



## **Can Dogs Learn Concepts the Same Way We Do? Concept Formation in a German Shepherd**

**Erica N. Feuerbacher and Jesús Rosales-Ruiz**

*University of North Texas*

Growing evidence shows that dogs can complete complex behavioral tasks, such as learning labels for hundreds of objects, readily learning the name of a novel object, and responding differentially to objects by category (e.g., “toy,” “ball,” “Frisbee”). We expand here on the evidence for complex behavioral abilities in dogs by demonstrating that they are capable of concept formation by strict criteria. A German shepherd responded differentially to two sets of objects (“toys” and “non-toys”) in Experiment 1. Additionally, the dog’s differential responding in Experiment 1 occurred from the first trial, indicating that he entered the experiment with this stimulus class already differentiated from his day-to-day exposure to contingencies. In Experiment 2, we used a common response (tug-of-war) with three objects that were not retrieved in Experiment 1 to attempt to add these objects to the stimulus class. After repeated sessions of tug-of-war, the dog began retrieving all three objects in the retrieval test, although the rates of retrieval varied between objects. Finally, in Experiment 3, we conducted a transfer of function test in which the dog emitted a new response to untrained exemplars suggesting that his differential responding in Experiment 1 was indicative of a concept by the strictest criteria. Additionally, he reliably emitted the new response in the transfer test to one of the three new objects from Experiment 2, suggesting this object had been reliably added to the conceptual class.

The ability to categorize or form concepts about the world can be a useful adaptation; it effectively allows those individuals that form concepts to more readily adapt their behavior to changing contingencies than individuals that have not formed categories (Herrnstein, 1990). For example, if a food item that was once associated with reinforcement now becomes associated with aversive stimuli (perhaps a fruit now has a bitter and nausea-inducing chemical sprayed on it) animals that behave based on categorization or concepts can more easily change their behavior towards the fruit than one that does not. Humans have the ability to form functional concepts, such as toys, furniture, and tools, objects that might not resemble each other in stimulus features but are united by their functions and how humans interact with them. Whether animals can form such functional concepts remains an empirical question.

One type of complex stimulus relations is equivalence. Equivalence refers to a specific set of stimulus relations that stem from mathematical relations, namely reflexivity ( $A = A$ ), symmetry (if  $A = B$ , then  $B = A$ ) and transitivity (if  $A = B$  and  $B = C$ , then  $A = C$ ; Sidman, 2000). In behavioral equivalence, untrained relations emerge. For example, we could teach a child to choose a picture of a flower from an array of other pictures when the experimenter says, “flower” ( $A = B$ ). We could also teach the child to choose the word flower from an array of other words when presented with the picture of the flower ( $B = C$ ). If the child’s behavior showed equivalence relations, without training, the child would choose the word flower from an array of other words when the experimenter says, “flower” ( $A = C$ ). Evidence from sea lions (Schusterman & Kastak, 1993, 1998) suggests that animals are capable of forming equivalence relations and learning the equivalence of arbitrary stimuli, and is thought to be an ability related to language acquisition (Dickins & Dickins, 2001) and possibly a precursor to more advanced linguistic ability (Schusterman & Kastak, 1993). To test for equivalence, a sea lion was trained  $A = B$  and  $B = C$  relations for 30 sets of stimuli. The sea lion responded by emitting a nose

touch to the stimulus. The sea lion was then tested for equivalence relations by testing for correctly choosing A when given C ( $C = A$ ); it succeeded in showing equivalence relations, including symmetry and transitivity in the first trial for 11/12 novel relations and full equivalence on the first trial in 16/18 novel relations. Such a performance is indicative that at least some animals could form functional concepts. A subsequent study in sea lions further indicated that having a common reinforcer for each set of stimuli allowed the animals to readily form functional relations between the stimuli of those sets (Kastak, Schusterman, & Kastak, 2001). This result follows from the equivalence formulation that Sidman (2000) provided, in which he suggested that groups of stimuli can be united by common responses or common reinforcers.

One strong candidate for showing functional concepts is the domestic dog. Dogs have surpassed primates in some cognitive tasks including using various markers to find hidden food (Hare, Brown, Williamson, & Tomasello, 2002) and correctly responding to imperative pointing (Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012). Additionally, they are a social species, which, at least in corvids, correlates with the complex cognitive task of transitive inference (Bond, Wei, & Kamil, 2010). Especially given Western dogs' requirements to survive in human homes, categorizing stimuli would enhance their ability to behave effectively with humans and our behavioral demands. For example, dogs need to be able to discriminate objects on which they are allowed to chew (e.g., bones or stuffed toys) and objects which they are not (e.g., shoes and furniture); similarly, dogs need to be able to discriminate animals which they are allowed to chase (e.g., wild birds) from animals they are not allowed to chase (e.g., family chickens). Not having to learn the contingencies each individual stimulus is involved in, and instead being able to generalize and respond to novel stimuli based on category membership, would be advantageous.

Furthermore, growing evidence indicates dogs can complete complex cognitive tasks, including learning the names of novel stimuli via fast-mapping (Kaminski, Call, & Fischer, 2004) and can learn over 1000 objects by name including responding differentially to objects by category (e.g., "toy," "ball"; Pilley & Reid, 2011). Dogs can also differentially respond to pictures with and without dogs (Range, Aust, Steurer, & Huber, 2008), including novel exemplars. Such performances are indicative of perceptual or open-ended categories, which include categories such as humans, trees, and cars. Category membership is based on physical characteristics and generalization across exemplars (Keller & Schoenfeld, 1950). Gorillas (Vonk & McDonald, 2002), monkeys (D'Amato & Van Sant, 1988) and pigeons (Herrnstein & Loveland, 1964) have experimentally demonstrated this type of stimulus control.

A higher-order category, variously called associative, non-similarity based, or functional concept or stimulus class (Goldiamond, 1966; Herrnstein, 1990; Zentall, Galizio, & Critchfield, 2002), is based on functional rather than perceptual properties (e.g., toys, furniture, and tools). That is, animals respond not on stimulus feature similarity but rather class membership is based on common responses, outcomes, functions or associations (Fields, Landon-Jimenez, Buffington, & Adams, 1995; Keller & Schoenfeld, 1950; Schusterman & Kastak, 1993; Schusterman, Reichmuth, & Kastak, 2002; Sidman, 2000; Urcuioli, 2001; Wasserman, DeVolder, & Coppage, 1992). In functional concepts each member of the class is representative of the entire concept; this means that if the function of one member of the class changes stimulus changes, the function of all the other members of that class also changes (Fields et al., 1995; Goldiamond, 1966; Wang, Dack, McHugh, & Whelan, 2011; Zentall et al., 2002). Herrnstein (1990) provided an example of this: if acorns were suddenly to become bitter, an animal with a categorization of acorns at the concept-level would more quickly change its behavior to all acorns (i.e., the function of acorns for the animal would change) than an animal that was not behaving based on categorization at the concept-level. Thus, a test for a functional category is a transfer of function test (Lea, 1984). In a transfer of function test, a new response is trained to one item of the purported concept and then other members of the concept are tested to determine whether they elicit the same new

response. If the animal emits the new response to other members of the class, this performance is congruent with a concept, but if the animal does not emit the new response to other members of the purported concept, the collection of stimuli are considered part of a category but not a concept (Lea, 1984).

There is compelling evidence that other animals share this level of categorization with humans (see Schusterman & Kastak, 1993; Urcuioli, 2001; Zentall, Wasserman, Lazareva, Thompson, & Rattermann, 2008). However, with respect to dogs, researchers have not yet experimentally demonstrated functional concepts. Pilley and Reid (2011) tested Chaser, a border collie, and found she could correctly identify objects as “toys,” objects of a various shapes and colors with which she had been allowed to play. While this is suggestive of a functionally based concept, the authors did not test for transfer of function; the question of whether dogs can form functional concepts remains.

Aside from the research on Chaser by Pilley and Reid (2011), one criticism of typical demonstrations of conceptual behavior in animals is that these behaviors were directly trained in experimental environments (e.g., match-to-sample, Galizio, Stewart, & Pilgrim, 2004; go/no-go, Range et al., 2008; or contingency reversals, Vaughan, 1988). While laboratory demonstrations of conceptual behavior in animals demonstrates their capacity to behave based on concepts, this does not necessarily mean that animals use concepts in their day-to-day lives; that is, despite animals’ capacity to form concepts, if the requisite contingencies are not in place in their daily lives, they will still not form concepts in their natural environment. Further research is needed to clarify whether interactions between the animal and natural environment lead to development of concepts and, if so, how (Zentall et al., 2002).

Sidman’s (2000) formulation of equivalence relations offers a viable mechanism for concept formation in everyday experiences. In this formulation, stimulus classes can be united through common reinforcers and responses to the stimuli, without the stimuli ever having been directly related to one another (e.g., conditional discriminative stimuli). For example, the concept of “sporting goods” mentioned by Premack (1983) could be united in this way, as the person would engage with those physically different objects in a similar way. Finally, Sidman’s formulation also parallels Herrnstein’s (1990) view of concepts that these complex performances are sensitive to the contingencies of reinforcement in which the conceptual stimuli are involved.

Sidman’s formulation that equivalence classes can be united by common responses or common reinforcers is supported by the use of common reinforcers to unite functional classes in sea lions (Kastak et al., 2002). Chaser’s categories of toys and non-toys also appeared to be based on common contingencies of reinforcement that served to unite objects within those categories while separating between the two categories (Pilley & Reid, 2011). The contingencies for Chaser were not experimentally programmed either, but rather a result of the contingencies of reinforcement she encountered through living with her owner. This suggests that common responses to stimuli or common reinforcers might be viable mechanisms for concept formation in dogs.

Thus, three questions can be identified for studying concept formation in dogs: 1) can dogs form a concept according to the criterion of showing a transfer of function; 2) do dogs form concepts in the natural environment, as opposed to experimentally induced concepts; and 3) can we identify a mechanism for concept formation in the dog’s natural environment?

In the current study, we addressed all three questions: whether a dog would respond differentially to two sets of stimuli (Experiment 1); whether we could identify a mechanism for concept formation by changing the function of objects for the dog through common response training to test whether we could expand the

putative concept (Experiment 2); and finally whether the dog would show a transfer of function within these sets of stimuli when the contingency of reinforcement was altered (Experiment 3), which would suggest a concept in the strictest sense.

