Searching high and low: Domestic dogs' understanding of solidity

[1] Julia Espinosa
[1] Department of Psychology, University of Toronto, Sidney Smith Hall, 100 St George Street, Toronto, ON, M5S 3G3, Canada.

[2] Emma C. Tecwyn
[2] Department of Psychology, Birmingham City University, Curzon Building, 4 Cardigan Street, Birmingham B4 7BD, UK

[3] Daphna Buchsbaum
[3] Cognitive, Linguistic & Psychological Sciences, Brown University, Metcalf Research Building, 190 Thayer Street, Providence RI, 02912

This is the in-press version of our manuscript accepted for publication with *Animal Cognition.*

Correspondence concerning this article should be addressed to Julia Espinosa, Department of Psychology University of Toronto, 100 St George Street, Toronto, ON M5S 3G3, Canada.

E-mail: jhespino@umich.edu

ORCID: 0000-0003-0780-2762

Abstract

Physical reasoning appears central to understanding how the world works, suggesting adaptive function across the animal kingdom. However, conclusive evidence for inferential reasoning about physical objects is limited to primates. We systematically tested a central feature—understanding of solidity—in domestic dogs, by adapting a validated procedure (the shelf task) previously used to test children and non-human primates. Dogs watched a treat dropped into an apparatus with a shelf either present (treat landing on the shelf) or absent (treat landing on the bottom surface) and chose where to search for it (above or below the shelf). Across four studies (*N*=64), we manipulated visual access to the treat trajectory and apparatus interior.

Dogs correctly inferred the location of treats using physical cues when the shelf was present (Study 1), and learned rapidly when visual cues of continuity were limited (Study 2), and when the shelf was absent (Study 3). Dogs were at chance when the apparatus was fully occluded, and the presence and absence of the shelf varied across trials within subjects, and showed no evidence of learning (Study 4). The findings of these four studies suggest that dogs may be able to make some inferences using solidity and continuity and do not exhibit proximity or gravity biases. However, dogs did not always search correctly from Trial 1, and failed to search correctly when the rewarded location varied within-subjects, suggesting a role for learning, and possible limits to their ability to make inferences about physical objects.

Introduction

Humans and nonhuman animals (hereafter animals) must reason about the physical properties of objects in order to successfully navigate their environment. Physical reasoning appears to be fundamental for understanding how the world works, and for abilities such as tool use and cause and effect reasoning. It therefore stands to reason that knowledge of basic features of physical objects such as *solidity* (solid objects cannot pass through each other), *continuity* (objects cannot jump between locations, they can only travel along connected paths in space and time), and *gravity* (objects that are dropped fall downward), may have evolved in response to the need to effectively interact with the environment, an ecological pressure that is shared across many species (Spelke, 2000). Evidence from developmental research indicates that humans are particularly adept at understanding the physical world, making inferences about the physical properties of objects from early infancy (e.g., Spelke, 2003; Spelke & Kinzler, 2007). Infants as young as 5 months appear sensitive to violations of physical laws like solidity and continuity, looking longer at impossible events where objects appear to pass through solid barriers (Baillargeon & Graber, 1987; Spelke, Breinlinger, Macomber, & Jacobson, 1992).

In fact, one predominant theory in developmental psychology argues that physical reasoning is among a small set of ‘core knowledge’ domains, which are particularly central to success across a wide variety of species and ecological niches. These domains of core knowledge are argued to be innate and evolutionarily ancient, in which case they should be found not only in human infants, but in an evolutionarily broad range of non-human species (Carey & Spelke, 1996; Spelke, 1994; Spelke & Kinzler, 2007). However, research so far has not conclusively established whether knowledge of physical properties and associated inferential reasoning is shared across species, and thus empirical evidence that it is an evolutionarily ancient ability is lacking. Most of the research on physical reasoning has focused on human children and non-human primates (hereafter primates), and has not examined more distantly related species, who nonetheless operate in the same physical environment. Further, the few studies of physical reasoning in non-primate species have used varying methodologies, producing mixed results that make cross-species comparisons challenging. In the present research, we explore whether domestic dogs make inferences on the basis of physical properties of objects, employing measures validated in other species for testing understanding of physical properties in domestic dogs.

Perceiving physical properties

Reasoning about the physical world is often taken for granted in fully developed adult humans (for a review of naïve physical reasoning in adults see Kubricht, Holyoak, & Lu, 2017). While there continues to be debate within the field of developmental psychology as to whether these intuitive domains of knowledge are innate, there is strong evidence that an intuitive knowledge of physical properties is early emerging. Human infants show an ability to make inferences about the physical world with little-to-no experience nor formal education, suggesting that this is an intuitive cognitive ability (Spelke, 1988; Spelke, Lee, & Izard, 2010; and Carey & Spelke, 1996), one that is key for developing an understanding of the environment and making sense of the world (e.g., Gopnik & Meltzoff, 1987; Wellman & Gelman, 1992; Wimmer & Perner, 1983).

A prerequisite for the ability to make inferences that guide object-directed behavior is having an awareness of the physical properties of objects on a basic, perceptual level (Cacchione & Rakoczy, 2017). This perceptual level of physical awareness has been extensively investigated in infants and primates using violation-of-expectation (VOE) looking time paradigms. In these tasks, longer looking time from the subject is interpreted as “surprise” at witnessing object interactions that defy physical laws like solidity and gravity.

Perceptual awareness of physical violations is present in human infants (for a review see, Baillargeon, 2002; Spelke, 1994; Valenza, Leo, Gava, & Simion, 2006), and primates tested in solidity tasks with comparable VOE methods responded similarly to infants (Santos & Hauser, 2002; Santos, Seelig, & Hauser, 2006), indicating that perceptual awareness of physical properties may not be unique to humans, but may be shared by closely related primates. This was demonstrated in rhesus macaques (*Macaca mulatta*) tested with a shelf apparatus where they watched apples drop behind an occluder (Santos & Hauser, 2002). Similar to infants, macaques looked significantly longer when the apple was revealed underneath the solid shelf (as though passing through it and violating solidity, unexpected outcome) than when the apple was shown on top of the shelf (expected outcome). These findings indicate that like infants, rhesus macaques appear to have a basic understanding of how principles of solidity, continuity, and gravity govern object interactions.

