Animal Behavior and Cognition

ABC 2024, 11(2):167-197
DOI: https://doi.org/10.26451/abc.11.02.03.2024

Exploring Dogs' Capacity to Innovate Actions on Cue

Allison Scagel*, Anna Sofia Hege, and Eduardo Mercado III

Department of Psychology, University at Buffalo, The State University of New York

*Corresponding author (Email: ascagel@buffalo.edu)

Citation – Scagel, A., Hege, A. S., Mercado III, E. (2024). Exploring dogs' capacity to innovate actions on cue. *Animal Behavior and Cognition, 11*(2), 167-197. https://doi.org/10.26451/abc.11.02.03.2024

Abstract – Creative animals can flexibly respond to novel circumstances, a capacity that is closely tied to cognitive complexity. We examined the capacity of four dogs to learn to innovate actions when instructed to do so using a "create" cue. An action performed in response to this cue was considered novel if the action had not yet been offered within the current training session and was not already associated with a specific cue. The dogs were able to self-select many novel actions, similarly to what has been observed in trained dolphins. Training animals to innovate actions on cue can help reveal aspects of the cognitive processes underlying action selection, cognitive control, and behavioral flexibility. Successful learning of the "create" cue may indicate that the dogs were able to monitor their recent actions and form an abstract concept of avoiding recently performed actions. Reinforcing dogs for spontaneity may increase the range of actions that they perform within the training context. We conclude that the dogs in this study were capable of learning to innovate actions in ways that are comparable to what has been shown in dolphins. Though it is unclear how other dogs would perform at this task, our results demonstrate it is possible for dogs to learn to innovate actions on cue.

Keywords – Event memory, Creativity, Inventiveness, Originality

Creativity, or the capacity to produce something novel, helps individuals solve problems in their environment by facilitating the discovery of solutions when previously used strategies fail (Kaufman & Beghetto, 2009; Kaufman et al., 2011; Kuczaj, 2017; Schusterman & Reichmuth, 2008; Yeater et al., 2024). Behavioral innovation (self-selection of an action that is in some way new) can occur when animals play (Kuczaj & Eskelinen, 2014a; Yeater et al., 2024), use tools to obtain food (Berthelet & Chavaillon, 1993; Sanz et al., 2014; Shumaker et al., 2011), or discover solutions to novel problems (Köhler, 1926; Yeater et al., 2024). As many animals are neophobic, behavioral innovation may be limited to situations that demand exploring new strategies and may be linked to bold personalities or risk-taking (Kuczaj, 2017; Brosnan & Hopper, 2014; Burghardt, 2015). Creativity and behavioral innovation can sometimes be facilitated by play, leading to new actions or combinations of behaviors that may or may not be functionally useful (Burghardt, 2015; Burghardt, 2005; Fagen, 1981).

Observations of animals innovating actions feature prominently in discussions of animal cognition. For example, Kummer and Goodall (2003) described an instance of a male chimpanzee (*Pan troglodytes*) that banged cans together during play, leading to the discovery that the banging noise intimidated other chimpanzees in the group. The chimpanzee then used the banging behavior outside of the play context to gain status within the social hierarchy and to become the dominant male. Among a group of killer whales (*Orcinus orca*) observed capturing seagulls that came to the surface of the water to feed on fish, one whale appeared to innovate by leaping several feet out of the water to catch gulls in mid-air rather than continuing to wait for the gulls to come to the surface (Kuczaj & Walker, 2012). Such episodes, while provocative, are often viewed skeptically by many scientists, following the examples of Lloyd Morgan and Edward

Thorndike. However, it is exactly these kinds of novel action sequences that provide the most compelling evidence of creativity in experimental studies of behavioral innovation.

The processes underlying behavioral innovation are largely unknown (Burghardt, 2015). Animals can be taught, however, to produce novel behaviors through shaping, a process that involves reinforcing successive approximations of a desired behavior (Pryor, 2000; Pryor & Chase, 2014; Pryor et al., 1969). Pryor and colleagues showed that two rough-toothed dolphins (*Steno bredanensis*) could be trained to innovate behaviors on cue by reinforcing actions the dolphins had not produced before within a given training session. By reinforcing a new behavior during successive public demonstrations of training sessions, researchers found that one dolphin began offering novel and complex behaviors on her own during each session. These actions ranged from spinning while swimming to jumping out of the water upside down. Researchers subsequently aimed to use operant conditioning to encourage response variation in pigeons (*Columba livia domestica*) and rats (*Rattus spp.*) (Balsam et al., 1998; Gutiérez & Escobar, 2022; Morris, 1989; Neuringer, 1991, 2004; Stokes, 1995), though not always successfully (Schwartz, 1982). Close examination of responses suggested that pigeons and rats were simply offering random responses rather than following a general rule of producing sequences not offered in previous trials.

The extent to which an individual can innovate actions is constrained by the individual's capacity for flexible, voluntary motor control (Mercado et al., 2014, 2022), and by the knowledge (e.g., action repertoire) that the individual has acquired through past experiences (Bailey et al., 2007; Lawrence et al., 2016). For example, the behavioral variability of pigeons and rats in past operant conditioning studies was limited in that the reinforced actions only involved key pecks or lever presses (Pryor & Chase, 2014). Within experimental contexts, creativity is also constrained by subjects' understanding of task requirements (Mercado & Scagel, 2022a), and the problem-solving strategies that different subjects employ during training and testing (Kuczaj & Eskelinen, 2014b). Some subjects may offer actions that require minimal physical effort while others may physically challenge themselves (Kuczaj & Eskelinen, 2014a, 2014b; Hill et al., 2022). The specific criteria and approaches that trainers use when attempting to shape individuals to innovate actions on cue also can strongly affect the relative creativity of an individual's performance during testing (Dudzinki et al., 2018; Lawrence et al., 2016; Hill et al., 2022), as well as observers' subjective impressions of how creative an individual's actions are (Kaufman, 2021). Animals that have learned to associate numerous actions with specific cues may gravitate toward selecting combinations of those actions when instructed to innovate (Braslau-Schneck, 1994; Hill et al., 2022; Van Steyn, 2018).

Training animals to innovate actions on cue, beyond providing enrichment, may provide insight into the cognitive processes underlying behavioral flexibility and creativity (Mercado & DeLong, 2010). Bottlenose dolphins (*Tursiops truncatus*) have been trained in several studies to either repeat a previous self-performed action, innovate an action not performed within the current session, or to perform one of five possible actions without performing any of the five actions twice consecutively (Braslau-Schneck, 1994; Cutting, 1997; Kuczaj & Eskelinen, 2014b; Lawrence et al., 2016; Mercado et al., 1998; Mercado et al., 1999; Taylor, 1995; Van Steyn, 2018). Dolphins also may innovate actions in natural contexts (Patterson & Mann, 2015). Dolphins trained to innovate actions in response to a specific gesture produced a variety of actions, performing some actions from their previously trained repertoire and other actions they had never been trained to perform (Braslau-Schneck, 1994; Kuczaj & Eskelinen, 2014b; Van Steyn, 2018). Dolphins may rely on internal representations of their own recently performed actions to avoid reproducing previous actions when innovating (Kuczaj & Eskelinen, 2014b; Mercado et al., 1998; Mercado et al., 1999; Van Steyn, 2018). A recent study of creativity in killer whales found that they were able to avoid repeating recently performed actions for sequences of more than fifteen cued innovations (Hill et al., 2022).

The current study sought to examine the capacity of domestic dogs (*Canis lupus familiaris*) to innovate actions on cue within a training session. Dogs are good candidates for the study of animal creativity and behavioral innovation because of their widespread availability and willingness to cooperate with humans (Aria et al., 2021). Through the process of domestication, dogs have developed a greater threshold for human interaction than their wild counterparts, enabling them to interact with people in ways that are not feasible for many other species (Coppinger & Coppinger, 2001). Dogs are also excellent at responding to cues provided by humans, seem to enjoy human social contact, and are often motivated to perform tasks for food rewards and social attention (Cook et al., 2016; Riedel et al., 2008; Thompson et al., 2016). Dogs are known to be able to innovate actions when reinforced for doing so (Pryor & Chase, 2014; Scagel & Mercado, 2022a; Willgohs et al., 2022), but how their creative capacities compare with those of other animals remains unclear. Here, we measured dogs' propensities for innovating actions when tested using methods closely matching those used in past experimental tests of creativity in cetaceans.

Method

Ethics Statement

All experiments were approved by the University at Buffalo (IACUC protocol #TR202100018).

Subjects

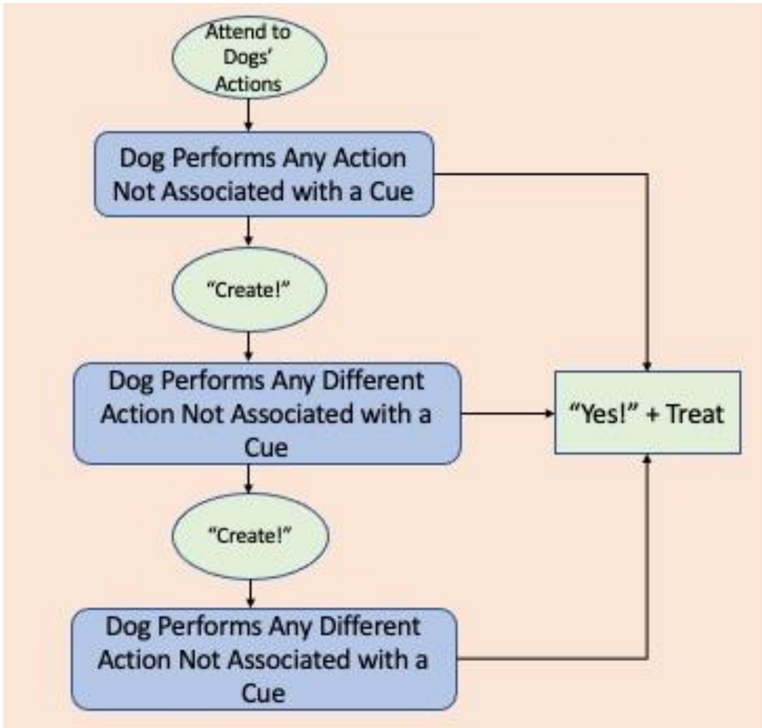
The subjects in this experiment were four pet dogs: Todd, a seven-year-old male chihuahua, Layla, a six-year-old female golden retriever, Gizmo, a three-year-old male Havanese mix, and Snickers, a nine-year-old male Cavalier King Charles spaniel. These dogs were selected for their availability to participate in all required training and testing sessions as well as for their previous experience with cognitive tasks (see Supplemental Materials for additional details on the training histories of subjects). Todd and Layla had previously been successfully trained to repeat actions on cue, including innovated actions (Scagel & Mercado, 2022a), though the range of behaviors they could innovate was not explored. Todd and Gizmo were trained and tested in their owner's homes, and Layla was trained and tested in the Neural and Cognitive Plasticity Laboratory on campus at the University at Buffalo. Each of these three dogs were trained and tested in the same location. Snickers was trained and tested in his owner's home as well as the laboratory. Additionally, each dog was trained and tested by one trainer to ensure responses in each session were assessed consistently; this trainer had six years of experience training dogs in professional contexts. Dogs were fed their normal diet throughout the experiment and always had access to water, though owners fed their dog a smaller amount of their normal food on training and testing days to prevent satiation. Training and testing lasted approximately four months for each dog.

General Training Procedure

During training with the "create" cue, the dogs were required to produce a novel behavior in response to the word "create." In this instance, an action was considered "novel" if it was a behavior that could be voluntarily performed, had not been produced before within the current training/testing session and was not already associated with a specific cue. Variations on previously offered "create" responses were also counted as novel. The trainer began teaching the "create" cue by attending to the dog's actions and rewarding any actions performed that were not associated with a cue (Figure 1). This could include potentially inadvertent actions such as moving a paw, turning the head, sneezing, vocalizing, touching an object with the nose, etc. Once the dog was reliably innovating actions, the "create" cue was added, and the trainer rewarded any action produced in response to the "create" cue that was not already associated with a specific cue and had not been offered previously within the current training session. Any actions judged to be novel by the trainer were rewarded with a verbal "yes" and a small dog treat approved by the dog's owner (see the Supplemental Materials for a more detailed discussion of the benefits and limitations of this training approach). Actions the dog had already produced within the given training session were not reinforced.

Figure 1

Procedure for Training Dogs to Innovate Actions in Response to "Create"



Note. Dogs were trained to innovate actions through a process of shaping in which they were reinforced for producing any action other than those associated with cues or that they produced in response to the "create" cue earlier within a training session.

Each training session lasted one hour, and each dog typically was trained in two sessions per week. During each training session, "create" training continued if the dog was motivated to voluntarily participate. Any signs of disengagement, such as refusing to offer any behavior in response to the "create" cue, walking away from the training area, getting a drink of water, or indicating a need to relieve themselves resulted in a five to 10-minute break, depending on the dog's motivation. Layla and Todd underwent 19 training sessions prior to testing, Snickers experienced 18 training session prior to training, and Gizmo underwent 17 training sessions prior to testing due to scheduling conflicts.