We investigated the concept “toys” because toys are a class that is not strictly feature based, although feature-based generalization can still occur within the class, but rather one that is defined through the contingencies in which the animal encounters a given object. Additionally, a pet dog likely has had extensive history with a variety of play objects and may have formed a concept of toys concept in the dog’s day-to-day experiences.

To investigate whether dogs have formed a concept of toys, we first tested whether a dog would differentially respond to two sets of objects, which we labeled as toys and non-toys. Second, we investigated whether we could add non-toys to the toy class by having the dog specifically play with non-toys. Finally, we used a reassignment task to test for transfer of stimulus function (Fields et al., 1995). Aero, a male German shepherd, was selected because of his extensive history playing “tug-of-war” and “fetch” with toys, which we used as common responses to add new exemplars to the concept. We suspected that these activities, which are likely reinforcers for dogs, could be common reinforcing responses that serve to unite the class toys. Such a possibility was suggested also by Chaser’s ability to discriminate toys from non-toys based on their contingencies of reinforcement and not their physical features.

If dogs can form concepts, this would help us understand the extent of their cognitive abilities and how they compare to humans. It also is a practical question as we are training dogs to complete more and more complex tasks, which might require concept formation (e.g., complex scent detection). Understanding if and how they form concepts could let us train them for these tasks more effectively.

## **Method**

### **Subject**

Aero, a male German shepherd dog, *Canis lupus familiaris*, participated in this study. Aero was 6 years old at the start of the study. He was owned by the first author and was selected as the study’s participant because of an extensive training history (clicker training), as well as an extensive history of playing with toys, largely through games of “tug-of-war” and “fetch.” Given his history with the first author, she was able to provide insight into possible common contingencies between objects. However, he was adopted from a rescue at 6 months old; thus, his entire history of reinforcement with various objects is not fully known.

### **Settings**

Sessions took place in the living room and kitchen of the dog’s home (Figure 1). For Experiments 1 and 2, the experimenter placed a small woven rug (0.5 m x 1 m) on the carpet in the living room. When applicable, the experimenter conducted training sessions on that rug, and placed objects in the center of the rug.

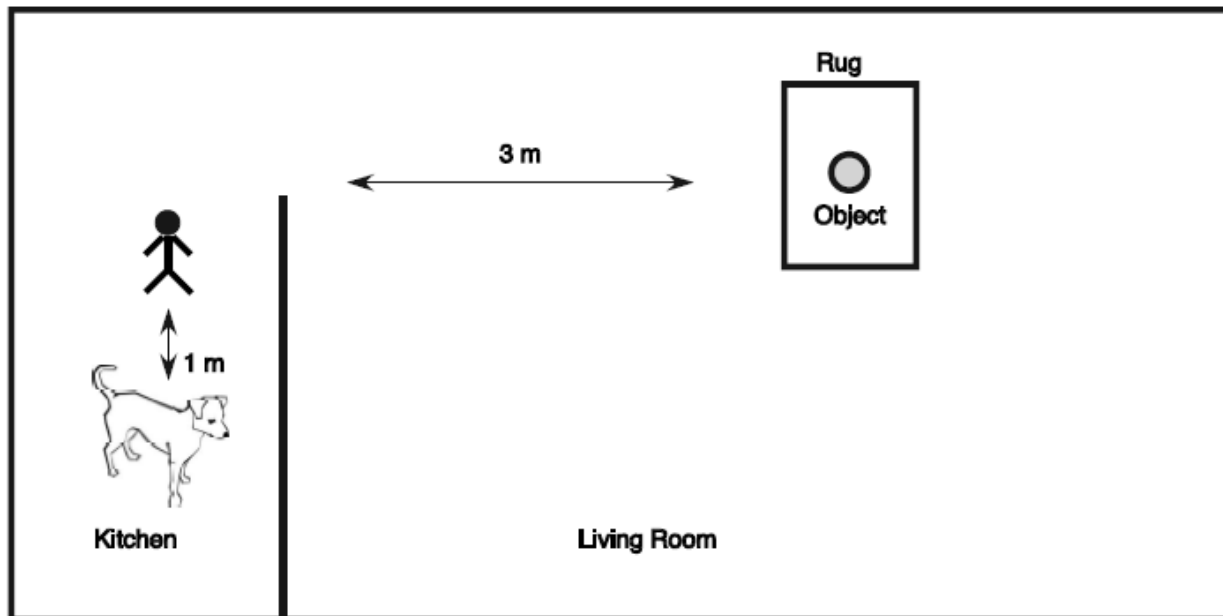


Figure 1. Diagram of the experimenter, dog, and object placement during the retrieval test sessions (Experiments 1 & 2). Diagram is not drawn to scale.

## Stimuli

Through five preliminary trials, we identified six objects that Aero retrieved on all five preliminary trials when told, “Go get it,” by the experimenter while she pointed towards the object. We also identified three objects that Aero reliably did not retrieve when tested the same way (i.e., told to “go get it” by the experimenter while she pointed; see Table 1 for a description of all objects). The retrieved objects were: a black rubber squeaky dumbbell (black dumbbell), a plush orange bear (orange bear), a plush round alligator (alligator), a piece of fake lambswool tied in a knot (lambswool knot), a set of interlaced rope rings with tennis balls on each ring (rope rings), and a plush toy with four small ropes spiking out of it (plush geometric). During the course of the experiment, four of the original retrieved objects were traded out for identical, new objects because the original began to show wear and tear, with no effect on Aero’s response to those objects. The non-retrieved objects were a medium-sized red rubber KONG® (Kong), a plastic Frisbee (Frisbee), which was turned upside down to ensure the dog could pick it up, and a blue rubber tug toy (blue tug).

As the experiment progressed we discontinued testing on some retrieved objects and instead introduced novel non-retrieved objects to increase the number of non-retrieved objects and keep that number similar to the original number of non-retrieved objects. The added objects included a plastic football (football), a honeycombed red rubber ball (red ball), a blue round plastic squirrel (squirrel), and a rubberized plastic blue geometric shape (blue geometric). To find these four non-retrieved items, we added in the new object to our array and tested it in our normal retrieval test. Three other objects we added in when looking for non-retrieved items Aero retrieved so we discontinued those (he usually began to retrieve them within three trials of being added to the array). We do not report on those items.

All of the objects used were novel except the black dumbbell 219 and the Kong. Of these two items, preliminary retrieval trials indicated that the black dumbbell was likely to be retrieved and the Kong was not. Thus the retrieved and non-retrieved groups were balanced for familiarity. For the duration of the experiment, Aero did not have any access to the objects used in the study other than during the training and test sessions. All objects were handled only by the first author and were stored together in a box between sessions, reducing the possibility that Aero could respond on olfactory cues.

All sessions were video recorded and responses were scored from the videos. We double scored 30% of the videos using a trained, independent observer. We calculated interobserver agreement (IOA) by session on a trial-by-trial basis. IOA initially ranged

from 90% to 100% with a mean of 99.5%. We reviewed the disagreements, which occurred the few times when the dog picked up the item and carried it the required distance off the rug but dropped it before crossing the threshold to the kitchen. The second observer had mistakenly coded these as nonretrievals. We retrained the second reviewer how to code those and subsequently agreement was 100%.

Table 1  
*Object Names and Descriptions*

Object Name	Object Description
<b>Toys</b>	
Black dumbbell	19 cm rubber dumbbell, 5.1 cm in diameter in center. Either end has small bumps on them
Plush alligator	11.5 cm diameter spherical plush green alligator
Rope rings	Two rope rings each with a tennis ball on them. Approximately 30.5 cm when rings fully extended
Plush geometric	7.6 cm diameter plush sphere with five 8.9 cm rope extensions protruding from center sphere
Lambswool knot	25.6 cm piece of fake lambswool tied into a knot
Plush bear	22.9 cm long orange plush bear
<b>Non-Toys</b>	
Plastic squirrel	12.7 cm blue spherical rubber squirrel with a handle protruding 5.6 cm from one side
Football	22.9 cm x 8.9 cm brown football with indentations on each side
Blue geometric	Blue, soft, rubber, hollow shape with six protruding arms. 16.5 cm arm tip to arm tip
Frisbee	Blue 24.1 cm diameter soft rubber disc
<b>New Toys</b>	
Blue tug	25.6 cm flat rubber tug toy
Kong	Red rubber KONG® 21.6 cm x 12.7 cm (at bottom)
Red ball	14 cm diameter honeycombed, hollow, soft rubber ball

## Experiment 1: Stimulus Class Test

We first investigated whether Aero would differentially retrieve two sets of objects. This tested whether Aero had a stimulus class of toys (objects he would retrieve) versus non-toys (objects he would not retrieve). Such a performance would meet the first criterion for demonstrating concept formation.

### **Trials and Sessions**

The experimenter stood at the junction of the kitchen and living room and faced into the kitchen towards Aero, who was in a sit-stay in the kitchen approximately 1 m from the experimenter. The experimenter then placed the object to her left approximately 3 m away in the middle of the small rug. Because of the dividing wall (Figure 1), Aero was not able to see into the living room where the objects were placed. When the experimenter had returned to the designated position facing Aero she cued him to retrieve the object by pointing to the living room and saying, “Go get it.” Aero ran past the experimenter into the living room toward the object. The experimenter remained stationary in the kitchen and was behind Aero when he responded to the object. If he picked up the object and began to return to the experimenter she said, “Yes” when he was within 0.25 m of her and then gave him a piece of Natural Balance®

(approximately 1 cm<sup>3</sup>) from a bag on the counter, regardless of whether he had carried the object all the way back, or dropped it along the way. If he picked up the object and carried it at least 1 m off the rug, the trial was scored as retrieval.

Alternatively, if he met one of the following four nonretrieval criteria, the experimenter said, “here” to recall him and give him a piece of Natural Balance® when he returned to her: 1) if he ran out to the object and turned around to come back (the most common response); 2) stood over the object but scanned back and forth for a few seconds without putting his mouth on the object (the second most common response); 3) ran across and past the object; or 4) picked up the object but dropped it again on the rug and did not pick it up again within 5 s (least common response). All of these trials were scored as nonretrieval.