VOE tasks have also been adapted to domestic dogs, where results showed that dogs reacted with surprise when a solid screen rotated through the space where they had previously seen a solid bone (Pattison, Miller, Rayburn-Reeves, & Zentall, 2010). Demonstrating similar visual behavior to infants and rhesus macaques by looking longer at impossible events, dogs seemed to detect that a violation of physical properties had occurred, which suggests that they are aware of physical properties (such as solidity) at a perceptual level. Dogs’ expectations about objects has been further supported by eye-tracking studies that suggest dogs anticipatorily track object movement in a horizontal plane (Völter, Karl, & Huber, 2020) and form expectations about object interactions relating to solidity and support, indicating an awareness of physical properties (Völter & Huber, 2021).

Acting on physical inferences

For visual awareness of objects' physical properties to be useful, an individual needs to be able to make appropriate inferences based on that perception, in order to guide their behaviour. However, an interesting dissociation between looking and acting measures has frequently been observed in studies of physical reasoning in humans and animals alike (e.g., Keen, 2003; Santos & Hauser, 2002). On the one hand, when shown an impossible outcome like an object that is dropped magically appearing below a solid shelf, instead of on top of the shelf, individuals "look" surprised by the physical property violation (measured by VOE). On the other hand, when required to search for an object that has been displaced, subjects often act as though unaware of physical properties, or as if relying on naïve biases such as a gravity bias (i.e., falling objects travel straight down regardless of intervening barriers) and perseverate in searching the lowest point directly underneath where the item was dropped (e.g., Cacchione & Burkhart, 2012; for a discussion across species see Tecwyn & Buchsbaum, 2018).

Looking time studies with infants suggest a perceptual awareness of physical properties emerges shortly after birth, but this is complicated by findings in toddlers who, when allowed to search for vertically displaced items, do not behave as though they are guided by an understanding of solidity until at least 2.5-3 years of age (Hood et al., 2000). Hood and colleagues (2000) examined toddlers' understanding of physical properties using a shelf design, where dropped toys could land on top of the shelf when it was present, or at the bottom of the apparatus when the shelf was absent. Two-year-olds searched the top and bottom locations at chance whether or not the shelf was present, while 2.5-year-olds searched correctly, appearing to factor in the solidity of the shelf as a barrier the toys could not pass though. These results suggest that awareness of solidity and continuity, while present early on, does not necessarily translate into more complex inferential abilities supporting goal-direct action until later in development.

As with human children, there are also mixed results for successfully searching for displaced objects on the basis of solidity in monkeys and great apes (Santos, 2004; Santos et al., 2006; Southgate & Gomez, 2006; Cacchione, Call, & Zingg, 2009). When presented with a horizontal motion-based search task, instead of searching for the object that rolled down a ramp in the cup at the end, macaques (*Macaca mulatta* and *Macaca arctoides*) showed a strong preference to search underneath the ramp itself, and this behaviour did not diminish with repeated test trials (Southgate & Gomez, 2006). In vertical drop scenarios similar to Hood & colleagues' shelf task, apes (*Gorilla gorilla*, *Pongo pygmaeus, Pan troglodytes*, *Pan paniscus*) were significantly above chance within 4-5 trials (Cacchione et al., 2009). Taken together, the rapid learning shown by some primate species could suggest a basic understanding of physical object properties, and an ability to translate visual cues of solidity and continuity into successful action, but it could also reflect learning of the correct search location within the experimental context. Overall, knowledge of solidity and continuity may not be a functional knowledge for directing behaviour for all primate species.

One explanation for the discrepancy between looking and acting across tasks in toddlers and primates is that there may be a naïve expectation about gravitational force overriding their representation of solid object events and influencing search behaviour. It has often been observed that in vertical motion tasks, subjects appear to have naïve expectations that dropped objects fall straight down (gravity bias), and search directly below the release point or at the lowest possible location (common marmosets (*Callithirix jacchus*), Cacchione & Burkart, 2012; toddlers, Hood et al., 1995; cotton-top tamarins (*Saguinus oedipus*), Hood et al., 1999; dogs, Osthaus, Slater, & Lea, 2003). Reorienting solidity tasks from vertical to horizontal object trajectory, and thus removing conflicting gravity information from the scenario, appeared to help performance in some cases (3-year-old children, Hood, Cole-Davies, & Dias, 2003; great apes, Cacchione et al., 2009). Subjects searched correctly for objects they had seen roll behind a screen, seeming to account for physical barriers in their expectation of where the object could be located. Thus, differences in performance on looking and searching paradigms may at least partly result from strong biases overriding behavior in vertical drop scenarios.

A handful of studies have examined non-primate species' ability to reason on the basis of solidity, primarily using the inclined screen task. In this task, subjects are presented with opaque screens that are either laid flat or set at an incline, with the goal of inferring the location of hidden rewards. These inferences are based on the understanding that a solid object located under one of the screens will alter the orientation of that screen (by propping it up), thus they should select the inclined over the flat screen. This task has produced mixed results across species, domestic horses (*Equus ferus caballus*) were unsuccessful in locating a human with a bucket of food behind a screen (Haemmerli, Thill, Amici, & Cacchione, 2018), whereas chicks (*Gallus gallus*) (Chiandetti & Vallortigara, 2011) as well as domestic pigs (*Sus scrofa domestica*) (Albaich-Serrano, Brauer, Cacchione, Zickert, & Amici, 2012) were able to select the correct screen in order to find food (pigs) or a small object (chicks), suggesting some knowledge of physical properties. All three studies adapted the paradigm to the species being tested, and thus without sufficient overlap in design or the animals’ prior experience with physical objects, it is challenging to conclude whether an ability to make choices guided by inferences about physical object properties such as solidity is present outside of primates and shared broadly across branches of evolution.