Behavioral Innovation Testing

Dogs were tested on their ability to innovate actions using methods closely matching those developed by Braslau-Schneck (1994) to test innovativeness in bottlenose dolphins. Test sessions consisted of 24 discrete trials, 16 of which were "create" trials. In the other eight trials of each session, deemed "non-create" trials, each dog was cued to produce one of two trained actions they were familiar with and proficient at performing (see Table S1 in Supplementary Materials). Non-create actions were to be performed only once prior to reinforcement in each non-create trial. Non-create trials were included to provide opportunities for the dogs to easily earn rewards and avoid frustration. Prior to each trial, the dog was instructed to sit facing the trainer in a "ready" position. All non-create cues were combined gestural and verbal cues to remain as consistent as possible across subjects. The "create" cue consisted of only the word "create" with no accompanying hand gesture. The "create" trials were divided into three blocks – an initial block of eight consecutive "create" trials followed by two blocks of four consecutive "create" trials. The first and last two trials within a testing session were "non-create" trials, and each block of "create" trials was separated by two “non-create” trials. The order of the two actions cued in “non-create” trials was randomized. Every “non-create” trial was reinforced with a verbal “yes” and a treat when the dog performed the cued behavior. Incorrect performance of “non-create” trials, including a failure to perform any action, received no reinforcement and was considered the end of the trial. “Create” trials were only reinforced if the action produced was not already offered in any prior trials within the test session and did not already correspond to a specific cue (i.e., a dog who can “spin” on cue was not be reinforced if he performed a “spin” in response to the “create” cue), as was required during training. During test sessions, the trainer judged whether offered actions were correct or incorrect. Though this introduced the potential for incorrect or missed reinforcement, real-time judgement allowed dogs to receive the same schedule of reinforcement they received during training and kept dogs motivated to participate. All responses to the “create” cue were recorded on video and in writing.

In this study, the reinforcement contingency during testing was the same as that used in training. This differs from the method of Braslau-Schneck (1994), who did not reinforce dolphins for responses to “creative” during test sessions. The change in procedure was made because she found that dolphins became more hesitant to produce actions to the “creative” cue when they were not rewarded. After being given the “create” cue, dogs had five seconds to perform an action. If no action was offered within these five seconds, meaning the dog remained in the “ready” position and did not produce any discernable movement or vocalization, the trial ended, and the dog received no reinforcement. Because dogs often produced multiple actions simultaneously or in rapid succession, all actions produced within the first five seconds after the “create” cue were recorded and considered when judging the novelty of the response. For instance, a dog could lift a paw while turning its head and vocalizing at the same time, making it impossible to only count one component action as correct or incorrect. Action sequences that included previously performed actions as well as completely novel actions were still considered novel responses as long as that particular action sequence had not been already performed within the given test trial. Recording sequences of behaviors offered within a given period after the “create” cue is consistent with previous methods of training dolphins to vary their behavior when asked (Kuczaj & Eskelinen, 2014b). Dogs were tested in a total of 96 “create” trials over six test sessions. Each test session was recorded via computer webcam (Supplemental Materials provide a more extended discussion of the rationale for this testing approach, as well as its limitations, and a more detailed example of the structure of a testing session).

Controls to Limit and Assess the Possible Role of Social Cueing of Innovated Actions

Though the “create” cue itself did not inform the dogs what they were supposed to do, it is possible the trainer, who was responsible for judging the “creativity” of each response to determine when to provide reinforcement, could have inadvertently cued subjects and guided them to a response deemed correct. To reduce the chance that dogs were being influenced by the behavior of the trainer, the trainer and dog began each trial in the exact same location within the testing room, and the trainer did not move or speak during any trials unless she was giving a cue or rewarding the dog.

To assess whether a dog could innovate actions in responses to a “create” cue in the absence of a knowledgeable trainer, control tests were conducted with Layla in which all access to relevant social cues was removed.

In control sessions, a naïve trainer who did not know what was considered a correct response to the “create” cue gave all cues to the dog. The trainer who determined whether Layla’s responses were correct observed control sessions from another room via a webcam. Two control sessions were conducted, and these were structured the same way as all other tests. During each control trial, the naïve trainer gave each cue to Layla. Layla then had five seconds to respond, and the trainer watching via webcam informed the naïve trainer whether to provide reinforcement or not. If Layla did not offer any response within five seconds, the trial ended and the next trial began. Control session responses were scored the same way as all other test sessions.

Scoring

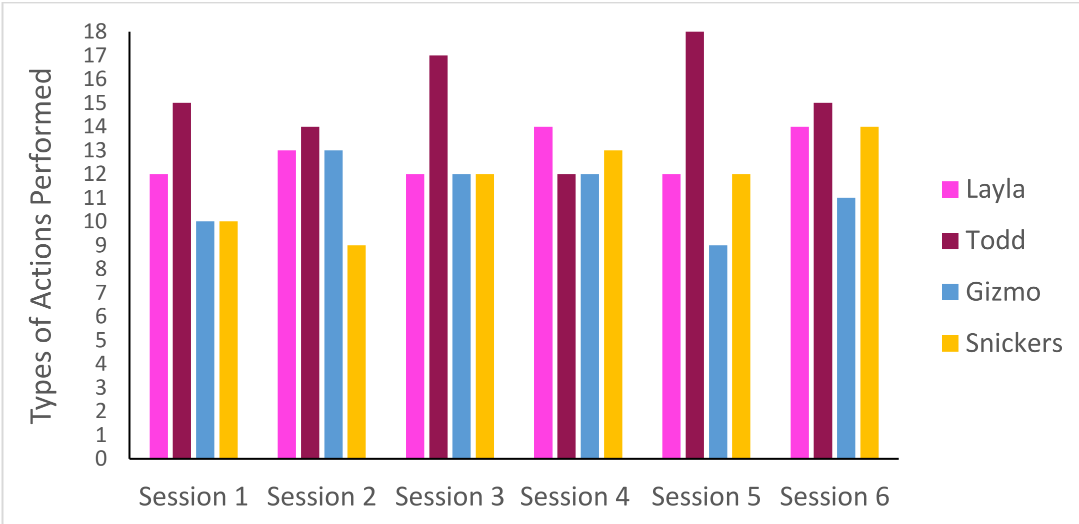
Videos were used to describe and categorize the behaviors the dogs offered in response to the "create" cue by type of action. For instance, moving the head in a circle or turning the head to one side would be considered two different actions, but would both be categorized as head movements. Dogs were scored primarily using metrics developed by Braslau-Schneck (1994) to describe responses of bottlenose dolphins to a "creative" cue. First, the number of correct responses to "create" was recorded for each dog. Within a given test session, actions produced in response to "create" trials were considered correct if the dog produced a novel action or sequence of actions within five seconds after the cue was given. The number of trials in which a dog repeated a response he or she had already produced in response to "create" within the current test session as well as the number of trials in which a dog repeated the response given in the preceding trial were also calculated. The total number of actions produced in response to the "create" cue for each dog was recorded and categorized according to type of behavior. The absolute and relative frequency of each kind of behavior was calculated for each dog, as well as the average number of different kinds of behaviors offered to "create" per test session. A "creativity index" was also calculated for each dog by dividing the number of different behaviors offered by the total number of responses given for "create." The "creativity index" provided a measure of the proportion of all behaviors offered in response to "create" that were novel. This value could range from a score of 0 if the dog did not respond at all to a score of 1 if every action produced in response to the "create" cue was different. Lastly, the distribution of responses to the first "create" cue per test session were recorded for each dog to examine if dogs first defaulted to any certain type of behavior when initially given the "create" cue within a session and then modified their behavior in subsequent test trials, or if they varied their initial response to the cue across sessions.

Results

Dogs correctly produced responses they had not already performed within the current test session in 64-78 trials (66-81%) and produced an average of 11.79 (74%, SD = 1.59) novel actions/action sequences during each test session out of a possible 16 "create" trials (Figure 2). Generally, they did not repeat actions or action sequences they had already produced within a given test session (Figure 3) and did not respond with behaviors already associated with specific cues. When dogs produced responses they had already offered, they were more likely to do so during the second half of a test session (M = 2.75, SD = 1.45) than during the first half of a test session (M = 1.33, SD = 0.92), $t(23) = -4.30$, $p < .001$, Cohen's d = 1.17, calculated using a two-tailed t test. Layla, Todd, and Gizmo responded correctly in 100% of non-create trials and Snickers responded correctly in 98% of non-create trials. The dogs often produced a sequence of actions in response to the "create" cue before the end of a trial. If these sequences were divided into discrete actions, then Layla, Todd, Gizmo, and Snickers produced 212, 225, 184, and 283 total actions respectively in response to the "create" cue over all test trials with an average of 2.35 (SD = 0.47) actions per "create" trial. When categorized into types of actions, the dogs produced an average of 27 (SD = 5.29) different types of actions overall and an average of 12.73 (SD = 1.78) types of behaviors each test session. As there were six test sessions, there were six first "create" trials, and all dogs produced unique actions or action sequences in all six of these first trials, meaning they did not begin any test session with the same response. Most novel actions were produced during the first session of testing with the "create" cue (Figure 4). During control sessions, Layla correctly innovated actions in 18 of 32 trials (56.25%); she repeated a response she had produced in the previous "create" trial four times (12.50%). Layla's creativity index score during control sessions (0.26) was higher than that of any dog during standard test sessions (ranging from 0.08-0.15), which was the proportion of all actions produced that were novel. Inter-rater reliability for coding responses was 85.77% agreement for all dogs' test sessions with a Cohen's Kappa of 0.50. Each dog's results are reported individually within the Supplemental Materials. A video of sample test trials can be viewed here: https://doi.org/10.6084/m9.figshare.25854151.v1.

Figure 2

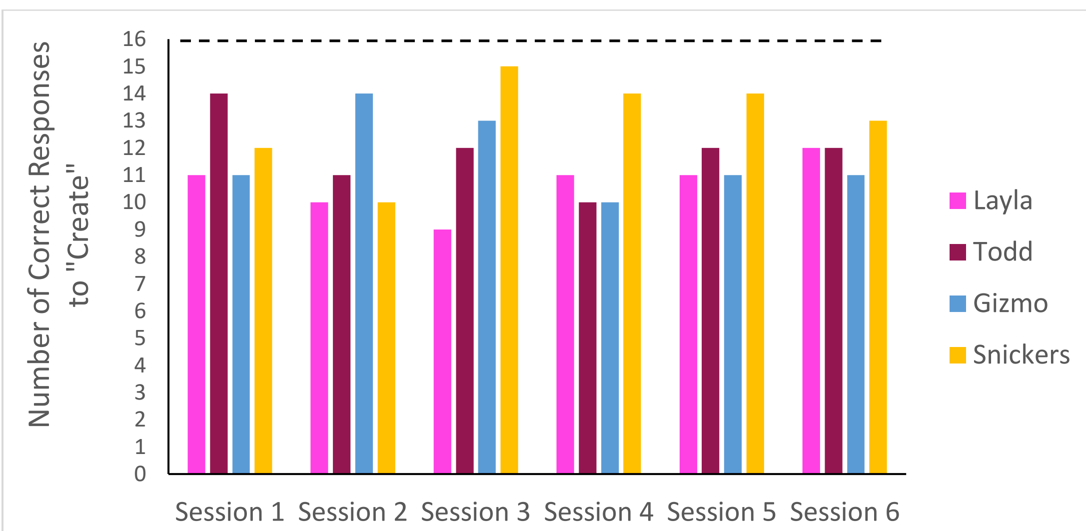
Number of Different Types of Actions Dogs Performed After Being Given the "Create" Cue per Test Session for Each Dog



Note. These actions were represented within each dogs' responses (actions/action sequences) to the "create" cue as multiple actions could be performed within the five seconds after a "create" cue was given. Represented within these actions are both correct and incorrect responses to the "create" cue. A correct response was an action or sequence of actions not already performed within the current trial and that was not already associated with a cue. An incorrect response was an action or sequence of actions already performed in a previous trial to the "create" cue, or an action already associated with a specific cue. The same types of actions can be included in the bars of multiple sessions. Layla produced an average of 12.8 types of actions per session, Todd produced an average of 15.2 types of actions per test session, Gizmo produced an average of 11.2 actions per session, and Snickers produced an average of 11.7 actions per session.