For objects Aero retrieved, he immediately ran to the object and put his mouth on it to pick it up. For objects that he did not retrieve, he would typically run out to them but not manipulate them in any way and then immediately turn back to come back to the experimenter. Thus, he was usually already returning to her before the “here” cue was given. If he did put his mouth on the object, the experimenter did not cue him in any way. She waited for him to either 1) return to her with the object to tell him “yes” and give him a treat (scored as retrieval), or 2) waited until he had instead released the object from his mouth and had not touched it again for 5 s before saying, “here” and giving him a treat (scored as a nonretrieval). This was to ensure that the experimenter did not interfere or prematurely call him away from an object he might retrieve. Additionally, because of these methods, Aero received a piece of food on every trial independent of retrieval behavior. After he returned to the kitchen the experimenter cued him to “sit and stay” to prepare for the next trial.

The experimenter initially cued Aero only once with a simultaneous “Go get it” and finger point. However, on some trials with non-toys Aero began to not run all the way out to the object. Thus, beginning in Session 10, the experimenter would re-cue Aero when necessary until he came within 15 cm of the edge of the rug (approximately 40 cm from the center of the object). This ensured that Aero came into contact with all objects. Typically if a re-cue was necessary, only one re-cue per trial was required.

Each object was tested once per session and only one retrieval session per day was conducted. We randomized the sequence of presentation of the objects, except that we did not present more than two objects that we expected him to retrieve or two objects we did not expect him to retrieve in a row.

## Results

We conducted a total of 112 retrieval test sessions (this includes all retrieval test sessions across all three experiments). One session (Session 23) was a probe session in which we only tested the blue tug (see Experiment 2). Thus, a total of 111 full retrieval sessions involving all stimuli were conducted. Figure 2 shows the cumulative retrieval responses for the all the objects tested. For the three objects that we used in common response training in Experiment 2 (blue tug, Kong, red ball), we have plotted here only the trials prior to that training. Each session in which he retrieved the object is plotted by an uptick on the line for that object: a higher slope of the line indicates a higher rate of retrieval for that object. Figure 3 shows the summary data for those same objects as percentage of trials on which he retrieved that object.

During all phases of the three experiments, Aero reliably retrieved the black dumbbell, (111 retrievals out of 111 trials), the plush alligator, (111 retrievals out of 111 trials), the rope rings, (106 retrievals out of 111 trials), the plush geometric, (104 retrievals out of 111 trials), the lambswool knot (54 retrievals out of 55 trials), and the plush bear (53 retrievals out of 54 trials). The lambswool knot and the plush bear were discontinued, which accounts for the reduced number of total sessions in which these objects were tested.

Aero never retrieved the Frisbee (0 retrievals out of 111 trials), the blue geometric (0 retrievals out of 52 trials), the football (0 retrievals out of 49 trials), the squirrel (0 out of 34 trials), the blue tug (0 retrievals out of 10 trials), the Kong (0 retrievals out of 34 trials), or the red ball (0 retrievals out of 7 trials). The red ball, blue geometric, football, and squirrel, were added as stimuli in sessions 58, 60, 63, and 78, respectively.

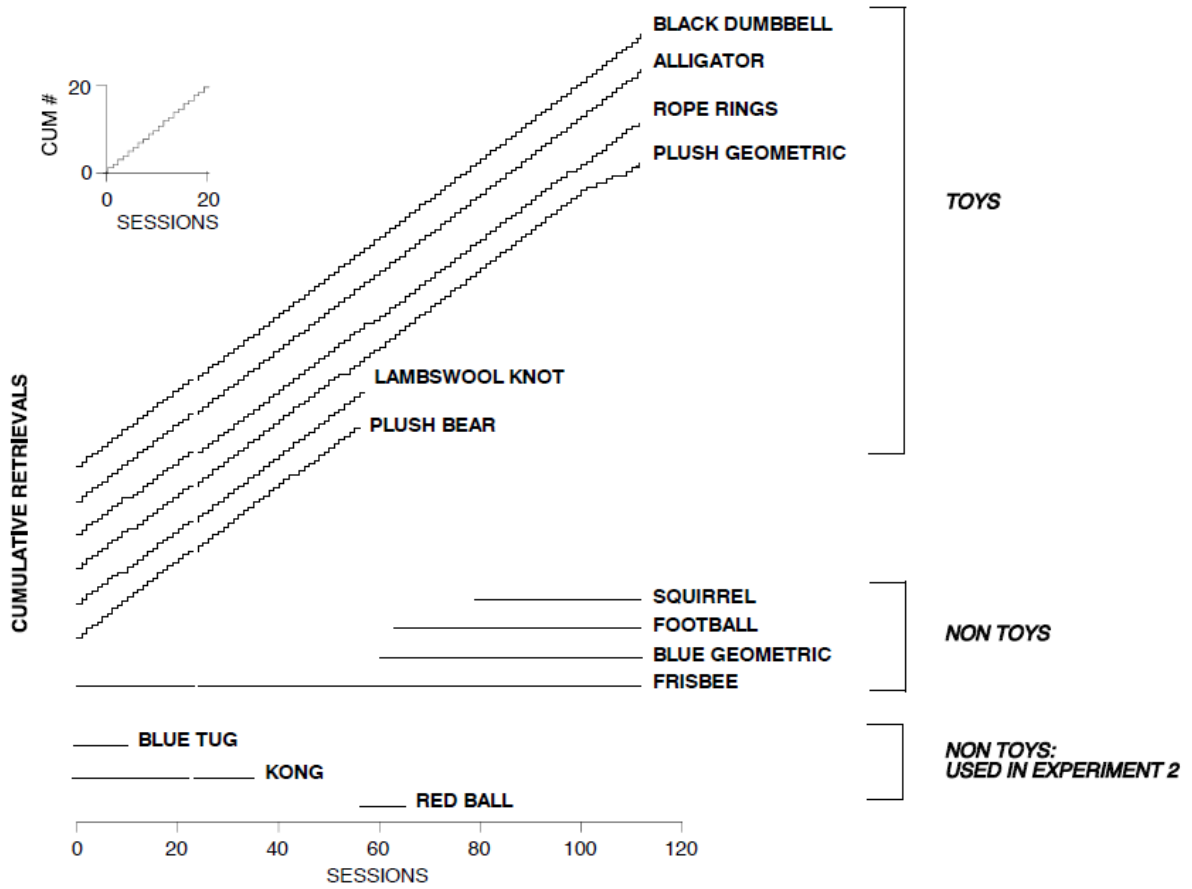


Figure 2. Cumulative record of retrievals of “toys” and “non-toys” in the retrieval tests. Objects that were reliably retrieved were labeled toys, and those that were not retrieved were labeled non-toys (Experiment 1). For ease of identification, we have separately identified non-toys that were not experimentally manipulated in Experiment 2 and the three objects that were.

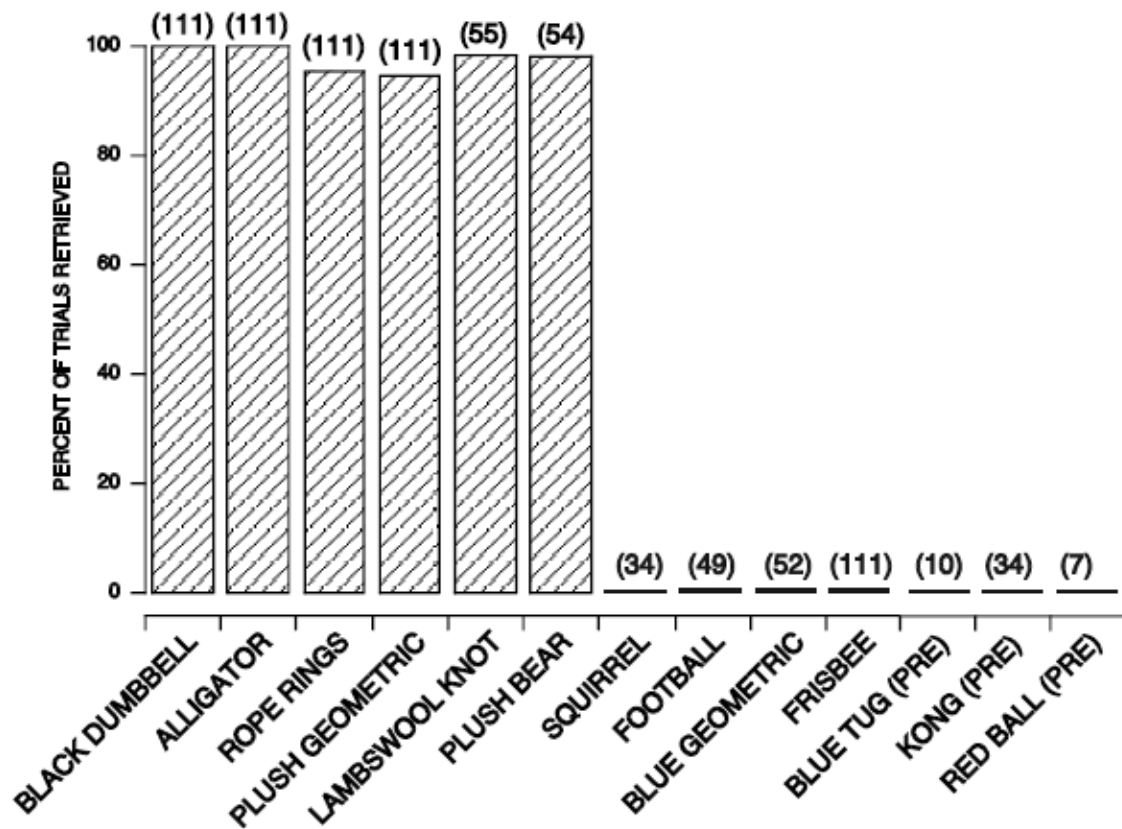


Figure 3. Percentage of retrievals for toys and non-toys across all retrieval test sessions in which the object was involved (Experiment 1). For the three toys that were used in Experiment 2 (blue tug, Kong, and red ball), the trials used to calculate this are the trials that occurred prior to any retrievals in the common response training phase. The number of sessions in which the particular object was tested is given in the parentheses.