Physical inferences in domestic dogs

A potential promising avenue for evaluating physical reasoning across species lies in replicating and extending primate investigations in a non-primate species that has shared the same ecology as humans for the last several millennia: the domestic dog (*Canis lupus familiaris*). Dogs are evolutionarily distant from primates, but they have been selected for human-like social traits and often outperform nonhuman primates (e.g., chimpanzees) in tasks of socio-communicative reasoning (Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006). While examining physical inferences in dogs cannot entirely disambiguate the question of whether physical knowledge is ancient and broadly shared, or an ability acquired by dogs through co-evolution with humans and experience with human artifacts, it addresses the hypothesis that physical knowledge is used by a wide range of species and fills a gap in what we know about foundational cognitive abilities in dogs.

Dogs have been the subject of many studies examining social cognition, but their physical reasoning has received considerably less attention and remains underexplored. Despite the paucity of evidence, there is a pervasive belief that dogs have limited physical reasoning skills (for a review see Bensky, Gosling, & Sinn, 2013; Bräuer et al., 2006; Lampe, Bräuer, Kaminski, & Virányi, 2017). However, results are inconsistent, which may be partly due to the idiosyncratic methodologies seen across tasks. For example, only two published studies have directly examined dogs’ understanding of the physical property of solidity in search-based tasks. One study used a horizontal motion set-up similar to designs used with toddlers (Hood et al., 2003), but included a noise-based confound which may have aided the dogs in successfully solving the problem (Kundey, De Los Reyes, Taglang, Baruch, & German, 2010). In the second investigation, a complex, novel apparatus design was used, that required dogs to be trained on lever pulling prior to test. This may have increased the difficulty of the task, changing it from simply requiring inferences about solidity, to a more sophisticated if-then sequence of inferences about the outcomes of the dogs' own actions (Müller, Riemer, Range, & Huber, 2014). With only these two conflicting findings (in addition to the looking time study of perceiving physical properties, described above, Pattison et al., 2010), it is unclear if dogs can both perceive causal information about solidity and use solidity cues to make inferences that support actions.

Interestingly, despite some suggestions to the contrary (e.g. Kundey, Reyes, Taglang, Baruch, & German, 2010; Range, Möslinger, & Virányi, 2012), dogs do not seem prone to consistent gravity errors in the same manner as toddlers and cotton-top tamarins, when tested in vertical motion tasks such as the tubes task (used to study naïve expectations of gravity in multiple species, Hood et al., 1995). A first study using the tubes task in dogs found that dogs also initially searched as though guided by gravity expectations (Osthaus, Slater, & Lea, 2003). However, in both this study and a follow up investigation (Tecwyn & Buchsbaum, 2019), researchers found that the dominant strategy was to search the middle location instead of directly below where the reward had been released, and preferred the middle location over the location suggested by gravity, when the two were pitted against each other. Though the two studies reported some differences in the range of search strategies dogs utilized in the task, there was overall agreement that dogs did not appear to understand the containment properties of the tube in item displacement.

While gravity may not exert the strongest pull on dogs' search strategies, means-ends connectivity tasks (i.e., selecting the string that is connected to a reward out of other, unconnected strings) brought to light another simple heuristic potentially underlying dogs' behavior in physical inference tasks: a proximity bias. In these tasks, dogs appeared to choose strings based on reward proximity rather than the reward-string connectivity (Osthaus, Lea, & Slater, 2005), though there was some individual capacity to learn how to choose the correct string when researchers controlled for proximity of the reward (Riemer, Müller, Range, & Huber, 2014). Dogs were somewhat more successful in a support variation of the task, overcoming potential proximity biases and selecting the board with a reward on it over one with a reward placed next to it (Range, Hentrup, & Virányi, 2011). However, dogs did not show a proximity bias in the tubes task, and were not more likely to pick the search location closest to where the treat was released (Tecwyn & Buchsbaum, 2019). Given these various accounts of physical inferences across domains, there is no clear consensus on what abilities dogs display in these areas, or whether they are indeed making inferences or relying on simpler heuristics.

The present research circumvents the problems with prior canine solidity studies in two ways. First, we adapted a simple search-based task (the 'shelf' task) that has previously been used in primate and toddler studies of solidity and continuity (Hood, Carey, & Prasada, 2000; Santos & Hauser, 2002), allowing for a direct cross-species comparison of naïve physical reasoning using validated methods that did not require extensive pre-training over multiple visits. Second, we controlled for auditory and olfactory cues that could help dogs solve the task without reasoning about physical properties. We examined dogs' physical reasoning about solidity over the course of four studies, systematically manipulating the availability of visual cues about solidity and continuity of motion. Each successive study provided increasingly restricted cues, which allowed us to discern if dogs are factoring in solidity and continuity when making inferences about the location of displaced objects. Considering the results from means-ends studies of physical reasoning (Osthaus et al., 2005), we also explicitly tested if dogs based their search strategies on simpler heuristics such as reward proximity rather than inferring the outcome of solid object interactions.

Study 1: Locating dropped objects when continuity and solidity cues are available

Our first study examined two central features of physical reasoning – the understanding of solidity and continuity – in dogs, adapting the search-based shelf task used with toddlers and primates (Cacchione et al., 2009; Hood et al., 2000): Subjects watch an object fall into the occluded apparatus with the shelf present, and can then choose to search for the object either on or under the solid shelf. In the present study, the partially occluded apparatus provided visual cues of solidity (visible shelf on either side of the occluder) and continuity (falling treat that did *not* reappear in the gap beneath the shelf and was not seen falling towards the lower compartment). If dogs make intuitive inferences about solidity and continuity of motion, similarly to primates, they should infer that falling treats will continue moving along a path until encountering a solid barrier, and search on top of the shelf to find the hidden treat. Otherwise, they should search randomly (like young toddlers) or perhaps show a bias to search the incorrect bottom compartment as a default strategy that may be rooted in a gravity bias (though evidence for such a bias in dogs is lacking in the tubes task, Tecwyn & Buchsbaum, 2019) or possibly some foraging behaviour that may have been reinforced by finding food on the ground (as has been suggested for rhesus macaques, Southgate & Gomez, 2006).