Figure 3

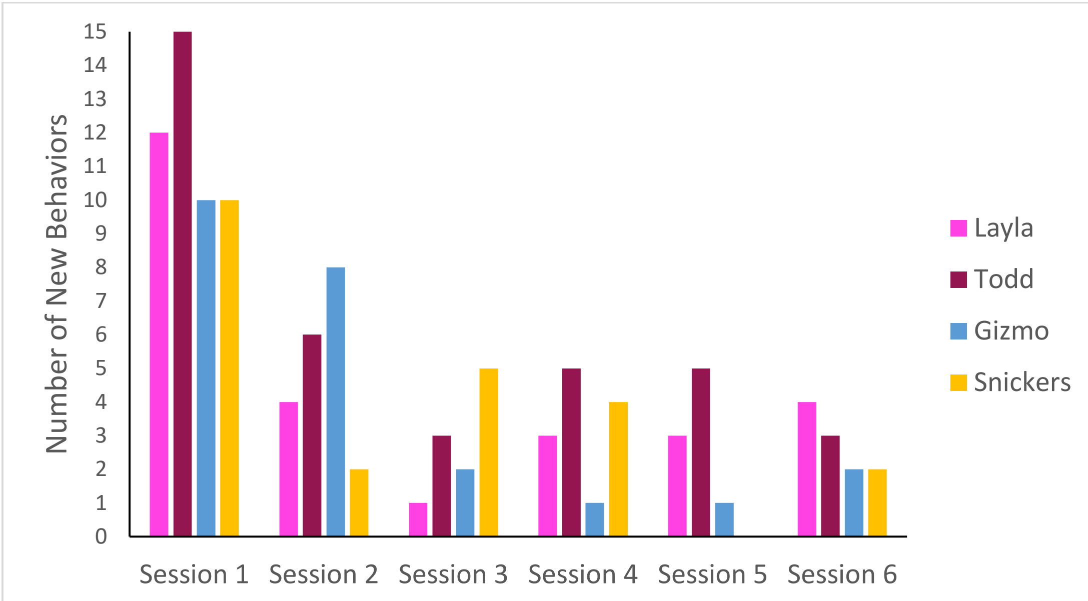
Number of Correct Responses to the "Create" Cue Given by Each Dog in Each Test Session



Note. A correct response was any action or sequence of actions produced within the first five seconds after a "create" cue was given, provided that action or action sequence had not already been performed in any prior trials within the current test session and was not a singular action that was already associated with a specific cue. The dotted line depicts the maximum possible correct responses, which was 16.

Figure 4

Number of New Types of Behaviors Produced Each Session by Each Dog in Response to the "Create" Cue



Note. Number of new types of behaviors produced each session by each dog in response to the "create" cue. Each unit is a unique behavior not performed during any prior test session. Layla performed a total of 28 behaviors, Todd performed a total of 34 different behaviors, Gizmo performed a total of 24 different behaviors, and Snickers performed a total of 22 different behaviors.

Discussion

The four dogs assessed in this investigation were able to produce a variety of behaviors in response to the "create" cue. Most responses offered within a given test session were untrained actions that the dogs had not already produced in other "create" trials during that session. Overall, our results are consistent with findings from a recent study of dogs' capacity to innovate actions in less formalized testing conditions (Willgohs et al., 2022). The degree to which these findings would apply to all dogs in general outside of an experimental context is not clear, but the fact that four pet dogs were able to successfully learn to innovate behaviors on cue indicates it is possible other dogs could do this as well.

The training requirement that dogs could not perform a response already associated with a cue to receive reinforcement in the current study may have encouraged them to actively avoid previously trained actions. This differs from the definition of "novel" used in other animal creativity studies (Kuczaj & Eskelinen, 2014b; Lawrence et al., 2016; Pryor et al., 1969; Willgohs et al., 2022), in which any action not already performed within the current session was considered "novel," even if it was a behavior associated with a cue that subjects had been previously trained to perform. We decided not to consider these types of actions "novel" to encourage the dogs in our study to produce actions they had never been explicitly trained to perform. In comparison, four bottlenose dolphins trained to be "creative" on cue offered only 25-52 % untrained responses (Braslau-Schneck, 1994), and killer whales produced 0-4 % untrained actions (Hill et al., 2022). Neither dolphins nor killer whales were required to avoid trained behaviors in previous studies. However, dogs trained to innovate on cue using the same criteria for reinforcement as in past studies of dolphin creativity still proved to be more likely to offer untrained actions (Willgohs et al., 2022). The factors that determine the range of actions that individuals select from when innovating in such tasks thus remain unclear. Dogs in the current study did sometimes include previously trained actions within novel action sequences, combining these behaviors with novel ones in response to a "create" cue. Still, such occurrences were rare (2-8% of all actions produced). The dogs in the current study seemed to learn that avoiding familiar actions was part of the task requirements.

Behaviors such as sneezing, barking, shaking, or scratching, among others, were counted as responses in the current study because each of these actions can be placed under stimulus control. Alternatively, these actions could be interpreted as displacement behaviors — actions performed with no discernable function related to the animal's current situation that may indicate feelings of conflict or stress (Pedretti et al., 2022). If so, this could indicate that these particular actions were perhaps a by-product of a dog feeling stressed or frustrated when he or she did not know the "correct" response to the "create" cue. Behaviors like sneezing, shaking, or scratching could also simply be the result of other physiological motivations that happened to occur during a testing session. Whether the dogs voluntarily chose to perform these behaviors because they met the criteria for reinforcement or inadvertently performed them because dogs were stressed or otherwise motivated is difficult to determine and is a limitation of our methodology. These responses represented only a minor portion of the actions dogs produced in response to the "create" cue, such that excluding them as correct responses would not substantively change the results of the current study.

All dogs tested produced the greatest number of novel types of actions during the first two sessions of testing and produced an average of 2.8 additional types of novel actions per session during subsequent test sessions. This finding indicates that the dogs generally chose from actions they had been previously reinforced for producing in response to the "create" cue but combined these actions in different ways to respond with novel action sequences. There were particular classes of actions that each dog tended to perform more frequently than others, as was the case with dolphins (Braslau-Schneck, 1994; Kuczaj & Eskelinen, 2014b; Hill et al., 2022). These actions may have been more frequently reinforced after the "create" cue during training, a pattern consistent with classical learning theory, given that reinforcement increases the frequency of behaviors (Lawrence et al., 2016; Skinner, 1965). Alternatively, certain actions may have been preferred by each dog because those actions required relatively less effort to perform (e.g., Willgohs et al., 2022, found that dogs were more likely to self-select "low energy" actions when asked to innovate). Another possibility is that dogs offered most of their different innovative actions toward the beginning of testing because at some point they reached a plateau of creativity or "ran out of ideas," and then defaulted to certain actions when they did not know what to do. Even adult humans struggle when faced with tasks in which they are supposed to respond randomly (Tune, 1964; Wagenaar, 1972). Still, dogs' persistent variations in the way they produced these actions across sessions suggests they learned they would not be reinforced for performing the same action or action sequence when the "create" cue was given (see also Willgohs et al., 2022).The dogs did not respond the same way to every first "create" cue within a testing session indicating that, like dolphins (Braslau-Schneck, 1994), they were not simply responding in a stereotyped manner to the "create" cue according to a certain order of behaviors they had previously been reinforced for producing. In fact, all the dogs produced a unique response to every initial "create" cue within each testing session.

The dogs avoided repeating responses and few response repetitions were of actions or action sequences produced in the just-prior trial. These results suggest the dogs may have used a strategy of avoiding recently performed actions (see also Kuczaj & Eskelinen, 2014b; Hill et al., 2022; Van Steyn, 2018) and could be an indication that performance during earlier trials was more likely to be forgotten than performance during the immediately preceding trial. In that scenario, most repeated responses should have occurred later in test trials, and in fact this was the case. Though the exact working memory span of domestic dogs for recently performed actions is unknown, performance by dogs on working memory tasks typically decreases with increasing passage of time (Krichbaum et al., 2021; Scagel & Mercado, 2022a; Tapp et al., 2003). There are several other factors that may constrain an individual's ability to avoid repeating earlier actions, such as a preference for seeking the path of least resistance or a bias toward performing specific actions. For instance, Pryor (2000) noted that when she attempted to train a pigeon to produce innovative actions it initially only produced species-typical behaviors. Only after days of reinforcing the pigeon for producing different actions did it begin expanding its behavioral repertoire to include novel actions. Similarly, Manabe and colleagues (1997) found that budgerigars (*Melopsittacus undulatus*) trained to innovate vocalizations offered more complex actions later in training and only when reinforcement was restricted to calls that did not match the previous three offered. One bird even began including pecks on a wall within the novel sounds it produced. Walruses (*Odobenus rosmarus divergens*) trained to produce novel vocalizations gradually came to produce novel sounds rapidly only after the animals became more experienced with the task (Schusterman & Reichmuth, 2008). Animals operantly conditioned in puzzle boxes tend toward producing more stereotyped actions as they gain experience with a task (Guthrie & Horton, 1946), suggesting that learning mechanisms favor uniformity in actions produced in the context of familiar problems. In short, learning to innovate actions on cue likely requires inhibiting multiple natural tendencies (Brosnan & Hopper, 2014), as well as some capacity to maintain memories of recent self-performed actions.

In several ways, the number and variety of actions produced by domestic dogs in the "create" task were comparable to the performance of bottlenose dolphins in similar tests (Table 1), despite differences in how dolphins were reinforced during testing, the extensive prior training dolphins received, the absence of domestication, and the large differences in action repertoires and living conditions between dogs and dolphins. Braslau-Schneck (1994) trained four dolphins to respond to a "creative" cue by producing a "behavior that is 'different' from the 'previous' behaviors," (how unique each behavior needed to be and what counted as a "previous" behavior was not strictly defined). The dolphins she tested produced an average of 4-11 different actions per session, whereas the dogs in this study produced an average of 11-15 different behaviors per session. Although methodological differences between these two studies preclude direct statistical comparisons between the performances of dogs and dolphins, subjects' reactions to "create" cues showed more similarities than differences. Recent work comparing action innovation in preschool children and dolphins likewise revealed similar levels of creativity, further suggesting cross-species continuity in creative capacities (Melzer et al., 2022; see also Goetz & Baer, 1973; Holman et al., 1977).

Table 1

Performance in Response to "Create" by Dolphins (Braslau-Schneck, 1994) and Dogs (Present Study)

Subject	Number of Trials	Total Number of Actions Produced to "Create"	Types of Actions Produced in Response to "Create"	Creativity Index Score
Dolphins				
Hiapo	144	219	23	0.1
Elele	144	323	72	0.2
Akeakamai	94	102	22	0.2
Phoenix	92	125	16	0.1
Dogs				
Layla	97	212	28	0.1
Todd	98	225	34	0.2
Gizmo	96	184	24	0.1
Snickers	96	283	22	0.08

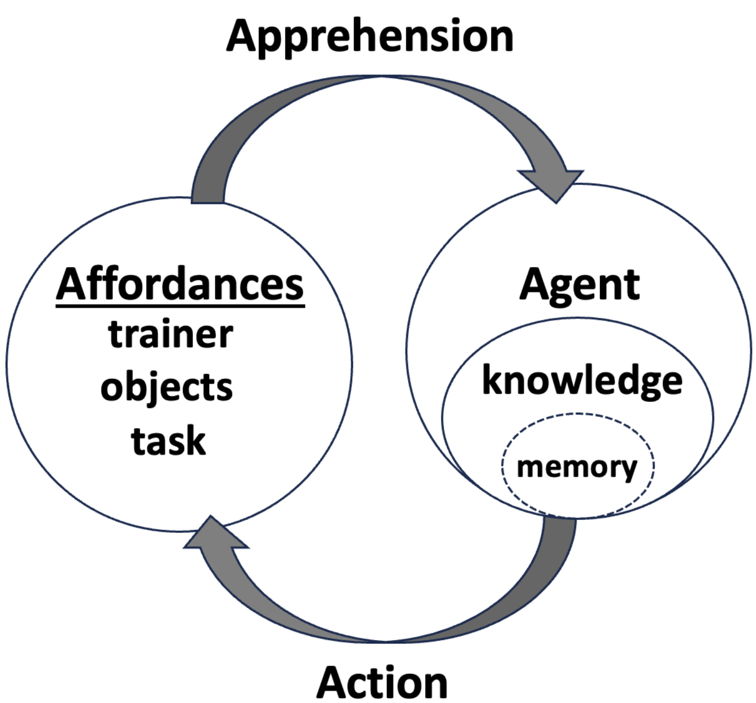
Dogs and dolphins have demonstrated the capacity to innovate actions on cue in controlled experiments, but many questions remain as to what cognitive mechanisms enable them to do so and whether those mechanisms differ significantly from those engaged by innovative children. Are dogs' and dolphins' innovations intentional or accidental? Are they improvising, problem solving, cheating, or reflexively reacting to states of frustration or confusion? Are all individuals selecting actions (or reacting) to "create" cues in the same way? Or, might some individuals have formed different impressions of what trainers are requesting when they produce this cue? How much of the variety in actions produced in response to such cues is determined by an individual's creativity versus the individual's ability to inhibit natural tendences toward performing habitual or default responses? What exactly are dogs, dolphins, and children doing when they respond to requests to innovate?