## Discussion

Based on the retrieval responses, the experimenter labeled those items that Aero retrieved as “toys” and those that he did not retrieve as “non-toys.” Using the criterion of differential responding to two sets of stimuli (Herrnstein & Loveland, 1964; Huber, 2001), the results suggested that Aero did have a “toy” concept.

Additionally, Aero responded differentially from the first session and we did not differentially reinforce retrievals or non-retrievals, suggesting he entered the experiment with this concept. In fact, despite there being no contingency in place to retrieve (i.e., whether he retrieved or not he received a piece of food), he continued to reliably retrieve the objects we called toys. Given that we do not know Aero’s history prior to 6 months of age, we cannot know for sure what first united the group of stimuli that Aero retrieved compared to the stimuli he did not retrieve. That is, were the retrieved toys united by tug-of-war compared to objects that were never played tug-of-war with, or were they united by a lack of punishment whereas nonretrieved items

were associated with punishment? Nevertheless our results indicate that these objects are members of different stimulus classes, similar to the toys and non-toys of Chaser (Pilley & Reid, 2011).

Of the objects used, only two were familiar in any way to Aero; the others likely had stimulus features in common with other objects with which Aero had interacted, but not this specific configuration or features. Aero's differential responding to these novel objects is likely a product of stimulus feature generalization from other toys he had played with and retrieved in his life. The objects that were non-toys likely did not share sufficient stimulus features with other objects that Aero responded to as "toys." The non-toys were typically rubber or plastic compared to the rope or plush objects in the toy class.

Additionally, Aero's performance did not change when we replaced the four worn objects with new replicas, indicating that Aero's performance was not a function of any olfactory cues from having previously retrieved that item. His performance paralleled that of Chaser who also entered the experiment with a toy class (Pilley & Reid, 2011).

## **Experiment 2: Adding Objects to the Class**

We next investigated whether we could identify a mechanism by which the toy stimulus class might have been established through Aero's natural, day-to-day experiences. To do this, we investigated whether we could add new items to the toy class; that is, could we take currently non-retrieved items and change the function of 334 those objects for Aero so that he would now retrieve them?

We hypothesized that the objects that Aero retrieved were similar to objects with which he had engaged in tug-of-war play and that we could add non-toys to the toy class by engaging Aero in a common response (tug-of-war; Sidman, 2000; Zentall, Galizio, & Critchfield, 2002) with those non-toy items.

### **Method**

We used the retrieval test (Experiment 1) as our metric for whether the non-toy objects had changed function to toys and would now be retrieved by Aero. We used a multiple baseline across three objects design (blue tug, Kong, and red ball) to test whether Aero would start to retrieve the target objects after common response training. Only one target object was trained at a time, such that after the training phase ended for one particular object, no further training sessions occurred with that object. For example, after the phase ended in which the blue tug was the target object for common response training, training for the blue tug stopped and the experimenter switched to training the Kong.

### **Common Response Training**

All training sessions took place in the living room on the small rug. The experimenter stood facing Aero, no more than 30 cm in front of Aero. To clarify the exact conditions under which we could change a non-toy to a toy, we used two experimental conditions: tug-of-war and interspersed training.

**Tug-of-war.** In tug-of-war, the experimenter played with Aero using the targeted object, pulling backwards or side to side on the object for approximately 30 s. Initially in the tug-of-war training sessions the experimenter presented an object with both hands at Aero's eye level and approximately 1 m from Aero's nose to initiate a game of tug with Aero. This way of initiating play was used for the first four training sessions involving the blue tug object, but all subsequent sessions were initiated by the experimenter placing the object on the rug and making a fast hand movement toward the object. This was a more reliable way of initiating play with Aero. When Aero grabbed the object, the experimenter engaged in a typical bout of tug of war, pulling backwards or side to side on the object. After approximately 30 s, the experimenter cued Aero to drop the object by saying, "Leave it." When Aero released the object the experimenter offered it to him again, and another bout of tug-of-war ensued. This sequence was repeated until 5 min elapsed, at

which point the experimenter cued Aero to drop the object. After he dropped the object she gave him either a piece of Natural Balance® or a preferred object not in the study to reinforce his behavior of relinquishing the object. In each 5-min session approximately 10 bouts of tug occurred.

**Interspersed.** The interspersed training sessions were similar to the tug-of-war sessions, except play bouts alternated between the target (non-toy) object and a toy (an object that was already consistently retrieved in the retrieval test trials of Experiment 1), which could possibly enhance the non-toy interaction as it would occur in close temporal proximity to interaction with a current toy. The experimenter initiated play bouts in one of two ways: for play bouts involving a toy, the experimenter placed both the toy and the target object on the rug and initiated play by making a fast hand movement toward the object. Aero would typically pick up the current toy in such situations. Thus, for play bouts involving the target object the experimenter placed only the target object on the rug to preclude Aero from continuing to engage only with the current toy.

The first tug bout always involved the current toy. After Aero picked up the toy, the experimenter engaged in a bout of tug with Aero. After approximately 30 s of tug-of-war, the experimenter cued Aero to drop the toy and subsequently initiated another play bout with the target object. After 30 s of tug with the target object, the experiment cued him to drop the target object and initiated a play bout with the toy. Play with the two objects thus alternated throughout the 5-min session. In each 5-min session approximately 10 alternations occurred, so that each object was involved in five bouts of tug per session. The toys that the experimenter used during this training were 1) the black dumbbell or 2) the rope rings.

### **Blue Tug**

Tug-of-war was used initially with the blue tug but after eight sessions with no change in retrieval responses, we then changed to interspersed training. Because after four sessions of interspersed training Aero began to retrieve the blue tug, it was not clear whether the interspersed training was the effective variable or whether more sessions of tug-of-war would have also resulted in Aero retrieving the object.

### **Kong**

Because we could not determine whether it was the longer exposure to common response training or the use of interspersed training that resulted in Aero retrieving the blue tug, we used interspersed training only with the Kong during the initial training. The retrieval rate of the Kong slowly decreased after training ended, thus we conducted a second training with the Kong after training the third object (red ball). During this second training, we used tug-of-war training with the Kong rather than interspersed.

### **Red Ball**

To assess whether longer exposure to tug-of-war would also result in retrieval of the object, we used tug-of-war training with the red ball.

### **All Three Objects Final Training**

We conducted a refresher training for all three objects during the last twelve sessions when we are also conducting the transfer of function test (Experiment 3). During this refresher training, we used interspersed training. These interspersed training sessions consisted of play alternations as described in the Interspersed Training section. However, now they involved multiple objects. In these interspersed training sessions we used the plush geometric and the black dumbbell (toys) along with the blue tug, Kong, and red ball. The black dumbbell was included in these sessions because it was the object being used to train a novel response during the transfer test (Experiment 3) and we wanted to ensure that Aero's retrieval response to that object was not disrupted by the training of a second response to it.

These multiple interspersed training sessions consisted of alternating 30 s play bouts between the various objects until all of the objects had been played with once, at which point the sequence of play was repeated two more times. The sequence of object play was the same throughout the three cycles of a play session, but varied between sessions. Because of the additional toys, training sessions were not limited to 5 min, but rather continued until all objects had been played with for 30 s three times each.

## Retrieval Testing

During the common response training sessions, we continued to conduct retrieval sessions to determine if Aero would start to retrieve the targeted objects. We conducted common response training sessions first and later the same day conducted the retrieval session. Retrieval sessions were run exactly as they were in Experiment 1.

## Analysis

To aid in visual analysis, we used a structured criteria method developed by Saini, Fisher, and Retzlaff (2017) to handle binary data (whether the behavior occurred or not). To do this, we calculated a mean response rate for each of the three objects targeted for class inclusion. Then, we added the number of trials in the control phase for each object (baseline) in which the dog responded in accordance with the object not being a toy (below the mean) plus the number of trials after common response training commenced in which the dog responded in accordance with the object being a toy (above the mean). Because the mean fell between 0 and 1 for each object, this meant that nonretrievals during baseline (below the mean) were in accordance with the function of non-toy and retrievals after common response training commenced (above the mean) were in accordance with the function of a toy. The required number of trials to show an effect, (i.e., in which the response was in the direction of the predicted function of toy or non toy) was based off the binomial distribution using the total number of trials in which the object was tested (Fisher, Kelley, & Lomas, 2003).

## Results

**Blue tug.** Figure 4 shows the cumulative retrieval responses for the three objects we targeted for class inclusion. Figure 5 summarizes the percentage of retrievals before and after common response training began for each of these objects. After 10 baseline sessions during which Aero never retrieved the blue tug, the experimenter conducted eight sessions of tug training (Sessions 11-18), each followed by a retrieval test session in which all objects in the study were tested (see Methods for Experiment 1). Aero did not retrieve the tug in a 446 ny of these trials. In Session 19 the experimenter began conducting interspersed training sessions prior to retrieval testing. A total of 12 retrieval test sessions were conducted during the interspersed training phase, nine of which were preceded by the interspersed training (Sessions 24, 26, and 30 were not preceded by a training session).