Participants

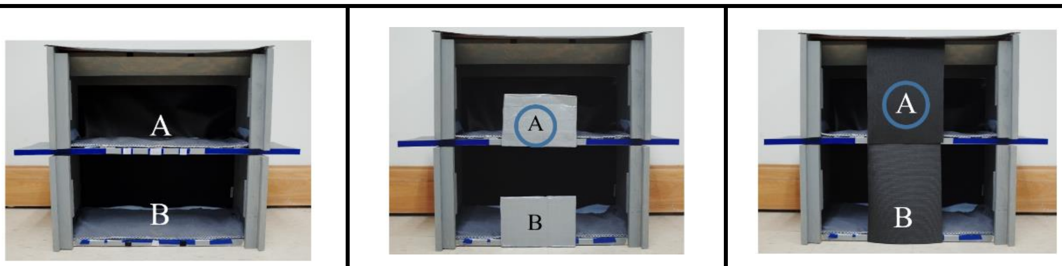
We recruited pet dogs from the Greater Toronto Area through online questionnaires advertised via social media and paper flyers. To participate, dogs were required to be at least six months old. No exclusions were made based on breed or training experience. Dogs were tested individually, and all data collection took place in the lab at the University of Toronto (for a full list of dogs in the study, see Supplementary Table S1). The study was approved by the University Animal Care Committee (UACC), and all procedures were in accordance university and federal ethics guidelines.

Sixteen dogs (F = 9, M_{age} = 3.71 years) were included in Study 1. Two additional dogs were excluded from analyses, due to experimenter error when baiting the apparatus during test trials (N = 1), and lack of motivation to complete training trials (N = 1). Lack of motivation was defined as failure to search the apparatus within a 30 second period after verbal release and encouragement to search for treats by the experimenter on two consecutive trials. Low motivation manifested as lying down at the start position and ignoring vocal encouragement, or walking away and lying down by the door of the room. In both situations the sessions were terminated to avoid distressing the dog through repeated attempts to re-engage them. We used a between-subjects design and dogs provided data for one of the four studies.

Materials and Procedure

The shelf apparatus used in our study consisted of a wooden frame and a removable shelf that generally resembled a short bookcase (Fig. 1A; see Supplement for details). The removable shelf defined two vertically aligned search locations within the apparatus (A and B, Fig. 1A). In Study 1, we used two small occluders that partially obscured the apparatus interior by attaching to the shelf and base of the apparatus (Fig. 1B). Given previous mixed results from dog physical reasoning studies, this modification lowered the point of entry to the task that, pending the outcome of dogs' performance, could be progressively raised in subsequent studies to closely replicate primate and toddler studies. These occluders were similar to the design used with 2-

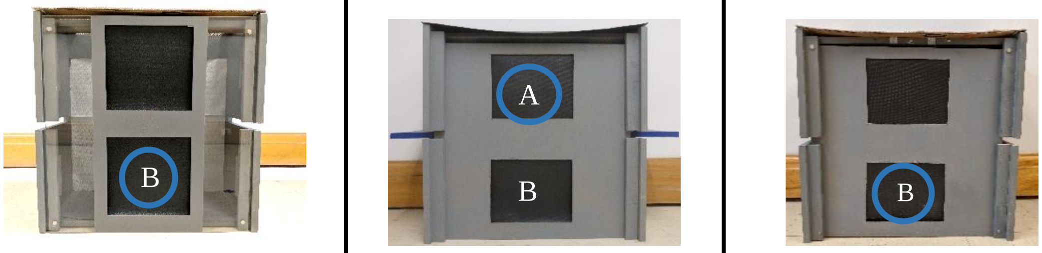
year-olds (Experiment 2, Hood et al., 2000) with one key difference: They did not span the height of the apparatus, and the spaces above each search location were visible, allowing dogs to see that the treat did not fall through the shelf.



A. Apparatus with removable shelf used during Familiarization trials.

B. Occluders with shelf in place, used in Study 1 test trials.

C. Occluders with shelf in place, used in Study 2 test trials.



D. Occluder with shelf removed used in Study 3 test trials.

E. Occluder used in Study 4 with shelf in place for 50% of trials.

F. Occluder used in Study 4 with shelf removed for 50% of trials.

Fig. 1 The shelf apparatus with the occluders used to systematically restrict visual cues of solidity and continuity across four studies, B – most, to E/F - fewest. The circled letters in B-F indicates the correct search location during test trials in that study/configuration. Panels E and F display the two configurations used in Study 4 (fewest visual cues) to test within-subject responses when the shelf was present (E) and absent (F).

Following the protocol used with apes (Cacchione et al., 2009), and toddlers (Hood et al., 2000; Exp 1 and 2), the shelf was kept in place throughout the session. Once the occluders were attached at the beginning of training, they remained in place until the test concluded. The small occluders were constructed of cardboard and covered in duct tape (height: 16cm, width: 24cm) which allowed them to be easily pushed aside by the dog’s muzzle while searching the apparatus. The top panel of the apparatus folded back for treats to drop inside where they would fall behind the occluder onto top of the shelf (Fig 1B; location A). We controlled for auditory cues by running a fan in the room, as well as extensively padding the search locations to absorb the sounds of dried liver treats landing. Additional treats were layered in the padding at each search location to control for odour cues should the dogs attempt to solve the task using olfactory information.