At a minimum, dogs in the current study were monitoring the actions of trainers, recognizing that certain sounds and gestures produced by the trainers differed from others, and that a subset of those cues were predictive of opportunities to receive treats and social praise. The variety of actions that dogs produced in response to the “create” cue relative to the stereotypy of actions provoked by other instructions further suggests that they learned the “create” cue was in some way different from most other cues that trainers produced. Dogs’ ability to avoid repeating actions offered in response to the cue could indicate they were also monitoring their own recent actions and/or that they ramped up their variations in exploratory actions (or reactions to confusion and uncertainty). The fact that the dogs avoided offering trained actions while producing a subset of untrained actions with differential frequency argues against the possibility that they randomly selected actions from their behavioral repertoire. These findings imply that the dogs were in some way self-selecting actions rather than relying only on external stimuli, and that the “create” cue (along with cues from objects and places within the room) became associated with this process of self-selection.

What distinguishes self-selection of innovative acts from everyday action selection is a kind of impulse control that enables the actor to deviate from the norm. In the context of animals interacting with trainers, the selection of innovative actions is more of a joint process because each agents’ actions are contingent on the actions of their partner. This joint process can be conceptualized within existing frameworks for understanding creativity in which acquired improvisational skills are engaged as part of a social interaction (Figure 5). From this perspective, the capacity of a dog or a dolphin to innovate actions on cue emerges from repeated social interactions that change the ways in which individuals represent and select actions. Action innovation within experimental tasks can then be viewed as a cognitive skill, because the dog or dolphin learns novel, context-dependent ways of selecting actions within its existing behavioral repertoire rather than learning how to perform new perceptual-motor skills (Mercado, 2008).

Figure 5

Dynamic Interactions Contributing to Action Innovation in the Context of an Experimental Task



Note. A dog (agent) participating in a task that requires innovating actions on cue must learn to apprehend when the production of self-selected actions is desired based on observations of the trainer’s behavior, the objects within the area, accumulated experience about the task demands within the experimental context. Repeated experiences responding to a “create” cue, followed by feedback in the form of treats, praise (or lack thereof) shapes the dog to select actions without repetition. Successful performance of this task requires the dog to learn to select actions sufficiently different from each other that the trainer judges them to differ from preceding actions.

Limitations

Because our study included only four dogs, it is unclear how our results would generalize to dogs of different breeds, sexes, ages, or training histories. Our sample size is consistent, however, with other studies in which nonhuman species perform complex cognitive tasks due to the extensive training required for each animal (Lyn et al., 2011; Mercado et al., 1998; Pepperberg, 1994; Pepperberg & Gordon, 2005; Pepperberg, 2015; Pilley & Reid, 2011; Pilley, 2013; Ramos & Ades, 2012; Scagel & Mercado, 2022a, 2022b; Willgohs et al., 2022). Training took place over several months and relied on the availability of the dogs' owners.

A second limitation of the study is that control tests were conducted only with Layla, in part because the other dogs were either unavailable or not comfortable interacting with an unfamiliar trainer. Given that the goal of these control tests was to establish that a dog can innovate actions in the absence of relevant social cues, establishing this is possible for a single dog shows that dogs as a species are capable of innovating actions without relying on such cues (just as experiments with a single horse, Clever Hans, was sufficient to establish that animals have the capacity to "solve" math problems using social cues, Beran, 2012; Pfungst, 1911). Although it is possible, in principle, that the three dogs not tested in control trials were using social cues to somehow guide their innovations, in the absence of any clear differences in the training procedures or test performance across subjects, the assumption that all dogs learned the task in similar ways is simpler and more consistent with the evidence.

When assessing inter-rater reliability, a second trainer scored each dog's responses to "create" trials by watching the recordings of test sessions. This inevitably led to the second rater having a different vantage point from which to score behavior than the trainer who conducted test sessions in-person with each dog and judged responses in real time, which could have led to each rater scoring some responses differently and could have contributed to a moderate Cohen's Kappa. A delay in reinforcement that could have resulted from incorporating more controlled evaluation of responses during test sessions could have hindered dogs' responding to the "create" cue, as increasing delays in reinforcement can result in a decrease in responding (Yamamoto et al., 2009). In the future, methods of scoring responses could be made more comparable, for instance by having the trainer wear a small camera that records the dog's actions from her viewpoint.

The setup of testing sessions may have inherently constrained the way dogs responded to "create" cues during testing. First, during the extensive training dogs received, responses deemed "correct" by the trainer were reinforced. This could have led to previously reinforced responses being preferred (Lawrence et al., 2016; Neuringer, 2004); the observation that each dog did offer one or two types of actions more frequently than others could support this. However, no one type of action was produced with a frequency of over 23%, and each dog also produced a wide variety of other types of actions. Additionally, Schwartz (1982) found that specifically reinforcing pigeons for offering pecking responses that varied from previous trials prevented stereotyped responding. In a study that required pigeons to vary the pattern of pecking responses with two different reinforcement contingencies, reinforcement appeared to both enhance and hinder response variability depending on how it was delivered (Neuringer, 2004). The choice was ultimately made to provide reinforcement to dogs in this study to ensure they continued to be motivated to respond to the "create" cue throughout the study, because when Braslau-Schneck (1994) did not provide any reinforcement for dolphins during testing, the dolphins became hesitant to offer any responses over time. Additionally, reinforcement only hindered response variation in pigeons when it was initially withheld and then only delivered after the fourth varied response (Neuringer, 2004). These examples suggest withholding reinforcement may be more detrimental than providing it.

Conclusions and Future Directions

Collectively, our results provide further evidence that "creativity" is not a characteristic unique to humans and instead is more likely an adaptive problem-solving process that evolved in a wide variety of species (Pryor & Chase, 2014). The dogs likely learned that they needed to avoid performing certain actions and monitor recent actions they had already performed to get treats (Kuczaj & Eskelinen, 2014). The current results, along with those showing that dogs can repeat self-selected actions (Scagel & Mercado, 2022a), suggest that dogs possess mental representations of their own actions that they can flexibly use to influence their behavior.

Today, training in behavioral innovation is used in zoos, aquariums, and by dog trainers as an enrichment task for animals in captivity (Pryor, 2009). Different species are likely to produce different types of actions when cued to innovate based on their own physical capabilities and the affordances of their environments (Montesano et al., 2008). Examining the different factors that determine how individuals innovate actions in experimental contexts can clarify the mechanisms that enable individuals to "be creative." Future research could examine how innovations performed by highly trained working dogs, service dogs, or show dogs, who are trained to perform a wide array of behaviors (Hall et al., 2021), compare to those of companion dogs and wild animals that have been trained to innovate actions on cue. Experiments that expand the range of contexts within which animals (and their trainers) are encouraged to try new things may provide further insights into the creative potential of non-humans.

Acknowledgments

The authors thank Avani Gulati for her assistance in training and testing the dogs as well as all dog owners for allowing their dogs to participate.

Author Contributions: EM and AS conceived and designed the experiment. AS and AH trained dogs in the experiment and analyzed the behavioral data. AS tested the performance of dogs after training. EM, AS, and AH contributed to the writing and editing of the manuscript.

Data Availability: Data from test sessions are available upon request.

Conflict of Interest: The authors declare no conflict of interest.

References

Aria, M., Alterisio, A., Scandurra, A., Pinelli, C., & D'Aniello, B. (2021). The scholar's best friend: Research trends in dog cognitive and behavioral studies. *Animal Cognition*, *24*, 541-553. https://doi.org/10.1007/s10071-020-01448-2

Bailey, A. M., McDaniel, W. F., & Thomas, R. K. (2007). Approaches to the study of higher cognitive functions related to creativity in nonhuman animals. *Methods*, *42*(1), 3-11. https://doi.org/10.1016/j.ymeth.2006.12.003

Balsam, P. D., Deich, J. D., Ohyama, T., & Stokes, P. D. (1998). Origins of new behavior. In W. T. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 403–420). Allyn & Bacon.

Bateson, P., & Martin, P. (2013). *Play, playfulness, creativity and innovation.* Cambridge University Press. https://doi.org/10.1017/CBO9781139057691

Beran, M. J. (2012). Did you ever hear the one about the horse that could count? *Frontiers in Psychology*, *3*, 357. https://doi.org/10.3389/fpsyg.2012.00357

Berthelet, A., & Chavaillon, J. (Eds.). (1993). *The use of tools by human and nonhuman primates.* https://dx.doi.org/10.1093/acprof:oso/9780198522638.001.0001

Braslau-Schneck, S. (1994). *Innovative behaviors and synchronization in bottlenosed dolphins*. Unpublished master's thesis, University of Hawaii, Honolulu.

Brosnan, S. F., & Hopper, L. M. (2014). Psychological limits on animal innovation. *Animal Behaviour*, *92*, 325-332. https://doi.org/10.1016/j.anbehav.2014.02.026

Burghardt, G. M. (2005). *The genesis of animal play: Testing the limits*. Cambridge, MA: MIT Press.

Burghardt, G. M. (2015). Creativity, play, and the pace of evolution. In A. B. Kaufman, & J. C. Kaufman (Eds.), *Animal creativity and innovation* (pp. 129-161). Academic Press.

Cook, P. R., Prichard, A., Spivak, M., & Berns, G. S. (2016). Awake canine fMRI predicts dogs' preference for praise vs food. *Social Cognitive and Affective Neuroscience, 11(12),* 1853-1862. https://doi.org/10.1093/scan/nsw102

Coppinger, R., & Coppinger, L. (2001). *Dogs: A startling new understanding of canine origin, behavior & evolution*. Simon & Schuster.
Cutting, A. E. (1997). *Memory of self-selected behavior in a bottlenosed dolphin (Tursiops truncates)*. Unpublished master's thesis, University of Hawaii, Honolulu.
Dudzinski, K. M., Yeater, D., Bolton, T., Eskelinen, H., & Hill, H. (2018). Defining creativity and confirming understanding of the concept in dolphins: Research and training perspectives. *Aquatic Mammals*, *44*, 426-436. https://dx.doi.org/ 10.1578/AM.44.4.2018.426
Fagen, R. (1981). *Animal play behavior*. New York, NY: Oxford University Press.
Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and the emergence of new forms in children's block building. *Journal of Applied Behavior Analysis*, *6*, 209-217.
Gutiérez, B., & Escobar, R. (2022). Effects of training different levels of operant variability in rats. *Behavioural Processes*, *201*, 104730. https://doi.org/10.1016/j.beproc.2022.104730
Guthrie, E. R., & Horton, G. P. (1946). *Cats in a puzzle box*. New York, NY: Rinehart & Company
Hall, N. J., Johnston, A. M., Bray, E. E., Otto, C. M., MacLean, E. L., & Udell, M. A. (2021). Working dog training for the twenty-first century. *Frontiers in Veterinary Science*, 834. https://doi.org/10.3389/fvets.2021.646022
Hill, H. M., Weiss, M., Brasseur, I., Manibusan, A., Sandoval, I. R., Robeck, T., Sigman, J., Werner, K., & Dudzinski, K. M. (2022). Killer whale innovation: Teaching animals to use their creativity upon request. *Animal Cognition*, *25*, 1-18. https://doi.org/10.1007/s10071-022-01635-3
Holman, 1., Goetz, E., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In B. C. Etzel, 1. M. LeBlanc & D. M. Baer (Eds.), *New developments in behavioral research: Theory, method and application*. (pp. 441-471). Wiley.
Karwowski, M., Dul, J., Gralewski, J., Jauk, E., Jankowska, D. M., Gajda, A., Chruszczewski, M. H., & Benedek, M. (2016) Is creativity without intelligence possible? A necessary condition analysis. *Intelligence*, *57*, 105-117. https://doi.org/10.1016/j.intell.2016.04.006
Kaufman, A. B. (2021) Innovation in marine mammals. In: Kaufman AB, Call J, Kaufman JC (Eds.), *The Cambridge Handbook of Animal Cognition* (pp. 710-720). Cambridge University Press. https://doi.org//10.1017/9781108564113
Kaufman, J. C., & Kaufman, A. B. (2004). Applying a creativity framework to animal cognition. *New Ideas in Psychology*, *22*, 143-155.
Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four C model of creativity. *Review of General Psychology*, *13*, 1-12.
Kaufman, A. B., Butt, A. E., Kaufman, J. C., & Colbert-White, E. N. (2011). Towards a neurobiology of creativity in nonhuman animals. *Journal of Comparative Psychology, 125*, 255-272. https://dx.doi.org/10.1037/a0023147.
Kaufman, A. B., & Kaufman, J. C. (2014). Applying theoretical models on human creativity to animal studies. *Animal Behavior and Cognition*, *1*, 77-89. https://doi.org/10.12966/abc.02.06.2014
Köhler, W. (1926). *The mentality of apes* (2nd ed.). Harcourt, Brace & Company, Inc.
Krichbaum, S., Lazarowski, L., Davila, A., Cox, E., Smith, J. G., & Katz, J. S. (2021). Dissociating the effects of delay and interference on dog (*Canis familiaris*) working memory. *Animal Cognition*, *24*, 1259–1265. https://doi.org/10.1007/s10071-021-01509-0
Kuczaj, S. A. (2017). Animal creativity and innovation. In J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, & T. Zentall (Eds.), *APA handbook of comparative psychology: Perception, learning, and cognition* (pp. 627-641). American Psychological Association. https://doi.org/10.1037/0000012-028
Kuczaj, S. A., & Eskelinen, H. C. (2014a). Why do dolphins play? *Animal Behavior and Cognition, 2*, 113-127. https://dx.doi.org/10.12966/abc.05.03.2014
Kuczaj, S. A., & Eskelinen, H. C. (2014b). The "creative dolphin" revisited: What do dolphins do when asked to vary their behavior. *Animal Behavior and Cognition*, 1, 66-77. https://dx.doi.org/10.12966/abc.02.05.2014
Kuczaj, S. A., & Walker, R. T. (2012). Dolphin problem solving. In T. Zentall & E. Wasserman (Eds.), *Handbook of comparative cognition* (pp. 736-756). New York, NY: Oxford University Press.
Kummer, H., & Goodall, J. (2003). Conditions of innovative behaviour in primates. In R. M. Reader, & K. N. Laland (Eds.), *Animal innovation* (pp.223-235). Oxford, UK: Oxford University Press.
Lawrence M. K., Borger-Turner, J. L., Turner, T. N., & Eskelinen, H. C. (2016). Investigating the effects of applied learning principles of the "create" response in Atlantic bottlenose dolphins (*Tursiops truncatus*). *International Journal of Comparative Psychology*, *29*, 1–9. https://doi.org/10.46867/ijcp.2016.29.00.11