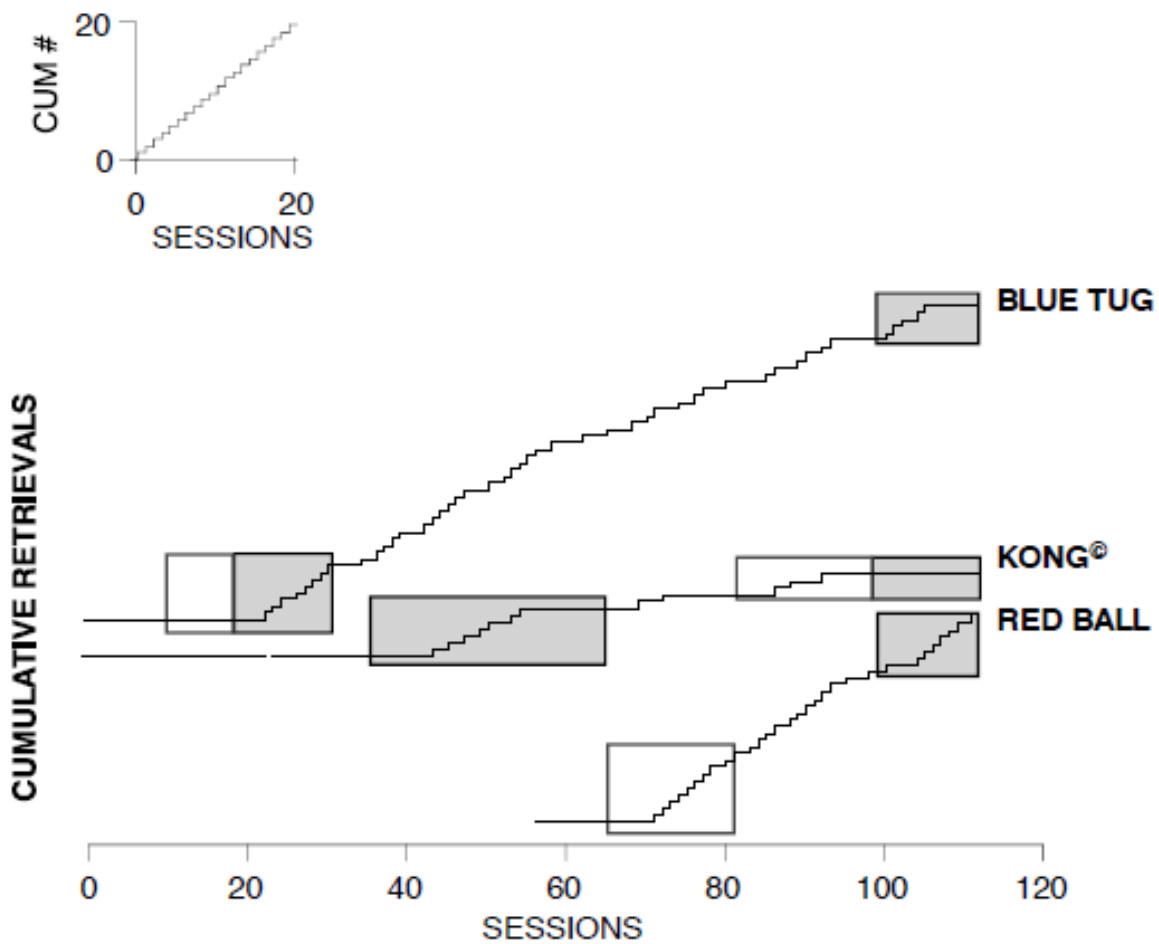
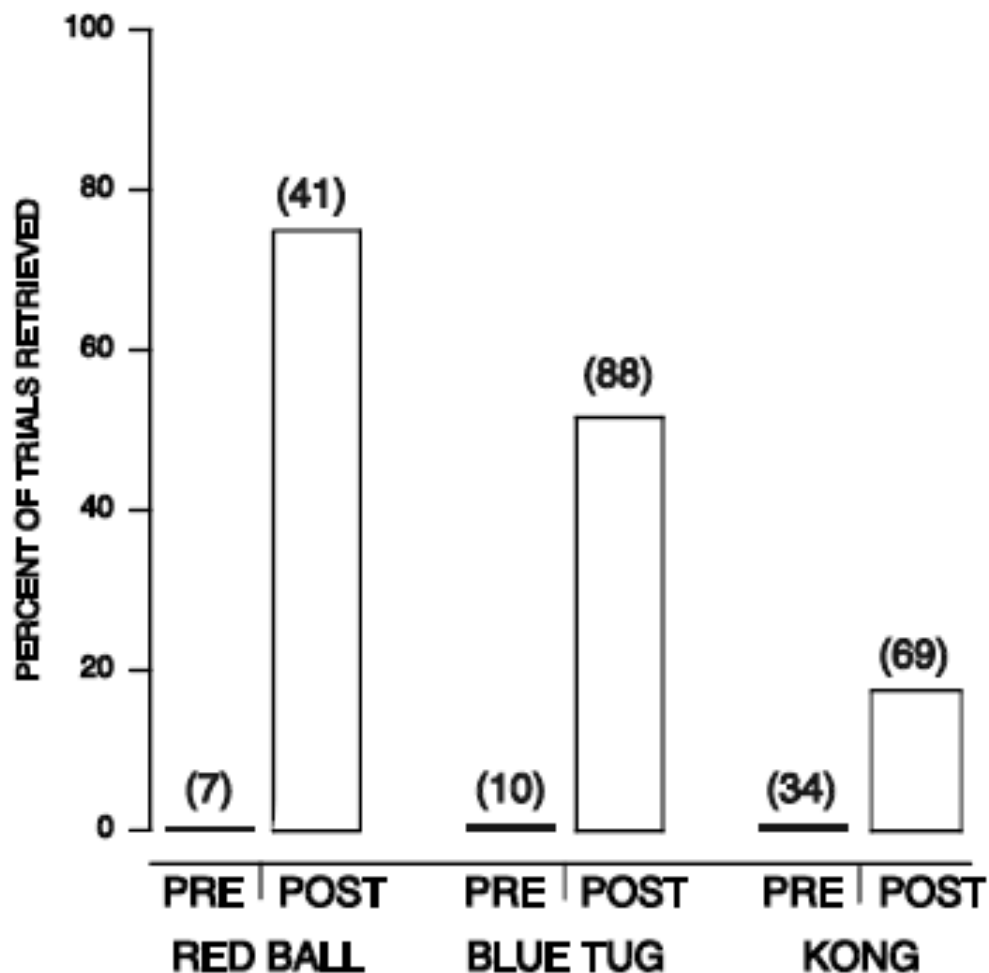


Figure 4. Cumulative record of retrievals of “new toys” (Experiment 2). The trials in which tug-of-war training was conducted prior to the retrieval test are demarcated by white squares. Trials in which interspersed training was conducted prior to the retrieval test are demarcated by shaded squares.



*Figure 5.* Percentage of retrievals for new toys (Experiment 2), showing both the percentage of trials in which Aero retrieved the object prior to common response training (Pre) and the percentage of trials in which Aero retrieved the object beginning with, and including, the session in which Aero first retrieved the object after initiation of common response training (Post).

After four interspersed training sessions (Session 22), Aero first touched the blue tug with his mouth and during Session 23 Aero first retrieved the blue tug. Including the session in which Aero first retrieved the blue tug, he retrieved the tug in seven out of the remaining eight sessions in the common response training phase. Training for the blue tug stopped after Session 30. After this, Aero continued to retrieve the blue tug 34 times in the following 69 trials, even though the blue tug was no longer being trained in the common response training. The rate of retrieval of the blue tug declined with the time since training had stopped. Interspersed training was resumed in Session 100, when all three objects were trained. Aero retrieved the blue tug an additional five times once that training resumed. Thus, Aero retrieved the blue tug a total of 46 times in 90 trials from the trial in which he first retrieved the blue tug.

Using the structured criteria method (Saini et al., 2017), 60 data points consistent with the predicted change (retrieval) would be required to demonstrate an effect (111 total trials). The blue tug had 10 nonretrievals in baseline plus 27 retrievals after common response training commenced; thus the 37 responses in the direction predicted did not meet the criterion level. As we indicated above, the retrieval rate decreased with time from training. This is borne out in an exploratory analysis in which the blue tug was closer to passing this analysis if we had ended data collection for the blue tug when we added the red ball and analyzed the responses to that point (37 responses in the direction predicted out of 65 possible trials).

**Kong.** After 34 baseline sessions in which Aero never retrieved the Kong, the experimenter conducted 29 interspersed training sessions, each of which was followed by a retrieval test session. After seven training sessions, Aero first put his mouth on the Kong (Session 42). After nine training sessions Aero first retrieved the Kong (Session 44). Including the session in which Aero first retrieved the Kong, he retrieved the Kong in seven out of the remaining 21 sessions in the common response training phase.

Overall, however, the rate of retrieval was inconsistent: of those 21 sessions, all seven retrievals occurred in the first 12 sessions, and no retrievals occurred in the last nine. Training was discontinued after Session 64 and Aero only retrieved the Kong twice after training ended. Because of this low retrieval rate, we implemented a second training phase for Kong in Session 82. Beginning with this session, the experimenter conducted 18 additional tug-of-war training sessions with the Kong, each of which was followed by a retrieval test session. During these sessions, Aero retrieved the Kong three times, in Sessions 87, 89 and 93. At Session 100, interspersed training was reinstated for all objects but did not result in any additional Kong retrievals. Thus, beginning with the first session in which Aero retrieved the Kong in the earlier training phase, Aero retrieved the Kong a total of 12 times out of 69 sessions.

Using the structured criteria method (Saini et al., 2017), 60 data points consistent with the predicted change (retrieval) would be required to demonstrate an effect (111 total trials). The Kong had 34 nonretrievals in baseline plus 12 retrievals after common response training commenced; thus the 46 responses in the direction predicted did not meet the criterion level.

**Red ball.** The red ball was first introduced during Session 58. After seven baseline sessions in which Aero never retrieved the red ball, the experimenter began conducting tug-of-war training for 17 sessions. After six sessions of training (Session 70), Aero first put his mouth on the red ball. After eight sessions of training (Session 72), he first retrieved the red ball. Including the first retrieval session, he retrieved the red ball in nine out of the remaining 10 sessions in the common response training phase. Training was discontinued after session 81. After training was discontinued, Aero continued to retrieve the red ball in 13 out of 18 sessions. When interspersed training was reinstated in Session 100 for all three objects, Aero retrieved the ball eight more times. Thus, beginning with the first session in which Aero retrieved the red ball in the earlier training phase, Aero retrieved the red ball a total of 30 times out of 41 sessions.

Using the structured criteria method (Saini et al., 2017), 33 data points consistent with the predicted change (retrieval) would be required to demonstrate an effect (56 total trials). The red ball had 7 nonretrievals in baseline plus 30 retrievals after common response training commenced; thus the 37 responses in the direction predicted met the criterion level.

## Discussion

The results indicated that common response training in one context was sufficient for Aero to begin retrieving the objects in the retrieval test context. We identified the three objects that received the common-response training and that he began to retrieve as “new toys.” The common response (playing tug-of-war) was one that Aero emitted in his daily life, suggesting that this might be what united the toy class identified in Experiment 1.