Dog guardians were present throughout the study. They were seated in chairs behind the dog and instructed to avoid social cuing through gaze or gestures (Supplementary Fig. S2). All researchers were trained to minimize body movements and to look at fixed locations on the walls or floor to eliminate the possibility of the Clever Hans Effect. Data collection and all phases of the study occurred in a single session in the lab (40-60 minutes).

Familiarization

Similar to procedures from solidity studies with macaque monkeys (Southgate & Gomez, 2006), each dog participated in a warm-up task at the beginning where they retrieved treats from both the top and bottom of the un-occluded apparatus (Fig. 1A). This allowed: 1) dogs to acclimate to the room and working on-leash with a trained handler; 2) dogs to learn to attend to the experimenter, orient toward the apparatus as a source of treats, and gain familiarity retrieving treats from either location; 3) us to see if dogs had an *a priori* location preference, measured by which location they searched first when both were baited simultaneously.

At the start of each familiarization trial, dogs were positioned by the handler approximately one meter in front of and facing the apparatus (Fig. S2). The female experimenter was seated to the side of the apparatus and showed two treats by extending her arms in front of locations A and B, then baited both locations simultaneously. The experimenter verbally released the dog to search and recorded the order in which the dog searched for the treats (top-bottom, or bottom-top), the handler recalled the dog to start and the sequence was repeated for a total of six trials (12 treats).

Training Trials

Training trials with the occluder began immediately after completing familiarization (Fig. 1B). Training trials served to: 1) familiarize dogs with the occluded apparatus; 2) teach dogs how to retrieve hidden food by using their head to push aside the occluder covering each search location; 3) ensure that the dogs recruited for the study had adequate object permanence prior to assessing physical reasoning (i.e., they could find a treat they observed being placed out of sight behind the occluder); and 4) that dogs were able to search both locations without exhibiting a location bias (see Supplement for location bias training procedure adapted from Seed, Call, Emery, & Clayton, 2009).

At the start of each training trial, the experimenter showed the dog one treat before placing it behind the occluder at either location A (on the shelf) or B (at the bottom of the apparatus), and showing the dog her empty hand while withdrawing from the apparatus. The dog was verbally released to search and coded as correct if they searched the baited location first, and incorrect if they searched the empty location first. Dogs were allowed to search exhaustively during training, receiving treats on every trial to motivate participation. Dogs passed training when they searched correctly on six consecutive trials within a maximum of 18 trials. This was designed as a sliding window, beginning with the first correct search and ending when the dog reached six correct searches in a row, searched incorrectly, or reached the 18 trials. Dogs that failed to search correctly six times in a row (not including location bias training) were excluded from the study and their data does not appear in subsequent analyses. None of the dogs in Study 1 failed to meet the training criterion or needed location bias training.

Test Trials

Dogs immediately began test trials after passing training. The objective of the test trials was to see if dogs used visual cues of solidity and continuity to spontaneously infer the location of a treat dropped through the top of the occluded apparatus. As in previous phases of the experiment, the dog was on-leash with the handler 1 meter in front of the apparatus and watched the experimenter present a single treat before dropping it into the apparatus and replacing the top panel. The experimenter verbally released the dog to search and they were allowed to search exhaustively, thus were rewarded on each trial regardless of which location they searched first (similar exhaustive search procedure used by Hood, Hauser, Anderson, & Santos, 1999, cotton-top tamarins; Osthaus, Slater, & Lea, 2003, dogs; Tecwyn & Buchsbaum, 2019). For Study 1, the shelf was in place for all 12 trials and location A was correct.

Data coding and analysis

We live-coded dogs' behaviour during the session and video recorded for offline reliability recoding. We coded searches to the top or bottom compartments when the dog's nose crossed the front plane of the apparatus in one of those locations, and scored performance as correct when the dog searched the baited location prior to the empty location. Our primary analysis examined whether as a group dogs searched the correct location at rates above chance, and whether their performance improved across trials. The sample size is conventional in comparative cognition research, and is appropriately powered to detect moderate-to-large effect sizes over repeated trials (power ≥ .80 for detecting average correct performance of 70% or greater relative to a chance level of 50%, e.g., Rosner, 2015, variance estimates taken from dogs' performance on similar 2-alternative forced choice tasks in our lab). Since dogs provided repeated responses, we performed generalized linear mixed models (binomial distribution:

correct/incorrect) with trials nested in dogs (analyses conducted using *lme4* package in *R*, Bates et al., 2015) to assess changes in performance over test trials. We used the same model approach to examine familiarization data within each Study to look for evidence of a pre-existing biases to either the top or the bottom location (binomial: bottom = 0, top = 1). As a secondary post hoc assessment of initial performance, we pooled data across studies to obtain sufficient power to examine Trial 1[1] (see analysis of initial performance section of Study 4) and used binomial tests to test Trial 1 at the Study level. A random subset of 25% of test sessions from each study were re-coded from video by naïve research assistants to confirm live coding results (*n* = 4 dogs/study, all test trials), with 100% agreement between live coding and recoding. Data and R analysis scripts are available on OSF.

Results & Discussion

During the familiarization trials, 12/16 dogs showed a qualitative preference for searching one of the locations and chose it first on at least 5/6 trials (bottom = 9 dogs; top = 3 dogs), the remaining four dogs exhibited random choices in no particular order. Overall, dogs did not choose to search the top or bottom first significantly more often than chance in familiarization, $B = -1.73$, $SE = 1.00$, $z = -1.73$, $p = .08$, suggesting that there was no strong *a priori* location bias. Since both locations were baited simultaneously and rewarded equally while in full view, dogs were not being reinforced to search in a particular location (nor was there a

---

[1] We recognize that testing performance on Trial 1 is often considered the most stringent assessment of inferential ability (though see e.g., Povinelli & Henley, (2020) for a recent argument against privileging Trial 1 data), resource limitations did not permit collecting a sample large enough to power first trial analyses and thus we elected to use repeated trials in order to increase power with a conventional sample size. For instance, to detect at least 70% correct performance at 80% power with $\alpha = 0.05$ on Trial 1 alone, a minimum of 47 dogs per study would be required, instead of the current 16 dogs per study.

correct first location choice in this phase). This phase allowed us to discern if there was a pre-existing preference at the group level that could influence searching in occluded trials.