Lyn, H., Greenfield, P. M., & Savage-Rumbaugh, E. S. (2011). Semiotic combinations in *Pan*: A comparison of communication in a chimpanzee and two bonobos. *First Language, 31(3),* 300-325. https://doi.org/10.1177/0142723710391872

Manabe, K., Staddon, J. E., & Cleaveland, J. M. (1997). Control of vocal repertoire by reward in budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology, 111(1)*, 50-62. https://doi.org/10.1037/0735-7036.111.1.50

Melzer, D. K., Yeater, D., Bradley, M., Hill, H. M., Guerra, G., Salazar, K., Bolton, T., & Dudzinski, K. M. (2022). A comparative test of creative thinking in preschool children and dolphins. *Animal Behavior and Cognition*, *9*, 349-362. https://doi.org/10.26451/abc.09.03.07.2022

Mercado, E., III (2008). Neural and cognitive plasticity: From maps to minds. *Psychological Bulletin*, 134, 109-137.

Mercado, E., III, Ashour, M., & McAllister, S. (2022). Cognitive control of song production by humpback whales. *Animal Cognition, 1-17*. https://doi.org/10.1007/s10071-022-01675-9

Mercado, E., III, & DeLong, C. M. (2010). Dolphin cognition: Representations and processes in memory and perception. *International Journal of Comparative Psychology, 23*, 344-378. https://doi.org/10.5070/P4233020463

Mercado, E., III, Mantell, J. T., & Pfordresher, P. Q. (2014). Imitating sounds: A cognitive approach to understanding vocal imitation. *Comparative Cognition & Behavior Reviews*, *9*, 1–57. https://doi.org/10.3819/ccbr.2014.90002

Mercado, E., III, Murray, S. O., Uyeyama, R. K., Pack, A. A., & Herman, L. M. (1998). Memory for recent actions in the bottlenosed dolphin (*Tursiops truncatus*): Repetition of arbitrary behaviors using an abstract rule. *Animal Learning & Behavior*, *26*(2), 210-218. https://doi.org/10.3758/BF03199213

Mercado, E., III, & Scagel, A. (2022). Second verse, same as the first: Learning generalizable relational concepts through functional repetition. *Animal Cognition*, 1-11. https://doi.org/10.1007/s10071-022-01702-9

Mercado, E., III, Uyeyama, R. K., Pack, A. A., & Herman, L. M. (1999). Memory for action events in the bottlenosed dolphin. *Animal Cognition*, *2*(1), 17-25. https://doi.org/10.1007/s100710050021

Montesano, L. Lopes, M., Bernardino, A., & Santos-Victor, J. (2008). Learning object affordances: From sensory-motor coordination to imitation. *IEEE Transactions on Robotics*, *24*(1), 15-26. https://doi.org/10.1109/TRO.2007.914848

Morris, C. J. (1989). The effects of lag value on the operant control of response variability under free-operant and discrete-response procedures. *The Psychological Record*, *39*(2), 263-270. https://doi.org/10.1007/BF03395067

Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*(1), 3-12. https://doi.org/10.1037/0097-7403.17.1.3

Neuringer, A. (2004). Reinforced variability in animals and people – Implications for adaptive action. *American Psychologist*, *59*(9), 891-906. https://doi.org/10.1037/0003-066X.59.9.891

Patterson, E. M., & Mann, J. (2015). Chapter 4 – Cetacean innovation. In A. B. Kaufman & J. C. Kaufman (Eds.), *Explorations in creativity research* (pp. 73-125). Academic Press. https://doi.org/10.1016/B978-0-12-800648-1.00004-8

Pedretti, G., Canori, C., Biffi, E., Marshall-Pescini, S., and Valsecchi, P. (2022). Appeasement function of displacement behaviours? Dogs' behavioural displays exhibited towards threatening and neutral humans. *Animal Cognition*, *26*, 943-952. https://doi.org/10.1007/s10071-023-01742-9

Pepperberg, I. M. (1994). Numerical competence in an African gray parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, *108*(1), 36-44. https://doi.org/10.1037/0735-7036.108.1.36

Pepperberg, I. M. (2015). Chapter 1 – Creativity and innovation in the grey parrot (*Psittacus erithacus*). In A. B. Kaufman, & J. C. Kaufman (Eds.), *Explorations in creativity research* (pp. 3-29). Academic Press. https://doi.org/10.1016/B978-0-12-800648-1.00001-2

Pepperberg, I. M., & Gordon, J. D. (2005). Number comprehension by a grey parrot (*Psittacus erithacus*) including a zero-like concept. *Journal of Comparative Psychology, 119(2)*, 197. https://dx.doi.org/10.1037/0735-7036.119.2.197

Pfungst, O. (1911). *Clever Hans (the horse of Mr. Von Osten): A contribution to experimental animal and human psychology*. Henry Holt and Company.

Pilley, J. W. (2013). Border collie comprehends sentences containing a prepositional object, verb, and direct object. *Learning and Motivation*, *44*(4), 229-240. https://doi.org/10.1016/j.lmot.2013.02.003

Pilley, J. W., & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, *86*(2), 184-195. https://doi.org/10.1016/j.beproc.2010.11.007

Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior, 12*, 653-661.

Pryor, K. (2000). *Lads before the wind*. Sunshine Books.

Pryor, K. (2009). *Reaching the animal mind* (pp. 82-87, 90-95). New York, NY: Scribner.

Pryor, K., & Chase, S. (2014). Training for variable and innovative behavior. *International Journal of Comparative Psychology, 27*(2), 218-225. https://doi.org/10.5070/P4272022389

Ramos, D., & Ades, C. (2012). Two-item sentence comprehension by a dog (*Canis familiaris*). *PLoS ONE, 7(2),* Article e29689. https://doi.org/10.1371/journal.pone.0029689

Riedel, J., Schumann, K., Kaminski, J., Call, J., & Tomasello, M. (2008). The early ontogeny of human-dog communication. *Animal Behaviour, 75(3)*, 1003-1014. https://doi.org/10.1016/j.anbehav.2007.08.010

Sanz, C. M., Call, J. & Boesch, C. (Eds.) (2014). *Tool use in animals*. New York, NY: Cambridge University Press.

Scagel, A., & Mercado, E., III (2022a). Do that again! Memory for self-performed actions in dogs (*Canis familiaris*). *Journal of Comparative Psychology, 136(2),* 121-139. https://dx.doi.org/10.1037/com0000310

Scagel A., & Mercado E., III (2022b). Same-different conceptualization in dogs (*Canis familiaris*). *Journal of Comparative Psychology*. Advance online publication. https://doi.org/10.1037/com0000332

Schusterman, R. J., & Reichmuth, C. (2008). Novel sound production through contingency learning in the Pacific walrus (*Odobenus rosmarus divergens*). *Animal Cognition, 11*, 319-327. https://doi.org/10.1007/s10071-007-0120-5

Schwartz, B. (1982). Failure to produce response variability with reinforcement. *Journal of the Experimental Analysis of Behavior, 37*, 171-181. https://doi.org/10.1901/jeab.1982.37-171

Shevlin, H. (2020). Rethinking creative intelligence: comparative psychology and the concept of creativity. *European Journal for Philosophy of Science, 11*. https://doi.org/10.1007/s13194-020-00323-8

Shumaker, R. W., Walkup, K. R., & Beck, B. B. (Eds.). (2011). *Animal tool behavior*. Baltimore, MD: Johns Hopkins University Press.

Skinner, B. F. (1965). *Science and human behavior* (No. 92904). Simon and Schuster.

Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior, 23*, 164-176. https://doi.org/10.3758/BF03199931

Tapp, P. D., Siwak, C. T., Estrada, J., Holawachuk, D., & Milgram, N. W. (2003). Effects of age on measures of complex working memory span in the beagle dog (*Canis familiaris*) using two versions of a spatial list learning paradigm. *Learning & Memory, 10(2)*, 148-160. https://doi.org/10.1101/lm.56503

Taylor, K. B. (1995). *Self-selection and self-monitoring of behavior by bottlenosed dolphins* (*Tursiops truncatus*). Unpublished master's thesis, University of Hawaii, Honolulu.

Thompson, H., Reimer, S., Ellis, S. L., & Burman, O. H. (2016). Behaviour directed towards inaccessible food predicts consumption – A novel way of assessing food preference. *Applied Animal Behaviour Science, 178,* 111-117. https://doi.org/10.1016/j.applanim.2016.02.008

Tune, G. S. (1964). A brief survey of variables that influence random generation. *Perceptual and Motor Skills, 18*. 705-710.

Van Steyn, R. J. (2018). An analysis of innovate training with bottlenose dolphins (*Tursiops truncatus*). Unpublished master's thesis, Hunter College, The City University of New York, New York City.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of the literature. *Psychological Bulletin, 77*, 65-72.

Willgohs, K. R., Williams, J., Franklin, E., & Highfill, L. E. (2022). The creative canine: Investigating the concept of creativity in dogs (*Canis lupus familiaris*) using citizen science. *International Journal of Comparative Psychology, 35*. https://doi.org/10.46867/ijcp.2023.35.5621

Yamamoto, M., Kikusui, T., & Ohta, M. (2009). Influence of delayed timing of owners' actions on the behaviors of their dogs, *Canis familiaris*. *Journal of Veterinary Behavior: Clinical Applications and Research, 4*, 11-18. https://doi.org/10.1016/j.jveb.2008.08.006

Yeater, D. B., Dudzinski, K. M., Melzer, D., Magee, A. R., Robinett, M., Guerra, G., Salazar, K., Bolton, T., & Hill, H. M. (2024). Training the concept of innovate in dolphins (*Tursiops truncatus*) is both creative and cognitively stimulating. *Animals, 14*, 896. https://doi.org/10.3390/ani14060896

Supplemental Materials

Method

Subjects

Todd, Layla, and Snickers had previous training with the "create" cue and participated in another cognitive study in which they were asked to repeat self-performed actions on cue, including actions performed in response to the "create" cue (Scagel & Mercado, 2022a). The experimental studies of action repetition involved training the dogs to perform several specific actions in response to verbal and gestural cues, as well as training them to repeat a recently performed action on cue. Todd and Layla were formally tested on their ability to repeat recently actions produced in response to the "create" cue in a small number of trials (four per dog) as a way to assess how flexibly they could apply the repeating rule. The cues were never combined during training sessions such that the four test trials represented the first four times the dogs had experienced these cues in combination. Both dogs successfully repeated untrained actions performed in response to the "create" cue. These tests and earlier training experiences are unlikely to have affected the dogs' responses to the "create" cue in the current experiment, except to the extent that the dogs became familiar with their trainers and formalized testing procedures; learning the "repeat" cue might also have increased dogs' attention to their own recently performed actions.

Layla and Gizmo participated in a prior study in which they were asked to classify sets of toys and other real-world objects as "same" or "different" (Scagel & Mercado, 2022b). The procedure of this task required training the dogs to perform two specific actions (placing paws on a chair or lying on a towel) to indicate whether the items were all the same or not. The dogs also learned to initiate their classifications of objects sets in response to a verbal cue "Choose." Both dogs successfully learned this task. These tests and earlier training experiences are unlikely to have affected the dogs' responses to the "create" cue, other than by acclimating them to directing their action based on a verbal cue and performing actions within formalized testing conditions.

Gizmo and Snickers had never been trained to innovate actions on cue prior to this study and none of the dogs had ever been tested on the variety of actions they produced in response to the "create" cue.