Importantly, common response training took place in a different context than testing and the cue used in testing (“go get it” and a finger point) was never used when playing with the three objects. Additionally, all test trials were reinforced, precluding the possibility that differential reinforcement resulted in Aero retrieving the items or forming a concept. In fact, the change in Aero’s behavior to the three objects runs counter to that predicted by his history of reinforcement with respect to those objects in the testing situation. That is, when the experimenter started the common response training for the blue tug, Aero had already been reinforced for 10 trials for not retrieving or ignoring it. Similarly, he had been reinforced for ignoring the Kong in 34 trials and the red ball in seven trials at the point at which common response training started for each of the objects. Additionally, we included the Frisbee as a control. Throughout all the experiments, Aero never retrieved the Frisbee, demonstrating that it was not prolonged exposure to the three target objects that resulted in him beginning to retrieve the objects.

We employed two types of common-response training (tug-of-war and interspersed). Although there were small variations in either how the object was presented or whether other toys were played with in the same training session, the same general response topography (all involved grabbing an object and playing tug-of-war) was similar. Across target objects, we found a similar number of common-response training sessions to first retrieval. Thus, the number of sessions in which the animal engages in the common response with the stimulus might be a critical variable for the object to acquire class membership, rather than the type of training (tug of-war or interspersed).

Although the results showed that the common response of playing was sufficient for initiating the retrieval of the three target objects, the rates of retrieval of the three objects differed: the red ball had the highest overall rate, and its rate approached that of the original toys; the Kong had the lowest overall rate; and the rate of retrieval for the blue tug fell between the first two objects. The variation in rates of retrieval indicates that factors other than the common response training might be responsible for the maintenance of that object in the class. For example, the new member might have to share the same contingencies as other members of that class in more than one context.

Three possible factors within the current experiment might have also contributed to the variation in retrieval rates. One possible factor is the interference from the prior history of reinforcement, as evidenced by the length of baseline for each object. The rates of retrieval of the three objects were inversely related to the length of baseline for each object. For example, the 34 session baseline of the Kong was nearly five times longer than that for the red ball, and the Kong showed the lowest rate of retrieval. This inverse relationship suggests that the non-retrieval responses reinforced during baseline interfered with the retrieval performance.

Another possible factor is that, although Aero’s tug-of-war response in the common response training sessions looked the same to the human eye for the three objects, the stimulation associated with each might have varied. The Kong, especially, was made of thicker, less pliable rubber so that Aero was not able to close his mouth as much with the Kong as with the red ball and blue tug.

A third possible factor is that longer and more frequent common-response training might be required. This is particularly evident in the blue tug, which had the longest time since training and the rate of retrieval declined with the time elapsed since training. Further support for this comes from the brief resumption of Kong retrievals when a second training phase for the Kong was conducted, as well as an increase in red ball and blue tug retrievals when these objects received additional training in the multiple interspersed training phase at the end. The effects of these potential variables on the rate of retrieval, and the persistence of retrieval in the absence of common-response training should be explored in future research.

### Experiment 3: Transfer of Function Test

To test whether Aero's differential responding to the two sets of objects (toys and non toys) constituted a concept, we conducted a transfer of function test. In this test, we trained Aero to emit a new response (nose touch) to one member of the toy class and not to one of the non-toy objects. We then tested if he would then differentially emit nose touches to other members of the toy class and not to the non-toy objects.

#### Method

**Pre-training.** To train the nose touch response used in the transfer test, we started by training Aero to nose touch a green piece of paper held in the experimenter's hand. During each correct response the experimenter clicked and delivered a treat to Aero. Once Aero was reliably touching the green paper in her hand, she taped the green paper to a monkey knot rope toy that Aero regularly played with but was not part of the study. Gradually the experimenter lowered the object with the green paper onto a small table (approximately 45 cm high x 85 cm long x 50 cm wide) and then faded her hand away from the object, until Aero was nose touching the object consistently without the experimenter's hand nearby. Once the object was on the table Aero had to push the object with the tip of his nose so that he displaced it by at least 0.6 cm. The experimenter then removed the target and Aero continued to respond correctly. Beginning with the trials in which the experimenter faded her hand away from the object, she also began alternating onto which end of the table she placed the object.

At this point the experimenter implemented the same training procedure, but with the black dumbbell (toy from Experiment 1). Once Aero was reliably touching the black dumbbell without the green paper or the experimenter's hand as prompts, the experimenter began discrimination training trials.

**Discrimination training.** The table was oriented lengthwise in front of the experimenter with the experimenter seated at the midpoint of the long side of the table, indicated by a small yellow dot placed on the table. Aero was on the other side of the table. The experimenter had also pre-marked the table with two blue dots on which she would place the two objects (black dumbbell and blue geometric). The dots were each 10 cm from the side of the table nearest Aero, and 16.5 cm away from the end of the long side of the table, so that the objects' midpoints were always 54 cm apart, although the distance between the edges of the two objects varied with the objects' size.

A nose touch was defined as the tip of Aero's muzzle (nose, lips, or chin) coming into contact with the object. The most common topography of the response was for Aero to touch the object with the extreme tip of his nose.

To begin a session, the experimenter tossed a treat approximately 3 m behind Aero. As Aero moved away from the experimenter she placed the two objects on the two blue dots and returned her hands to her lap, which held treats and a clicker. After finding the treat Aero returned to the table and made a nose touch to one of the objects. As Aero returned to the table, the experimenter remained motionless, staring straight ahead at the back wall of the room. If he emitted a nose touch to the toy the experimenter clicked and tossed a treat again approximately 3 m behind Aero to reset him for the next trial. Then the experimenter reset the objects on the table and another trial occurred.

If Aero made an incorrect response by touching the non-toy, the experimenter removed both objects and remained motionless for 2 s. She then began a correction trial; she tossed a treat to reset Aero and placed the objects in the same position they were in when

he made the error. If he made a correct response on this correction trial the experimenter clicked and tossed a treat and moved on in the trial sequence. If he made another error the same correction procedure was implemented again.

Sessions consisted of 20 trials and we counterbalanced which side the toy was placed on across trials. We pseudo-randomized the placement of the toy following Fellows (1967) with the added constraint that the toy was not placed on the same side in more than two consecutive trials. Training sessions continued until Aero met the criterion of responding at or above 90% correct for five consecutive sessions. Any correction trials were omitted from the calculation of percent correct in each session. After Aero met the response criterion, we began conducting transfer test sessions. We began pre-training Aero after Session 108 of the retrieval tests. We completed the pre-training and all but one session of discrimination training before resuming retrieval tests in Session 109. The same day we conducted Session 109 of the retrieval tests, we conducted the final discrimination training session (see below) before moving to the transfer test.

**Transfer test.** In the transfer test, we investigated whether Aero would transfer the nose touch response trained in discrimination training to other toys, including the three new toys we had trained in Experiment 2. The general procedures for the transfer test sessions were the same as those for the discrimination training sessions, the only difference being the stimuli presented. The transfer test consisted of presenting pairs of objects in which one member of the pair was a toy or new toy (one of the objects that received common response training 626 in Experiment 2), and the other member of the pair was always a non-toy.

We conducted a total of five transfer test sessions. Only one transfer test session was conducted per day. Sessions 1-4 each consisted of four blocks of trials. In Sessions 1 and 2, each block consisted of five trials. Each pair of objects was tested once per block (see Table 2 for object pairs). In addition to the pair of objects used in the discrimination training (black dumbbell and blue geometric), three new toys (Experiment 2) and one toy (Experiment 1) were each paired with the three non-toys not used in the discrimination training. Because we had only three non-toys (other than the blue geometric used in discrimination training) the football was used in two pairings as the non-toy. In each block, the first trial was the pair of trained objects (black dumbbell and blue geometric). The remaining trials were the pairs of toys/non-toys or new toys/non-toys. The sequence of pair presentations in these trials varied between trial blocks.

We conducted Sessions 3 and 4 to test whether Aero would emit nose touches to other toys (showing a transfer of function) as Sessions 1 and 2 only included one toy that was not used in discrimination training. Each session consisted of four blocks of four trials per block. Each pair of objects was presented once per block within the same parameters as Sessions 1 and 2. Along with the pair of objects used in the discrimination training, three toys were paired with three non-toys (Table 2).

Finally, in Session 5, we tested all four toys and all three non-toys used in the previous sessions in one complete session. The four toys (black dumbbell, plush geometric, alligator and rope rings) and the three new toys (blue tug, Kong, and red ball) were each presented once with the Frisbee (non-toy) and once with the squirrel (non-toy). We eliminated the blue geometric (non-toy) from the test pairs as Aero could have responded correctly based on exclusion alone, as it was the negative stimulus during discrimination training. The four toys (black dumbbell, plush geometric, alligator and rope rings) and the three new toys (blue tug, Kong, and red ball) were each presented once with the Frisbee and once with the squirrel. Thus, each trial in the 14-trial session involved a unique pair of objects.

All sessions were counterbalanced for the side of toy or new toy presentation. Sessions were also constructed so that no more than two consecutive trials had the toy or new toy on the same side. Prior to each transfer test session, the experimenter conducted a discrimination training session as detailed above.

Table 2  
*Frequency Table of Retrievals and Nonretrievals*

Object Name	Retrievals	Nonretrievals		
<b>Toys</b>				
Black dumbbell	111	0		
Plush alligator	111	0		
Rope rings	106	5		
Plush geometric	104	7		
Lambswool knot	54	1		
Plush bear	53	1		
<b>Non-Toys</b>				
Frisbee	0	111		
Blue geometric	0	52		
Football	0	49		
Squirrel	0	34		
<b>New Toys</b>				
	Retrievals (Pre)	Nonretrievals (Pre)	Retrievals (Post)	Nonretrievals (Post)
Blue tug	0	22	46	44
Kong	0	34	12	65
Red ball	0	7	30	19

**Retrieval testing.** We continued to conduct retrieval tests during the time of transfer of function testing. During one day of discrimination training, and three days of the transfer test sessions (transfer test Sessions 1, 2, and 5) the experimenter also conducted retrieval sessions (Sessions 109-112, Figures 2 and 4) to assess which objects were part of the concept “toy” during the transfer test. There were two days in which we only conducted the transfer test session preceded as usual by discrimination training, but we did not conduct a retrieval test. The retrieval sessions were conducted just as they were in Experiments 1 and 2. Two of these retrieval sessions (Sessions 109 and 111) were also preceded by a common response training session (multiple interspersed training, see Methods, Experiment 2).