Dogs completed took an average of 9.8 (*SE* = 1.03) training trials with occluders before meeting pass criterion (6 correct in a row) and moving to the test trials. For the majority of dogs, the range of training trials required to meet the criterion was from six ($n = 5$) to 16 ($n = 1$) (see Table S2). One dog was inadvertently given a single additional training trial to reach criterion and move to the test phase (completed a total of 19 training trials) (See Table S3 for list of possible errors).

In Study 1, when provided with salient visual cues about solidity and continuity of motion, dogs were able to search correctly from Trial 1 in a version of the shelf task adapted from humans and primates. We found that overall, across test trials dogs searched the correct location (top) first significantly more often than chance (chance = 6/12), $B = 4.04$, $SE = 1.37$, $z = 2.95$, $p = .003$ (Fig. 2). Looking at individual performance, 13/16 dogs scored significantly above chance across their test trials (≥ 10/12 trials correct), exact binomial test: $p \leq .039$. Looking at the performance of dogs on the first test trial, 87.5% (14/16 dogs) searched the correct (top) location first, exact binomial test: $p \leq .004$. These results suggest that dogs may be able to make inferences involving a solid barrier and may be using visual cues of solidity to infer the location of hidden treats dropped behind the occluder, even prior to learning over repeated trials. We next looked at changes over the course of the experiment by including trial number as a predictor, and found that dogs' performance improved slightly, suggesting some additional learning across trials, $B = .25$, $SE = .11$, $z = 2.39$, $p = .02$ (Fig. 3).

Together these results suggest that dogs make inferences using the physical properties of the shelf in this task. However, the small occluders used in Study 1 offered abundant visual cues about both solidity and continuity to aid dogs in their search, so it is unclear if dogs are using one or both of these physical properties to solve the task. It is also possible that dogs solved this task without any grasp of solidity at all, but through a simpler heuristic that led them to search the location closest to the point where the experimenter released the treat. We examine the role of continuity versus solidity cues in Study 2 by using a larger occluder design that did not show the treat passing through the top compartment and address the possibility of a search heuristic more directly in Study 3.

Study 2: Locating dropped objects when only solidity cues are available

In Study 2 we aimed to identify if dogs were able to use solidity in isolation to solve the shelf task. In a close replication of a search-based study with toddlers (Experiment 2, Hood et al., 2000) we extended the occluders to span the height of the apparatus (Fig. 1C), restricting continuity of motion cues to the space above the apparatus. This prevented dogs from watching the treat drop towards the shelf, and critically, did not allow them to see whether the treat re-emerged below the shelf. Thus, dogs needed to infer the location of the hidden treat based only on the presence of the shelf—i.e., based on knowledge of solidity. If dogs were relying on continuity of motion cues to locate the treat in Study 1, then their performance should decline compared to Study 1. This occluder design also allows for a closer comparison to the performance of children and other primate species, who were tested with limited continuity cues in comparable designs (Hood et al., 2000; Cacchione et al., 2009).

Participants

Sixteen dogs (F = 8, M_{age} = 3.44 years) took part in Study 2. Ten additional dogs were excluded due to (1) failure to pass the training trials (n = 4) (i.e., they did not search correctly on six consecutive trials within a maximum of 18 trials), (2) experimenter error during test (*n* = 1), and (3) lack of motivation to complete training trials (*n* = 5).

Materials and procedure

We used the same experimental procedure described in Study 1, with the following changes. To hide the treat's falling trajectory and final resting place, we used long occluders (width: 24cm) attached with Velcro to the top of apparatus frame and to the removable shelf that hung down and covered the entire middle section of the apparatus (Fig. 1C). The occluders were made of vinyl that allowed dogs to easily push them aside to access the interior of the apparatus. This maintained visual cues to solidity while removing continuity of motion cues. As in Study 1, the shelf was present throughout training and test.

Results & Discussion

Relative to Study 1, the continuity cues available in Study 2 were significantly reduced, requiring dogs to use the available information about solidity to solve the task. In the familiarization phase of Study 2, 12/16 dogs showed a qualitative preference for one of the locations by searching it first on at least 5/6 of trials (bottom = 9 dogs; top = 3 dogs), the remaining four dogs searched randomly. Similar to Study 1, we did not observe a statistically significant difference in searches to either location during familiarization, $B = -2.08$, $SE = 1.13$, $z = -1.83$, $p = .07$, which indicates no strong pre-existing bias to the top or bottom location. Dogs in Study 2 completed an average of 8.5 ($SE = 0.9$) training trials before passing on to the test trials. The number of trials required for individuals to progress to the test phase ranged between six (*n* = 9) and 16 (*n* = 2) trials (see Table S2).

We found that overall, across test trials dogs searched the correct top location significantly above chance, *B* = 2.75, *SE* = 0.45, *z* = 6.13, *p* < .001 (Fig. 2), indicating that across trials dogs were able to locate the treat correctly. As individuals, 94% (15/16) dogs searched correctly significantly above chance across test trials (≥10/12 trials correct, exact binomial test: *p* ≤ .039). Looking at the performance on the first trial, 69% (11/16) of dogs searched the correct (top) location first, exact binomial test: *p* ≤ .21. As a group, dogs were not significantly different from chance on their first attempt. We next examined changes over the course of the experiment by including trial number as a predictor, and as in Study 1 we found a significant improvement over trials, *B* = 0.36, *SE* = 0.11, *z* = 3.16, *p* = .002, (Fig. 3), suggesting that dogs may have learned the correct location across the test trials. However, in addition to associative learning, the rapid rate at which dogs learned in this task may also indicate some basic understanding of physical properties, a point that we will return to in the general discussion (see Supplement for additional analyses).