General Training Procedure

"Create" training sessions for Todd and Gizmo were conducted in their owners' homes in rooms with at least 10-15 feet of space within which to work, Layla was trained in the Neural and Cognitive Plasticity laboratory on campus at the University at Buffalo, and Snickers was trained in his owner's home as well as the lab. The only objects in the laboratory were a foam mat upon which dogs were trained and tested, four chairs, a desk, cabinets along two walls, and a water bowl. During training sessions for each dog, the trainer began teaching the "create" cue by attending to the dog's actions and rewarding any actions performed that were not associated with a cue. This could include moving a paw, turning the head, sneezing, vocalizing, touching an object with the nose, etc. Both voluntary actions, such as moving a paw, as well as potentially involuntary actions, like sneezing, were reinforced during training to increase the frequency of any actions being produced as trainers noticed that without any reinforcement, dogs tended to remain stationary in front of the trainer. Though some actions may have initially been performed involuntarily, reinforcing them increased the chances these or similar behaviors would be voluntarily produced in the future. Once the dog was reliably innovating actions, the "create" cue was added, and the trainer rewarded any action produced in response to the "create" cue that was not already associated with a specific cue and had not been offered previously within the current training session. Any actions deemed "novel" by the trainer were rewarded with a verbal "yes" and a small dog treat approved by the dog's owner. Actions the dog had already produced within the given training session were not reinforced. Occasionally, dogs seemed to become fixated on performing certain behaviors, particularly towards the end of training sessions. Novel behaviors performed after a bout of repeating previously performed actions were heavily rewarded with several treats and praise. To "inspire" new actions, objects such as cardboard boxes, chairs, or toys were sometimes introduced. Upon introduction of an object, dogs often investigated or interacted with the object, and new actions performed towards objects could then be reinforced to further encourage the production of new behaviors. Each training session lasted one hour and each dog typically was trained in two sessions per week. During each training session, "create" training continued as long as the dog was motivated to voluntarily participate. Any signs of disengagement, such as refusing to offer any behavior in response to the "create" cue, walking away from the training area, getting a drink of water, or indicating a need to relieve themselves resulted in a five to 10-minute break, depending on the dog's motivation. Typically, 40-45 minutes of each session was spent training "create." Layla and Todd underwent 19 training sessions prior to testing, Snickers underwent 18 training sessions prior to testing, and Gizmo underwent 17 training sessions prior to testing due to scheduling conflicts. Training was extensive, taking place over several months and relying on the availability of the dogs' owners.

Behavioral Innovation Testing

Test sessions consisted of 24 discrete trials, 16 of which were "create" trials. In the other eight trials, deemed "non-create" trials, each dog was asked to perform one of two actions they were familiar with and proficient at performing (see example test session schedule below).

Sample "Create" Test Session

Trial	Action	Correct? (1/0)	Description	Notes
1	Spin			
2	Rise			
3	Create			
4	Create			
5	Create			
6	Create			
7	Create			
8	Create			
9	Create			
10	Create			
11	Rise			
12	Spin			
13	Create			
14	Create			
15	Create			
16	Create			
17	Spin			
18	Spin			
19	Create			
20	Create			
21	Create			
22	Create			
23	Rise			
24	Rise			

Action Sequences within Test Sessions

When given the "create" cue, dogs would often perform multiple behaviors at the same time or in rapid succession. For instance, they could lift a paw while also turning their head and vocalizing all at once. This made counting only single behaviors nearly impossible and led to the decision to count all behaviors performed within the first five seconds immediately following the "create" cue. Additionally, a new variation on a previous behavior is still different than the previous behavior and would therefore be counted as "creative." A sequence that included a combination of previously performed actions and novel actions together would still be considered novel as long as that combination of actions had not been performed before. This is consistent with criteria used by Kuczaj and Eskelinen (2014b) when training three bottlenose dolphins to vary their behavior on cue.

Limitations of Real-Time Judging of Actions in Testing

Though the use of real-time judging of responses during testing can result in repeated responses garnering accidental reinforcement or truly novel responses missing reinforcement, the decision was made to allow the trainer to judge and choose whether to reinforce responses during test sessions to keep the training and testing procedures consistent for every subject and to avoid any possible loss of the dogs' motivation to continue to participate. Braslau-Schneck (1994) found that when dolphins were not reinforced for any responses during test sessions, the animals stopped responding to the "creative" cue over time. This decision to provide reinforcement in real-time is also consistent with other studies of animals' abilities to innovate behaviors on cue (Braslau-Schneck, 1994; Kuczaj & Eskelinen, 2014b; Lawrence et al., 2016; Pryor et al., 1969, Willgohs et al., 2022). Additionally, real-time judgement and reinforcement was used to ensure rapid reinforcement of actions to enhance learning acquisition and maintenance of the "create" response.

Controls to Limit and Assess the Possible Role of Social Cueing of Innovated Actions

Control tests were given to only Layla because the other dogs were either unable to come to the lab or were distressed by unfamiliar people, which would have hindered their performance with an unfamiliar trainer. We also had no reason to believe Layla's performance in the control condition would be somehow exceptional.

As there are no standardized methods for training or testing behavioral innovation, there are no standard control procedures. To avoid influencing the actions dolphins produced in response to a "creative" cue, bottlenose dolphins were not given reinforcement during testing for any responses to "creative" (Braslau-Schneck, 1994). Willgohs and colleagues (2022) did not employ any control measures with their dogs, as their subjects were solely trained and tested by dog owners who recorded and submitted videos of each session. Likewise, Lawrence and colleagues (2016), as well as Kuczaj and Eskenlinen (2014b) did not use any control measures when testing bottlenose dolphins' abilities to innovate behaviors on cue.

Scoring

Though the visual coding of actions is less controlled or standardized than the use of behavioral coding software, this is common practice when testing animals on their ability to produce "creative" responses on cue (Braslau-Schneck, 1994; Kuczaj & Eskelinen, 2014b; Lawrence et al., 2016; Pryor et al., 1969; Willgohs et al., 2022). This procedure can lead to disagreements in response characterization by different evaluators, but measures of inter-rater reliability can assuage these concerns.

Results

Layla

Due to experimenter error, Layla was tested in 97 total “create” trials rather than 96, and 48 non-create trials. Over these 97 “create” trials, Layla correctly produced responses that she had not previously performed within the current test session and were not already associated with a cue in 64 trials (66%) and produced an average of 10.7 (67%) “creative” responses during each test session out of a possible 16. Of all “create” trials, 30 (31%) consisted of actions or action sequences Layla had already produced within the current test session. Among these trials, Layla repeated an action/action sequence she had just performed in the preceding trial in eight trials (27%). Therefore, Layla only repeated a response from the preceding trial in eight of 97 “create” trials (8%). During the remaining three (3%) incorrect trials, Layla produced a behavior that was already associated with a cue (spin, down, or bow). Layla offered the correct behavior in all non-create trials.

Because she could feasibly produce multiple actions in response to a single “create” cue within the five seconds after the cue was given, Layla often produced a sequence of actions in response to the “create” cue before the end of a trial. If these sequences are divided into discrete actions, Layla produced a total of 212 actions in response to this cue over 97 “create” trials for an average of 2.2 actions per “create” trial. These 212 total actions were divided into 28 different classes of actions. Layla produced an average of 12.8 types of behaviors each test session. The complete list of actions Layla performed in response to the “create” cue is presented in Table S2 with the absolute and relative frequency of each action. The most frequent behavior produced was a bark (16%). Of the 28 different actions Layla performed, 25 were not associated with any previously trained cue. Layla’s creativity index score was 0.13, which was the proportion of all actions produced that were novel.

As there were six test sessions, there were six first “create” trials in the testing portion of this study. Of these, Layla produced a unique action/action sequence in all six of these first trials, meaning she did not begin any test session with the same action or action sequence she produced to begin another test session. Within first “create” trials, Layla performed eight different types of behaviors with a total of 14 actions produced. Of these 14 total actions, Layla performed a paw lift twice, bark twice, shift weight to hips twice, and foot taps four times, though none of these were combined into identical action sequences during initial trials. The remaining four types of behaviors were never performed more than once during initial “create” trials. The relative frequency of types of actions Layla performed during first “create” trials is shown in Figure S1.

The number of completely novel types of behaviors Layla performed per session that she had never produced in any prior test session is depicted in Figure 4. Most of the 28 types of actions she performed in response to the “create” cue were produced in the first session. During subsequent sessions, Layla only produced an average of three completely novel types of behaviors per session.

During training sessions, Layla offered some actions when given the “create” cue that she did not perform during testing. These included drinking water from a bowl, lying down on her side, rolling onto her back, jumping backwards, lifting the edge of a mat with her paw, high-stepping, moving to a seated position next to the trainer, and touching objects with her nose or paw.

Control Sessions

During control sessions, Layla was tested in 32 total “create” trials and 16 non-create trials over two test sessions. Of these, Layla performed actions or action sequences she had not already produced within the current session and that were not already associated with a cue in 17 trials (53%) and produced an average of 8.5 (53%) correct responses out of 16 possible “create” trials in each session. Of the 32 total “create” trials, Layla repeated responses already performed within the current test session in 11 (34%) trials. Of these 11 repeated responses, Layla repeated the response given in the preceding trial in four trials (36%).

In control trials, Layla did not perform any singular actions that were already associated with a cue. Layla performed each cued behavior correctly in all non-create trials.

As in other sessions, Layla often produced a sequence of actions in response to the "create" cue during control sessions before a trial was over. In total, Layla produced 46 actions in response to this cue over the 32 "create" trials. Of these, she produced 12 different kinds of behaviors and performed an average of 8.5 kinds of behaviors each control session. Of these 12 kinds of behaviors, none were already associated with a specific cue. Layla's creativity index score during control sessions was 0.26. Of the two first "create" trials during control sessions, Layla produced different responses to each, meaning she did not simply produce the same order of responses in each session.

These results provide strong evidence Layla was not relying on inadvertent social cues from a trainer to produce responses to "create." Even in the presence of a trainer who did not know what constituted a "correct" response to this cue, she innovated a variety of actions. It is important to note that this is the first time such a control condition has been implemented in any study of action innovation/improvisation in any species (including humans). Past studies universally have had trainers or test administrators requesting actions who were aware of the nature of the task and of the kinds of actions that would qualify as correct responses to those requests. Although unlikely, it is possible that every innovative action reported in those past studies was unintentionally cued by a trainer/test administrator. Given that Layla's creativity index during blind testing was twice what it was when she was tested by an informed trainer, her performance suggests that if anything inadvertent social cues during testing may have degraded her capacity to innovate actions. These findings increase confidence that past demonstrations of action innovation in children and non-humans provide a conservative estimate of their creative capacities and that they are not simply artifacts of experimental methods.

Todd

Due to experimenter error, Todd was tested in 98 total "create" trials rather than 96, and 48 non-create trials. Out of these 98 "create" trials, Todd correctly produced actions or action sequences that he had not previously produced within the current test session and were not already associated with a cue in 71 trials (72%) and produced an average of 11.8 (74%) correct "creative" actions/action sequences out of 16 possible "create" trials during each test session. Of the 98 "create" trials, Todd produced actions/action sequences he had already performed within the current session in 27 (28%) trials. Among these 27 trials, Todd repeated the action or action sequence he had just performed in the preceding trial in 11 trials (41%). Overall, Todd only repeated an action or action sequence from the preceding trial in 11 of 98 "create" trials (11%). Todd did not perform any discrete actions previously associated with a cue unless they were part of a sequence that included novel actions. Todd performed the correct behavior in all non-create trials.

Todd, like Layla, often produced a sequence of actions in response to the "create" cue before the end of a trial. When divided into discrete actions, Todd performed a total of 225 actions in response to this cue over 98 "create" trials for an average of 2.3 actions per "create" trial. Of these 225 total actions, Todd performed 34 different kinds of behaviors. Todd produced an average of 15.2 different kinds of behaviors each test session. The complete list of actions Todd produced in response to "create" is provided in Table S3 along with the absolute and relative frequency of each action. The most frequently performed behavior was a "head tip up" (18%). Of the 34 different actions Todd performed, 30 were not associated with any previously trained cue. Todd had a creativity index score of 0.15.

Of six first "create" trials, Todd produced a unique action/action sequence in every first "create" trial. In other words, he did not begin any test session with the same action or action sequence he produced in the first trial of another test session. Within first "create" trials, Todd performed nine different types of behaviors with a total of 10 actions produced. Of these 10 total actions, Todd performed a paw lift twice, though neither of these were combined into an identical action sequence during initial trials. The remaining eight types of behaviors were never performed more than once during initial "create" trials. The relative frequency of the types of behaviors Todd produced in first "create" trials is shown in Figure S2.

The number of completely novel kinds of behaviors Todd performed per test session that he had never produced in any prior test session is depicted in Figure 4. Fifteen of the 34 types of actions he performed in response to the "create" cue were performed in the first session. During subsequent sessions, Todd produced an average of 4.4 completely novel types of behaviors per session.