**Daily sequence of sessions.** On the day that a retrieval test and a transfer test were both conducted, we conducted the retrieval test first, followed by discrimination training, and finally the transfer test. On the days that a common response training, a retrieval test and a transfer test were all conducted, we conducted the common response training first, then the retrieval test, followed by discrimination training, and finally the transfer test.

Table 3  
*Transfer Test Pairs*

Transfer Test Sessions 1 and 2		
Pair	Toy	Non-Toy
1 (trained pair)	Black dumbbell	Blue geometric
2	Rope rings	Football
<u>Pair</u>	<u>New Toy</u>	<u>Non-Toy</u>
3	Blue tug	Frisbee
4	Kong	Football
5	Red ball	Squirrel
Transfer Test Sessions 3 and 4		
1 (trained pair)	Black dumbbell	Blue geometric
2	Rope rings	Football
3	Plush geometric	Frisbee
4	Alligator	Squirrel
Transfer Test Session 5		
1 (trained pair)	Black dumbbell	Frisbee
2	Black dumbbell	Squirrel
3	Plush geometric	Frisbee
4	Plush geometric	Squirrel
5	Alligator	Frisbee
6	Alligator	Squirrel
7	Rope rings	Frisbee
8	Rope rings	Squirrel
<u>Pair</u>	<u>New Toy</u>	<u>Non-Toy</u>
9	Blue tug	Frisbee
10	Blue tug	Squirrel
11	Kong	Frisbee
12	Kong	Squirrel
13	Red ball	Frisbee
14	Red ball	Squirrel

## Results

**Transfer test.** Figure 6 shows total percent of nose touches to the various objects in the transfer test across all five test sessions, along with retrieval data from Experiment 1 for comparison. Aero emitted a significantly greater number of nose touches to all but one of the toys when compared to the paired non-toy (binomial test): 16 nose touches out of 18 trials to the dumbbell,  $p < 0.01$ , 95% *CI* [1.0, 0.73]; 13 nose touches out of 18 trials to the rope rings,  $p < 0.05$ , 95% *CI* [0.58, 0.96]; and nine nose touches out of 10 trials to the plush geometric,  $p < 0.01$ . The only toy that did not meet the criterion of the binomial test was the alligator. Aero emitted eight nose touches out of 10 trials, such that his response to the alligator trended in the predicted direction but was not statistically significant,  $p = 0.054$ , 95% *CI* [1.0, 0.65]. Towards the three new toys, Aero emitted a significantly greater number of nose touches to the red ball than the paired non-toy: nine nose touches out of 10 trials to the red ball, binomial test,  $p < 0.01$ , 95% *CI* [1.0, 0.72]. However, he did not emit nose

touches to the other new toys at a level above chance: four nose touches out of 10 trials to the blue tug, and three nose touches out of 10 trials to the Kong.

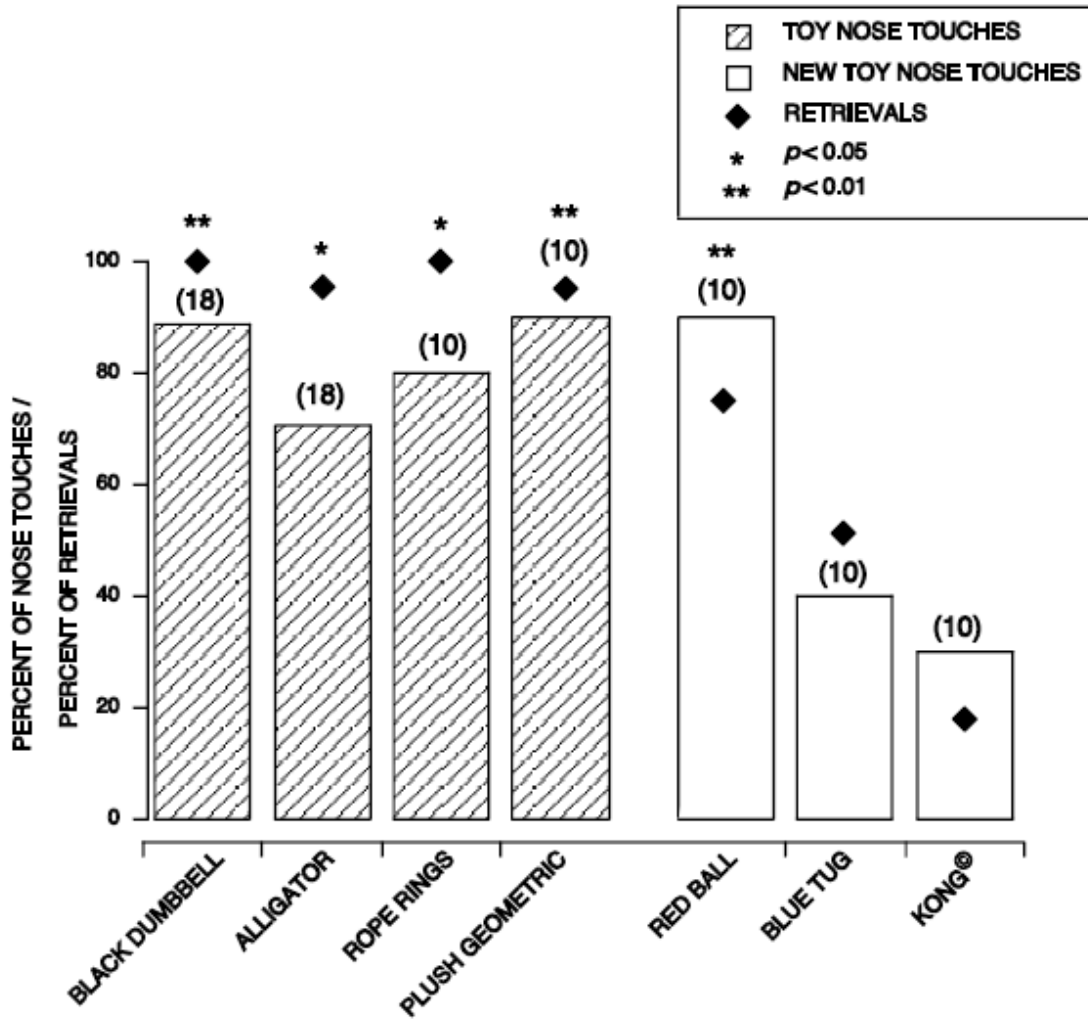


Figure 6. Percentage of trials in which Aero emitted a nose touch to the toy or new toy in the transfer test sessions (bars) (Experiment 3). The percentage of trials in which the object was retrieved in all retrieval test sessions (toys) or from the first retrieval (non-toys) (diamonds). \*  $p < .05$ , \*\*  $p < .01$ .

**Retrieval sessions.** In the four retrieval sessions (Sessions 109-112) conducted during discrimination training and the transfer test, Aero continued to retrieve all of the original toys and to not retrieve the non-toys (Figure 2). Aero retrieved the red ball in three out of the four sessions, but did not retrieve the blue tug or the Kong in any of these four sessions (Figure 4).

## Discussion

Aero's performance in the transfer test was consistent with a having formed a concept of toys. He reliably emitted nose touches to all but one of the toys when these were presented with non-toys, and the one toy that did not meet statistical significance trended in the correct direction (eight out of 10 nose touches) This demonstrated he was sensitive to the change in contingency of reinforcement for these objects, the criterion that Herrnstein (1990) outlined for showing concept formation. That is, Aero emitted a new response to untrained stimuli that was consistent with his patterns of retrieval. The results suggest that the differential responding we identified in Experiment 1 was in accordance with a concept.

Aero's percentage of nose touches emitted to the three new toys exactly paralleled his percentage of trials in which he retrieved the object, starting from the trial in which he first retrieved that object (Figure 6). That is, the red ball had the highest percentage of retrievals, which nearly approximated the rate of retrieval of the original toys, and the highest percentage of nose touches, followed by the blue tug, which had intermediate levels of both retrievals and nose touches, and finally the Kong, which had the lowest rate of retrieval and the lowest percentage of nose touches. The red ball was the only one of the three new toys that Aero retrieved in the retrieval tests that were conducted during the same time period as the transfer of function tests. This further suggests that the rate of retrieval is a valid indicator of class membership and that while the red ball seems to have been successfully added to the toy class, the Kong and the blue tug were not. As noted in the Discussion section for Experiment 2, there are several possibilities that might explain the success in adding the red ball but not the other objects. It would have been interesting to conduct transfer of function tests when the rates of retrieval for the Kong and the blue tug were high to investigate whether they were, a tone point, part of the toy class, but for some reason became excluded again.

We performed this test with the experimenter in sight of Aero. The experimenter always stared straight ahead to the wall behind Aero. While this does not preclude that there were other subtle cues that the dog used to choose correctly, given dogs' abilities to follow overt human gestures that are not hand or foot points, (e.g., elbow points, shoulder turns, or head turns) we do not think this is very likely (see Udell, Giglio, & Wynne, 2008). Additionally, the dog did not perform well on two out of the three new toys, which we might have expected him to perform well on if the experimenter was providing inadvertent cueing. To further reduce any concerns of inadvertent experimenter-cued responding, future studies could try to completely occlude the experimenter from the dog.

## General Discussion

Previous research has shown that dogs can experimentally form stimulus classes such as the presence of dogs in pictures (Range et al., 2008) and Chaser's performance demonstrated that such stimulus classes could be formed in the natural environment. Aero's performance expands on this. Our results suggest that not only can dogs form stimulus classes, they can form concepts, assessed even by the strictest criteria.

Aero's performance is similar to Chaser's performance with "toys" (Pilley & Reid, 2011), both having emerged from natural day-to-day contingencies. Herrnstein and Loveland (1964) also suggested that, based on the speed with which their pigeons showed discriminated responding to pictures of humans and pictures without humans, their pigeons likely already had formed the concept "human" from their daily contingencies. While Chaser's history with the toys compared to non-toys (allowed to play with toys and not with non-toys) suggested a contingency of reinforcement that united the toys and separated them from 738 non-toys, our results suggest that a common contingency can be used to add members to the concept of toys for Aero. These

results would support the suggestion by Pilley and Reid that a common response to toys created Chaser's concept of toys.