Taken together, dogs' performance is comparable to that of children 2.5 years of age and older (93% correct, Hood et al., 2000), and apes (59% correct trial 1, 66% correct trial 4; Experiment 1, Cacchione et al., 2009). In this study dogs also outperformed two-year-old children who searched either randomly (40% correct, Experiment 1, Hood et al., 2000), or persisted in searching the bottom location (20% correct, Experiment 2), and rhesus monkeys, who also persisted in searching the bottom location in a similar task (Southgate & Gomez, 2006).

Overall, dogs in Study 2 were able to accurately locate dropped treats in the shelf task even when continuity information was limited, providing suggestive evidence that they may use or rapidly learn to use solidity cues. Nonetheless, it is possible that dogs succeeded through learning the correct search location, rather than through the use of physical reasoning, a point we will return to in the general discussion. Besides associative learning across trials, there is another, simpler explanation for dogs' success in this task—they could have succeeded by searching based on proximity, as the top location was closest to where they last saw the treat before it disappeared into the apparatus (cf. Cacchione et al., 2009 for a similar alternative explanation for apes' performance). While there was no evidence of a pre-existing bias to the top location, quite the opposite in fact, reliance on a proximity heuristic when observing the treat being dropped could explain the successful search behaviour seen in Studies 1 & 2, when the top location was consistently correct. If dogs factor in the solid properties of the shelf as a barrier between the top and bottom locations, they should be able to correctly locate the treat both when a solid barrier is present *or* absent.

Study 3: Locating hidden items behind a long occluder with the shelf removed

To address the possibility that dogs used a simpler rule (searching on the basis of proximity) to solve the previous two studies instead of reasoning about physical properties, in Study 3 we removed the shelf during the test trials so that the bottom location was the correct search location (Fig. 1D). If dogs are indeed reasoning about solidity, then they should infer that the treat will fall all the way to the bottom when the shelf is absent and search that location. If, on the other hand, they use a simpler proximity-based strategy, then we would expect them to persist in searching the top location as observed in previous studies.

Participants

Sixteen dogs (8 F, $M_{\text{age}} = 3.63$ years) participated in Study 3. Seven additional dogs were excluded because (1) they did not meet the training criterion ($n = 4$), (2) they were location biased to the bottom and did not respond to location bias correction training (*n* = 1), or (3) because they were physically large enough to successfully reach the bottom of the apparatus and get treats when searching through the top opening of the occluder (*n* = 2). Three dogs showed a bias to the bottom location during training trials and required location bias correction training (see procedure details in Supplementary Information), and learned to search flexibly within the training session and their data were included in final analyses.

Materials and procedure

We used the same experimental procedure described in Study 1 with the following changes. As in Study 1, the shelf was in place for familiarization and training to screen for location preference and teach dogs how to search both the top and bottom. The shelf was then removed prior to the first test trial and the bottom location was correct for test trials (Fig. 1D). The occluder for Study 3 (adapted from Exp. 2 Hood et al., 2000) covered the entire front of the apparatus and had two vertically aligned windows with vinyl flaps that allowed access to the top (A) and bottom (B) compartments. It was constructed from a sheet of clear acrylic with a grey-painted center. This left an eight-inch transparent panel on either side of the search locations to allow the dogs a clear view of the interior of the apparatus. This held constant the visual information about solidity (shelf visible on either side of the occluder) as well as the lack of continuity information from Study 2. Like the procedures of Studies 1 & 2, the occluder was not removed during test trials.

Results & Discussion

In the familiarization phase of Study 3, 13/16 dogs showed a qualitative preference for one of the locations, searching it first on at least 5/6 trials (bottom = 8 dogs; top = 5 dogs), the three remaining dogs chose randomly. As in the two previous studies, while some individual dogs had a preference for the top or the bottom, overall across dogs there was no significant difference in which location dogs searched first during familiarization, $B = -0.64$, $SE = 0.55$, $z = -1.15$, $p = .25$, which suggested an absence of strong location biases. In the training phase, dogs in Study 3 completed an average of 11.69 ($SE = 1.45$) training trials before progressing to the test trials. The number of training trials required for dogs to reach criteria for the test phase ranged between the minimum possible six trials ($n = 3$) and maximum eighteen trials ($n = 2$) (see Table S2). Three dogs received location bias training to correct a bottom bias in training trials, learned to search flexibly, and were included in final analyses.

Overall, across test trials, we found that dogs searched the bottom location significantly above chance, $B = 2.45$, $SE = 0.69$, z = 3.55, $p < .001$ (Fig. 2), suggesting that dogs were able to find hidden treats when the shelf was absent (bottom location correct). As individuals, 75% (12/16) of dogs searched correctly above chance across test trials (≥ 10/12 trials correct, exact binomial test: $p \leq .039$). Looking at initial responses on Trial 1 we observed that 63% (10/16 dogs) searched the correct (bottom) location first on Trial 1, exact binomial test: $p \leq .45$, which suggests that they may have initially failed to account for the absence of the shelf. As in Studies 1 and 2, dogs in Study 3 showed significant improvement in search accuracy over trials, $B = 0.48$, $SE = 0.11$, z = 4.22, $p < .001$ (Fig. 3), suggesting a rapid learning rate in the task (see Supplement).

Though initially not different from chance, dogs' performance on this variation of the task improved rapidly and is notably better than that of 2-year-old children (0% correct, Experiment 2, Hood et al., 2000; However, toddlers in this study received a somewhat different familiarization, where they were familiarized to a scenario where the top location was correct, before being tested with the shelf absent). To our knowledge, no other species or age groups have been tested on the shelf-absent search task. Taken together, the observed effects across these three studies may suggest some level of physical inferences in dogs, as they are able to solve the shelf task with the shelf either present or absent, or exhibiting persistent location biases.