During training, Todd offered behaviors he did not produce during testing in response to the "create" cue. These included a "chirp" vocalization, picking up a toy with his mouth, standing on a toy, licking his lips, walking through the legs of a nearby chair, growling, jumping straight into the air, jumping over objects, standing on an object and pivoting around it, pulling an object with his paw, resting his head on the ground, snapping his teeth, scooting backwards while lying down, standing on his back legs and wiggling his body, rubbing the side of his face on an object, turning to face the opposite direction, and grasping an object with his front paws.

Gizmo

Gizmo was tested in 96 total "create" trials over six test sessions and 48 non-create trials. Out of the total 96 "create" trials, Gizmo correctly produced actions or action sequences that he had not previously produced within the current test session and that were not already associated with a cue in 70 trials (73%) and performed an average of 11.7 (73%) novel "creative" actions/action sequences during each test session out of the total 16 "create" trials. Of the 96 "create" trials, Gizmo produced actions/action sequences he had already performed within the current test session in 26 (27%) of all "create" trials. Among these, Gizmo repeated the action or action sequence he had just performed in the prior trial in 10 trials (38%). In total, Gizmo repeated an action or action sequence from the preceding trial in 10 of 96 "create" trials (10%). Gizmo did not perform any singular action previously associated with a cue unless the action was part of a sequence that included novel actions. He also performed correctly in 100% of non-create trials.

Like other dogs, Gizmo typically produced a sequence of actions in response to the "create" cue before a trial ended. When divided into discrete actions, Gizmo performed a total of 184 actions in response to "create" over 96 trials for an average of 1.9 actions per "create" trial. Of these 184 total behaviors, Gizmo produced 24 different kinds of behaviors (Table S4). Gizmo produced an average of 11.2 types of behaviors each test session. The complete list of actions Gizmo performed in response to "create" is listed in Table S4 along with the absolute and relative frequency of each action. Of the 24 types of actions Gizmo performed, 23 were not associated with any previously trained cue. Gizmo's creativity index score was 0.13.

Out of the six initial "create" trials Gizmo faced, he produced a different action/action sequence in all six first "create" trials. He did not start any test session with the same action or action sequence he had performed in the first trial of another test session. Across first "create" trials, Gizmo performed nine different types of actions. He did not perform any of these actions in more than one first trial. Figure S3 illustrates the relative frequency of types of actions Gizmo performed in initial "create" trials.

The number of completely novel type of actions Gizmo produced within a single test session that he had never performed in any prior test session is depicted in Figure 4. Of the 24 types of actions he performed in response to the "create" cue, Gizmo performed 10 in the first session and eight in the second. During the following sessions, Gizmo produced an average of 1.5 completely novel types of behaviors per session.

During training, Gizmo produced several types of actions he did not produce during testing. These included scratching himself with a back leg, lying on his back and writhing back and forth, jumping over objects, lying on his side and covering an eye with his front paw, jumping sideways, moving to a seated position behind the trainer, lying on the ground and crawling backwards, lying on the ground and crawling forwards, pushing or touching a box with his nose or feet, standing on a box, biting a box, walking around a box, and walking onto a dog bed and sitting on it.

Snickers

Snickers was tested in 96 total “create” trials and 48 non-create trials. Out of these 96 “create” trials, Snickers correctly produced actions or action sequences that he had not previously produced within the current test session and were not already associated with a cue in 78 trials (81.25%) and produced an average of 12.8 (80%) correct “creative” actions/action sequences out of 16 possible “create” trials during each test session. Of the 96 “create” trials, Snickers produced actions/action sequences he had already performed within the current session in 13 (13.5%) trials. Among these 13 trials, Snickers repeated the action or action sequence he had just performed in the preceding trial in 6 trials (46%). Overall, Snickers only repeated an action or action sequence from the preceding trial in 6 of 96 “create” trials (6%). During the remaining five incorrect trials, Snickers responded with discrete actions that were already associated with a cue (down, rise). Snickers performed the correct behavior in 47 of 48 non-create trials (97.92%).

Snickers often produced a sequence of actions in response to the “create” cue before the end of a trial. When divided into discrete actions, Snickers performed a total of 283 actions in response to this cue over 96 “create” trials for an average of 3 actions per “create” trial. Of these 283 total actions, Snickers performed 22 different kinds of behaviors. Snickers produced an average of 11.7 different kinds of behaviors each test session. The complete list of actions Snickers produced in response to “create” is provided in Table S5 along with the absolute and relative frequency of each action. The most frequently performed behavior was a “bow” (15%). Of the 22 different actions Snickers performed, 19 were not associated with any previously trained cue. Snickers had a creativity index score of 0.08.

Of six first “create” trials, Snickers produced a unique action/action sequence in every first “create” trial. In other words, he did not begin any test session with the same action or action sequence he produced in the first trial of another test session. Within first “create” trials, Snickers performed eight different types of behaviors with a total of 12 actions produced. Of these 12 total actions, Snickers performed a down four times, though none of these were combined into an identical action sequence during initial trials. The remaining seven types of behaviors were never performed more than once during initial “create” trials. The relative frequency of the types of behaviors Snickers produced in first “create” trials is shown in Figure S4.

The number of completely novel kinds of behaviors Snickers performed per test session that he had never produced in any prior test session is depicted in Figure 4. Ten of the 22 types of actions he performed in response to the “create” cue were performed in the first session. During subsequent sessions, Snickers produced an average of 2.6 completely novel types of behaviors per session.

During training, Snickers offered behaviors he did not produce during testing in response to the “create” cue. These included growling, tipping his head up, tilting his head to the side, scratching his ear, chasing his tail, whining, drinking, and licking objects.

Limitations of the Experimental Approach

The number of dogs tested in this study is smaller than in several other studies of dogs’ cognitive capacities and multiple reviewers of this manuscript questioned whether a sufficient number of subjects were tested to justify any scientific claims about the innovative capacities of dogs. The current data cannot support any broad claims about how innovative capacities vary across dogs of different breeds, sexes, training histories, or ages. Undoubtedly, individual dogs will vary in their learning capacities and innovativeness as a function of these factors and many others, including the skills of their trainers. Our experimental design was not intended to survey or explain such individual differences or to test a specific hypothesis about the mechanisms that might enable dogs to self-select actions. Our purpose was instead to establish whether it was possible for dogs to learn to respond to a specific cue by producing a wide variety of untrained actions. Anecdotal reports (especially in publications by Karen Pryor) suggested that this should be possible, but to our knowledge our study is the first to explore this capacity in a controlled context using methods previously used to test another species (dolphins). Our sample size matches that used in the earlier dolphin study and can be interpreted similarly. The exploratory nature of the experiment and the subjective aspects involved in identifying innovative actions do lead to limitations in what can be inferred from the results; several such limitations are discussed below.

One limitation of research on "creativity" in general is that there is still disagreement on the definition of "creativity," both conceptually and operationally (Dudzinski et al., 2018; Kaufman & Kaufman, 2014; Kuczaj, 2017; Shevlin, 2020). Karwowski and colleagues (2016) state that "creativity" is "a human capacity to produce ideas and products that are both novel and useful or appropriate." Others have described it as "acting in an open and nonpredetermined manner in ways that generate novelty" (Patterson & Mann, 2015). Bateson & Marton (2013) defined "creativity" as "having a new idea" and "innovation" as "applying that idea in a productive and novel manner so that a better outcome is developed." Kaufman & Beghetto (2009) even divided creativity into a "four-C" model which included mini, little-, Pro-, and Big-c creativity, essentially qualifying certain types of acts as "more" or "less" creative and reserving the entire category of Big-c creative acts to human societies. When defining the "creativity" of species other than humans, the characteristics of creativity have been described as fluency, flexibility, originality, and elaboration (Kaufman & Kaufman, 2004; 2014; Kaufman et al., 2011). In each study of non-human "creativity," researchers must operationalize the term to determine whether their animals have been "creative" or not. Pryor and colleagues (1969) defined "creativity" for their dolphins as "only those actions...which have not been reinforced previously." Braslau-Schneck (1994) described it as "the production of original or novel motor behaviors. This can involve recombination of previously produced behaviors." Schusterman & Reichmuth (2008) defined vocal "creativity" for their walruses as acoustic responses that were obviously different from those emitted before. Subtle changes in vocalizations did not count as novel.

The current study could only examine whether dogs could learn to be "creative" as it was defined in this methodology: producing behaviors not already produced in the current session that were not already associated with a cue. By this metric, Layla, Todd, and Gizmo did learn to "be creative." We describe the task the dogs learned as innovating actions on cue, in part because we did not attempt to evaluate the originality of any of the actions offered by dogs. Others might describe the task as involving behavioral improvisation (Shevlin, 2020). The current study could not definitively determine whether dogs learned to actively avoid performing recent or trained behaviors or to randomly vary their actions as rats and pigeons have done in the past (Morris, 1989; Neuringer, 1991; Schwartz, 1982). Which strategy an individual is using can be difficult to identify because multiple situations can lead to the production of a variety of behaviors. Still, the finding that dogs rarely repeated actions just performed in the previous trial and almost never responded with actions already associated with a specific cue indicates dogs likely did avoid performing certain actions in response to the "create" cue. Dogs also were more likely to repeat responses from earlier trials within the second half of test sessions rather than the first half, indicating that as a greater number of trials had passed, dogs either "ran out of ideas" or may have struggled to remember what they had done earlier in the session. If dogs were performing actions randomly, they should have been equally as likely to repeat an action from the just-preceding trial as they were to repeat an action from earlier trials, and they also should have performed some actions from their trained repertoire. Additionally, the fact that dogs generally defaulted to certain types of actions that they combined in different ways and that each dog had a particular type of action they seemed to prefer over others suggests they were not randomly selecting actions.

It could be argued that the dogs were not selecting actions at all, but merely learned to associate the "create" cue with states of distress or confusion that led to reflexive reactions (displacement activity, attempts to escape, etc.) that naturally were more variable, for example through processes of classical conditioning. This scenario is unlikely because: (1) trainers worked with the dogs extensively prior to testing, before the dogs learned to consistently produce a variety of actions in response to the cue, and thus had extensive experience with identifying actions that the dogs produced in response to confusion or frustration; (2) prior to learning the task, when presumably any confusion and frustration was near its peak, the dogs were more likely to repeatedly offer specific trained actions than to perform untrained actions; and (3) reflexive reactions to states of confusion and frustration fall within a relatively constrained set that would be expected to vary systematically either within or across test sessions.

The types of actions the dogs produced in response to the “create” cue depended on the objects available in their environments. During training sessions, the trainer occasionally introduced objects such as boxes, dog toys, paper bags, or chairs to the training context. When these objects were available, dogs offered behaviors like sniffing the objects, touching, pushing, or tapping the objects with their paws, biting or picking up the objects in their mouths, jumping over the objects, and even standing on top of and pivoting around the objects. Because not all the dogs had the same objects available to them, no objects were used in testing sessions, limiting the types of actions they might have produced. These observations indicate that the “creativity” of an individual may depend on their environment, as objects have their own affordances, or action possibilities that can be performed with a certain object based on its physical features (Montesano et al., 2008).

In everyday situations, individuals often have more time to think of creative actions and having this “incubation” period often seems to help spark creativity. To avoid influencing the dogs’ responses, the trainer remained still and did not express emotion during testing, simply saying the word “yes” and providing a treat if a response was correct. During training, the structure of trials was at the discretion of the trainer and trials could occur in rapid succession, with few pauses. Though this data was not recorded, it is possible less time between trials during training could have lowered the likelihood dogs would forget recently offered responses. If a dog produced a completely novel action during training, the trainer excitedly praised the dog and offered a “jackpot” of multiple treats. When Alex, an African grey parrot (*Psittacus erithacus*), was trained to innovate vocal utterances, excitement from trainers seemed to encourage him to continue attempting new words (Pepperberg, 2015). The fact that it was up to the trainer’s discretion to stop a trial to reinforce the dog could have influenced the number of actions each dog produced within a trial. In the future, an additional experimenter could be used to decide when to provide reinforcement, though this could result in delayed reinforcement. During test sessions, the trainer remained emotionless to prevent influencing the dog’s behavior, but this could have potentially stifled creativity. Sometimes, dogs appeared frustrated during training by lack of reinforcement for incorrect responses, but after being continually probed to “create,” seemed to try a new action. Similarly, a rough-toothed dolphin being reinforced for producing new actions during shows appeared to show signs of frustration after cycling through her typical repertoire of behaviors and subsequently appeared to have a “breakthrough” of creativity (Pryor et al., 1969). Pacific walruses (*Odobenus rosmarus divergens*) trained to produce varied vocalizations typically repeated a vocalization several times, and only shifted their acoustic response after appearing frustrated from lack of reinforcement (Schusterman & Reichmuth, 2008). Perhaps some frustration helps encourage a change in behavior. Because of the setup of test sessions, Layla, Todd, Snickers, and Gizmo did not appear to show signs of frustration, as the presence of intermittent “non-create” trials meant they never needed to wait long to receive an “easy” instruction and receive reinforcement. Additionally, while the trainer did their best to remain still, it is possible dogs could have detected inadvertent movements made by the trainer during trials that could have influenced the type of action each dog produced. In the future, studies could examine how dogs perform in this task without a trainer in view.