Our results also support Sidman's formulation of equivalence relations (Sidman, 2001). That is, stimuli can enter into a stimulus class (e.g., toys) if a common response is evoked by the various stimuli. To date, while experimental procedures, such as match-to-sample procedures (e.g., Galizio et al., 2004) and rapidly reversing contingencies on whole groups of stimuli (e.g., Vaughan, 1998), have provided ways to train new stimulus classes, these procedures seem unlikely to occur in the natural environment. In the case of Aero and the formation of a concept of toys, Aero played tug-of-war with many objects throughout his life and this common response could be the mechanism that established the concept initially as well as allowed new members to be added.

The concept of toys is particularly intriguing, as members appear to share no overarching physical features that can be used to discriminate between a toy and non-toy. Premack (1983) doubted that animals had the ability to discriminate based on the function of the object, stating, "...pigeons have never been shown to have functional classes—furniture, toys, candy, sports equipment—where class members do not look alike; they only recognize physical classes" (p. 358). Nevertheless, Aero's performance parallels the ability of Chaser to discriminate between "toys" and "non-toys" based on past functions of those objects for the dog (Pilley & Reid, 2011), a performance corresponding to the associative concept of Zentall et al. (2002). However, we expanded on the results of Pilley and Reid to demonstrate that Aero's performance meets the more stringent criterion of a transfer of function test required by some definitions of concept formation (Herrnstein, 1990; Lea 1984). It is possible that Chaser, if tested, would also show the requisite transfer of function to call the class "toys" a concept.

Still, we did not catalogue or measure the physical features of the objects used in this study, such as color, texture, odor, or shape. It remains a possibility, then, that there is some subtle physical feature that we did not notice between the toys and non-toys that controlled Aero's differential responding to the two sets of stimuli, and could possibly account for why the red ball seemed to be added to the concept and not the blue tug or Kong. One possible difference that accounted for some, but not all differential responding was that plush or rope toys were all members of the toy group, whereas plastic or rubber were more likely to fall in the non-toy group, although there were exceptions (the black rubber dumbbell was a retrieved item). A fruitful line of future research would be to determine which stimulus features dogs predominantly use to determine whether an object is a toy or not, and see if we could still add objects to the class that do not have that feature through common response training. A good dimension to start with might be non-rubber/non-plastic versus rubber/plastic.

Aero's performance also parallels the complex performance of some primates including a juvenile gorilla (Vonk & McDonald, 2002), which was able to differentiate sets of stimuli based on abstract discriminations (e.g., animals versus non-animals and food versus animals). When presented with novel sets of stimuli, the gorilla continued to correctly respond to the different categories. While the authors demonstrated transfer to new sets of stimuli, a transfer of function test was not conducted. Thus, whether the gorilla's performance would have met the criteria of concept formation remains to be tested.

As dogs are employed in more complex tasks that require higher cognitive abilities, such as concept formation, understanding how dogs form concepts and how we can train them is essential. For example, dogs that detect explosives or drugs that vary in exact content or concentration might be more efficiently trained by using their abilities to form concepts rather than training them on all possible exemplars. If common response

or common reinforcer training can enhance concept development, employing it with working dogs could enhance their detection performance.

Our results add to our growing understanding of complex behavioral repertoires in animals and specifically dogs, including putative fast-mapping in a border collie (Kaminski et al., 2004), a similar performance in a Yorkshire terrier (Griebel & Oller, 2012), dog visual stimulus classes (Range et al., 2008), and object names as verbal referents (Pilley & Reid, 2011), by demonstrating concept formation by the strictest criteria. Additionally, because Aero, like Chaser (Pilley & Reid, 2011), entered the experiment already differentially responding to two sets of stimuli, dogs and other animals likely form these complex discriminations in their day-to-day living and we have identified an ecologically-relevant mechanism for concept formation, which will hopefully be further explored for its generality in other animals.

Conspicuously, the current study and many of the studies cited that investigated dogs (e.g., Griebel & Oller, 2012; Kaminski et al., 2004; Pilley & Reid, 2011) have had only one subject. Thus, the generality of our findings and the findings of other studies showing complex cognitive capabilities in dogs needs to be explored by testing more dogs with diverse backgrounds. Nevertheless, these studies demonstrate the ability for at least some domestic dogs to engage in complex cognition. Our study demonstrated a viable mechanism for adding members to a concept; future research should also attempt to investigate whether this mechanism can be used to construct a novel concept in domestic dogs.

## References

- Bond, A. B., Wei, C. A., & Kamil, A. C. (2010). Cognitive representation in transitive inference: a comparison of four corvid species. *Behavioural Processes*, *85*, 283-292. doi: 10.1016/j.beproc.2010.08.003
- D'Amato, M. R., & Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 43-55. <http://dx.doi.org/10.1037/0097-7403.14.1.43>
- Dickins, T. E., & Dickins, D. W. (2001). Symbols, stimulus equivalence, and the origins of language. *Behavior and Philosophy*, *29*, 221-244.
- Fellows, B. J. (1967). Chance stimulus sequences for discrimination tasks. *Psychological Bulletin*, *67*, 87-92. <http://dx.doi.org/10.1037/h0024098>
- Fields, L., Landon-Jimenez, D. V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior*, *64*, 129-145. doi: 10.1901/jeab.1995.64-129
- Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis*, *36*, 387-406. doi: 10.1901/jaba.2003.36-387
- Galizio, M., Stewart, K. L., & Pilgrim, C. (2004). Typicality effects in contingency-shaped generalized equivalence classes. *Journal of the Experimental Analysis of Behavior*, *82*, 253-273. doi: 10.1901/jeab.2004.82-253
- Goldiamond, I. (1966). Perception, language, and conceptualization rules. In B. Kleinmuntz (Ed.), *Problem solving*. New York, NY: John Wiley & Sons.
- Griebel, U., & Oller, D. K. (2012). Vocabulary learning in a Yorkshire terrier: Slow mapping of spoken words. *PLoS ONE*, *7*, e30182. doi: 10.1371/journal.pone.0030182
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, *298*, 1634-1636. doi: 10.1126/science.1072702
- Herrnstein, R. J. (1990). Levels of stimulus control: A functional analysis. *Cognition*, *37*, 133-166. doi:10.1016/0010-0277(90)90021-B
- Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, *146*, 549-551. doi: 10.1126/science.146.3643.549

- Huber, L. (2001). Visual categorization in pigeons. In R. G. Cook (Ed.), *Avian visual cognition* [On-line]. Available: [www.pigeon.psy.tufts.edu/avc/huber/](http://www.pigeon.psy.tufts.edu/avc/huber/)
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping." *Science*, *304*, 1682–1683. doi: 10.1126/science.1097859
- Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior*, *76*, 131–158. doi: 10.1901/jeab.2001.76-131
- Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology*. New York, NY: Applebaum-Century-Crofts, Inc.
- Kirchhofer, K. C., Zimmermann, F., Kaminski, J., & Tomasello, M. (2012). Dogs (*Canis familiaris*), but not chimpanzees (*Pan troglodytes*), understand imperative pointing. *PloS One*, *7*, e30913. <http://dx.doi.org/10.1371/journal.pone.0030913>
- Lea, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 263-276). Hillsdale, NJ: Erlbaum.
- Pilley, J. W., & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, *86*, 184–195. doi: 10.1016/j.beproc.2010.11.007
- Premack, D. (1983). Animal cognition. *Annual Review of Psychology*, *34*, 351–362. doi: 10.1146/annurev.ps.34.020183.002031
- Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs. *Animal Cognition*, *11*, 339–347. doi: 10.1007/s10071-007-0123-2
- Saini, V., Fisher, W. W., & Retzlaff, B. J. (2017). Efficiency in functional analysis of problem behavior: A review. Manuscript submitted for publication.
- Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalophus californianus*) is capable of forming equivalence relations. *Psychological Record*, *43*, 823–839.
- Schusterman, R. J., & Kastak, D. (1998). Functional equivalence in a California sea lion: Relevance to animal social and communicative interactions. *Animal Behaviour*, *55*, 1087–1095.
- Schusterman, R. J., Reichmuth, C. J., Kastak, D. (2002). How animals classify friends and foes. *Current Directions in Psychological Science*, *9*, 1–6. doi:10.1111/1467-8721.00047
- Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, *74*, 127–146. doi: 10.1901/jeab.2000.74-127
- Udell, M. A., Giglio, R. F., & Wynne, C. D. (2008). Domestic dogs (*Canis familiaris*) use human gestures but not nonhuman tokens to find hidden food. *Journal of Comparative Psychology*, *122*, 84–93. <http://dx.doi.org/10.1037/0735-7036.122.1.84>
- Urcuioli, P. (2001). Categorization & acquired equivalence. In R. G. Cook (Ed.) *Avian visual cognition* [On-line]. Available: [www.pigeon.psy.tufts.edu/avc/urcuioli/](http://www.pigeon.psy.tufts.edu/avc/urcuioli/)
- Vaughan, W. Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 36–42. <http://dx.doi.org/10.1037/0097-7403.14.1.36>
- Vonk, J., & McDonald, S. E. (2002). Natural concepts in a juvenile gorilla (*Gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of Behavior*, *78*, 315–332. doi: 10.1901/jeab.2002.78-315
- Wang, T., Dack, C., McHugh, L., & Whelan, R. (2011). Preserved nodal number effects under equal reinforcement. *Learning and Behavior*, *39*, 224–238. doi:10.3758/s13420-011-0020-z
- Wasserman, E. A., DeVolder, C. L., & Coppage, D. J. (1992). Non similarity-based conceptualization in pigeons via secondary or mediated generalization. *Psychological Science*, *3*, 374–378. doi:10.1111/j.1467-9280.1992.tb00050.x
- Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: an introduction. *Journal of the Experimental Analysis of Behavior*, *78*, 273–248. doi:10.1901/jeab.2002.78-237
- Zentall, T. R., Wasserman, E. A., Lazareva, O. F., Thompson, R. K., & Rattermann, M. J. (2008). Concept learning in animals. *Comparative Cognition & Behavior Reviews*, *3*, 13-45. doi:10.3819/ccbr.2008.30002

**Financial conflict of interest:** No stated conflicts.  
**Conflict of interest:** No stated conflicts.

*Submitted: September 5<sup>th</sup>, 2016*

*Resubmitted: January 13<sup>th</sup>, 2017*

*Accepted: March 14<sup>th</sup>, 2017*