatts P (1984a) Stages in the development of intentional search by young infants. *Dev Psychol* 20:389–396
- Willatts P (1984b) The stage-IV infant’s solution of problems requiring the use of supports. *Infant Behav Dev* 7:125–134