Table S1

Behaviors to be Performed by Dogs in "Non-Create" Trials

Dog	Behavior	Description
Todd	Spin	Turning in a circle
	Pose	Sitting up on back feet
Layla	Spin	Turning in a circle
	Rise	Standing up on back feet
Gizmo	Twirl	Turning in a circle
	Down	Lying down on belly
Snickers	Down	Lying down on belly
	Rise	Standing up on back feet
	Spin	Turning in a circle

Table S2

All Behaviors Layla Produced in Response to the "Create" Cue Within Test Sessions

Actions Layla Produced in Response to "Create"	Behavior Category	Frequency	Relative Frequency
Bark *	Vocalization	33	0.156
Foot taps *	Paw movement	27	0.127
Head turn to side *	Head movement	26	0.123
Paw lift *	Paw movement	16	0.075
Groan/grumble *	Vocalization	16	0.075
Walk backwards *	Walking movement	11	0.052
Partial bow *	Bow	10	0.047
Step side to side *	Walking movement	8	0.038
Side step *	Walking movement	7	0.033
Slap floor with paw *	Paw movement	6	0.028
Soft bark/woof *	Vocalization	5	0.024
Forward hop *	Hop	4	0.019
Side hop *	Hop	3	0.014
Head shake *	Head movement	2	0.009
Spin	Spin	2	0.009
Grunt *	Vocalization	1	0.005
Scoot backwards on rump *	Shift position	1	0.005
Shift weight to hip while sitting *	Shift position	1	0.005
Full bow	Shift position	1	0.005
Down	Shift position	1	0.005
Slow paw wave *	Paw movement	1	0.005
Step forward *	Walking movement	1	0.005
Sniff ground *	Head movement	1	0.005
Touch nose to back paw *	Head movement	1	0.005
Body shake *	Shake	1	0.005
Scratch floor with paw *	Paw movement	1	0.005
Lick self *	Head movement	1	0.005
Tap feet while sliding into a down position *	Paw movement	1	0.005
Total = 28		212	1.000

Note. Asterisks indicate behaviors that were not already associated with a cue.

Table S3

All Behaviors Todd Produced in Response to the "Create" Cue Within Test Sessions

Actions Todd Produced in Response to "Create"	Behavior Category	Frequency	Relative Frequency
Tip head up *	Head movement	40	0.178
Paw lift *	Paw movement	27	0.12
Head circle *	Head movement	14	0.062
Hop in place *	Hop	13	0.058
Reach paw forward *	Paw movement	13	0.058
Side hop *	Hop	13	0.058
Touch trainer's knee with paw *	Touch	12	0.053
Sneeze	Vocalization	9	0.04
Stand on back legs *	Shift position	9	0.04
Hop with front legs while sitting *	Hop	8	0.036
Huff *	Vocalization	7	0.031
Skip *	Hop	6	0.027
Step forward *	Walking movement	5	0.022
Rest chin on trainer's knee *	Touch	5	0.022
Bark	Vocalization	4	0.018
Put both front paws on trainer's knee *	Touch	4	0.018
Side step *	Walking movement	4	0.018
Hop forward *	Hop	3	0.013
Partial down *	Shift position	3	0.013
Move head side to side *	Head movement	3	0.013
Walk to object *	Walking movement	3	0.013
Wave both front legs *	Paw movement	3	0.013
Sniff ground *	Head movement	3	0.013
Foot stomp *	Paw movement	3	0.013
Head turn *	Head movement	2	0.009
Down	Shift position	1	0.004
Nod head up and down *	Head movement	1	0.004
Hop on back legs *	Hop	1	0.004
Jump over trainer's knee *	Hop	1	0.004
Push off trainer's leg with paw *	Touch	1	0.004
Walk around trainer's knee *	Walking movement	1	0.004
Touch notebook with paw	Touch	1	0.004
Sway side to side *	Shift position	1	0.004
Touch trainer's knee with nose *	Touch	1	0.004
Total = 34		225	1.000

Note. Actions with an asterisk were not previously associated with any cue.

Table S4

All Behaviors Gizmo Produced in Response to the "Create" Cue Within Test Sessions

Actions Gizmo Produced in Response to "Create"	Behavior Category	Frequency	Relative Frequency
Single paw lift *	Paw movement	43	0.234
Sneeze *	Vocalization	18	0.098
Down	Shift Position	15	0.081
Reach paw forward *	Paw movement	14	0.076
Step forward *	Walking movement	10	0.054
Head shake *	Head movement	10	0.054
Forward lunge *	Shift position	10	0.054
Partial down *	Shift Position	7	0.038
Scoot *	Shift Position	6	0.032
Lift one paw then other *	Paw movement	6	0.032
Touch trainer's hand with nose*	Touch	5	0.027
Hop forward *	Hop	5	0.027
Woof *	Vocalization	4	0.021
Shift weight to hip *	Shift position	4	0.021
Lie on side *	Shift position	4	0.021
Stretch forward while lying down *	Shift position	4	0.021
Walk backwards *	Walking movement	4	0.021
Stand *	Shift position	4	0.021
Huff *	Vocalization	3	0.016
Touch trainer's knee with paw *	Touch	3	0.016
Touch trainer's knee with nose *	Touch	2	0.011
Sniff *	Head movement	1	0.005
Outstretch back legs while lying down *	Paw movement	1	0.005
Side step *	Walking movement	1	0.005
Total = 24		184	1.000

Note. Actions with an asterisk were not previously associated with any cue.

Table S5

All Behaviors Snickers Produced in Response to the "Create" Cue Within Test Sessions

Actions Snickers Produced in Response to "Create"	Behavior Category	Frequency	Relative Frequency
Bow*	Shift position	41	0.145
Down	Shift position	40	0.141
Paw lift *	Paw movement	34	0.12
Foot tap *	Paw movement	34	0.12
Shift forward *	Shift position	28	0.099
Wipe feet *	Paw movement	19	0.067
Shuffle paws *	Paw movement	18	0.064
Sit	Shift position	14	0.049
Half down *	Shift position	10	0. 035
Crouch *	Shift position	9	0. 032
Hop *	Hop	9	0.032
Step forward *	Walking movement	7	0.025
Touch trainer's leg with paw *	Touch	5	0.018
Sneeze*	Vocalization	3	0.011
Shuffle backwards *	Walking movement	3	0.011
Rise	Shift position	2	0.007
Huff *	Vocalization	2	0.007
Head turn *	Head movement	1	0.004
Paw slap *	Paw movement	1	0.004
Groan *	Vocalization	1	0.004
Lean back *	Shift position	1	0.004
Head down *	Head movement	1	0.004
Total = 22		283	1.000

Figure S1

Relative Frequency of Types of Actions Layla Produced to the First "Create" Trial of Each Test Session

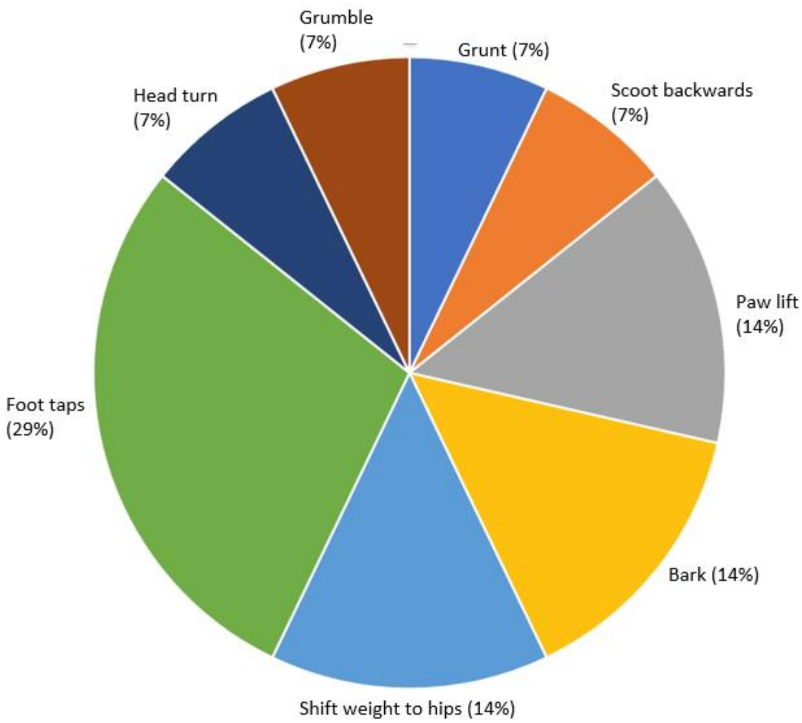


Figure S2

Relative Frequency of Types of Actions Todd Produced in the First "Create" Trial of Each Test Session

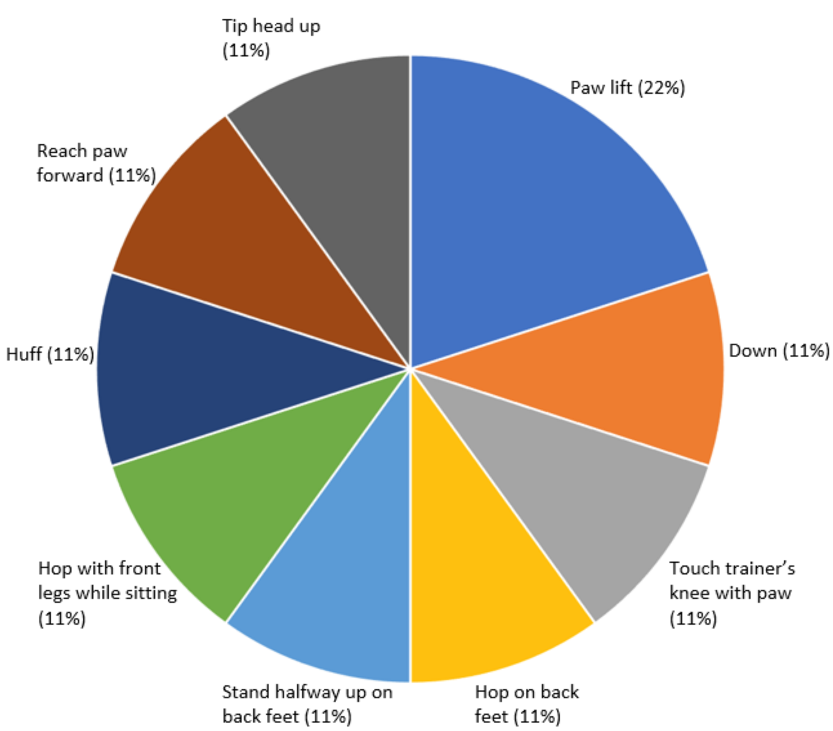


Figure S3

Relative Frequency of Types of Actions Gizmo Produced in the First "Create" Trial of Each Test Session

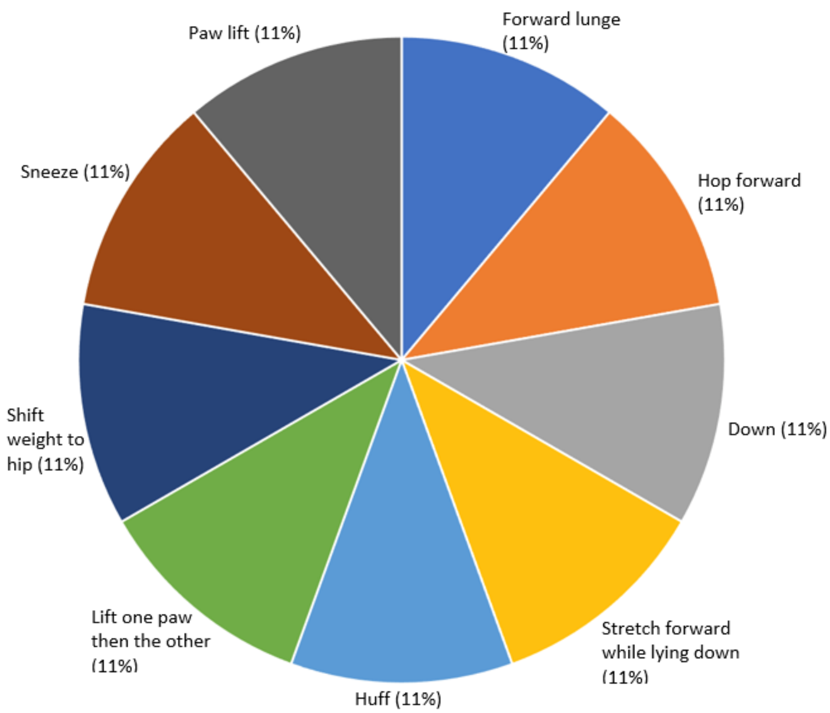


Figure S4

Relative Frequency of Types of Actions Snickers Produced in the First "Create" Trial of Each Test Session