Study 4: Locating Objects Dropped into a Fully Occluded Apparatus

Given dogs' success in Studies 1-3, for the final study in this series, we drew inspiration from previous solidity research (Experiment 4 of Hood et al., 2000; Mixed trial block, Kundey et al., 2010; Müller et al., 2014) and tested both types of events (shelf present *and* absent) within-subjects. Toddlers tested on this variation of the shelf task were at chance (Hood et al., 2000), but unlike the dogs in the present study, they also exhibited perseveration errors when required to reason about either the presence or absence of the shelf (Experiment 2, Hood et al., 2000). Randomly changing the correct location from trial-to-trial is a more challenging task, and would require dogs to search flexibly, tracking and remembering the changing configuration of the apparatus, and reasoning about the resulting presence or absence of the shelf to correctly infer where to search. Previous studies of solidity in dogs also varied the correct location within-subjects (Kundey et al., 2010; Müller et al., 2014), and while dogs may have succeeded or failed those tasks for various reasons, they nevertheless demonstrated flexible searching between the possible locations in a vertical motion task. Dogs were comparatively successful in the previous three studies of the current research. Nevertheless, the question remains if dogs would demonstrate knowledge of solidity and continuity in the shelf task when the correct location varies, and search flexibly depending on the apparatus configuration. When the shelf is inserted, they should search the top location, and when the shelf is absent they should search the bottom location.

Participants

Sixteen dogs (F = 8, M_{age} = 5.2 years) took part in Study 4. Four additional dogs were excluded because (1) they did not pass the training trials (*n*= 1), (2) they did not respond to location bias correction training (*n* = 1), or (3) because they were large enough to get treats from the bottom of the apparatus when searching through the top window (*n* = 2). Five dogs required location bias correction training (see procedure details in Supplementary Information), learned to search flexibly within the training session and their data were included in final analyses.

Procedure

The same materials and procedure described in Study 1 was used with the following modifications. The occluder for Study 4 was opaque and covered the front of the apparatus by sliding into the apparatus frame (Fig. 1E & F). When the shelf was inserted, the ends were visible on either side of the occluder to provide visual cues of solidity. The interior could be accessed through vertically aligned ‘doggy doors’ covered with vinyl flaps (identical to Study 3). Continuity of motion cues were restricted to the ~8-inch space between where the experimenter released the treat and the top plane of the apparatus.

The occluder and shelf were in place for all training trials, but unlike previous studies, they were necessarily removed during the test trials to take out and put in the shelf. The shelf was present for six of the twelve test trials (Fig. 1E) and absent for the other six (Fig. 1F), following a predetermined pseudorandom sequence so that the correct location could be either the top or the bottom on a given trial but either location would not be rewarded on more than two consecutive trials.

When the shelf was in place the treat would land on top of the shelf and could be retrieved through the top compartment. When the shelf was absent the treat would fall to the base of the apparatus and could be retrieved through the bottom compartment. When removed, the shelf was placed within sight on the floor. At the start of each test trial, even if the configuration did not change, the experimenter removed the occluder and showed the dog the apparatus interior. With the shelf present, the experimenter reached through the open top and knocked on the shelf to emphasize its presence (similar procedure by Hood et al., 2000 with toddlers, and Santos & Hauser, 2002 with rhesus macaques). With the shelf absent, the experimenter slowly waved her hand back and forth inside the apparatus to draw attention to the empty space.

Results & Discussion

Study 4 tested dogs' understanding of solidity and their ability to apply it flexibly by varying the correct location between the top and bottom locations within dogs, using a fully occluded apparatus that limited visual cues of solidity and continuity. In the familiarization phase, 14/16 dogs showed a preference to go to one of the locations over the other, choosing it first on at least 5/6 trials (bottom = 13 dogs; top = 1 dog), and the remaining two dogs chose randomly between the locations. In contrast to Studies 1-3, during familiarization dogs searched the bottom location first significantly more often than expected by chance, $B = -9.76$, $SE = 3.08$, $z = -3.17$, $p = .002$, which points to a pre-existing bias to search the bottom in this sample of dogs, before seeing the occluder or beginning the training and test.

Dogs completed an average of 12.31 (*SE* = 1.06) training trials on the occluded apparatus (not including location bias correction trials) before meeting pass criterion and progressing to the test trials. The range of training required was between six (*n* = 3) and 18 (*n* = 6) trials (see Supplementary Table S2). All dogs (*n* = 5) that required location correction bias training before the test were biased towards the bottom compartment, in line with the results of the familiarization data analysis.

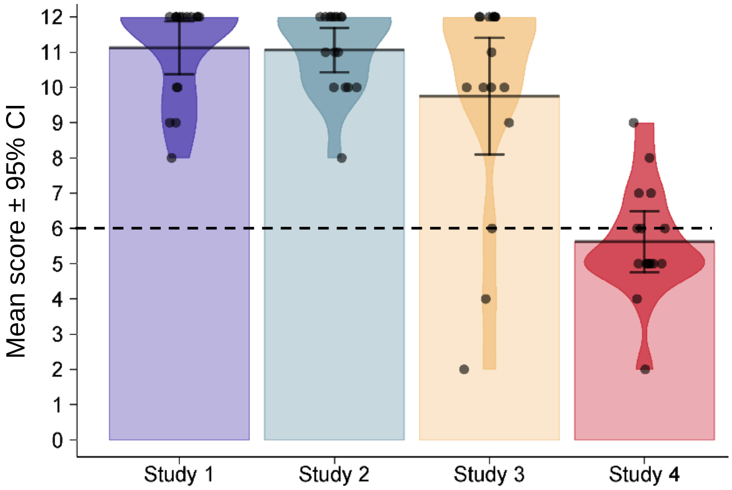


Fig. 2 Mean correct score by Study with 95% CI. Distribution of individuals' totals indicated by the superimposed beans and dogs within studies represented by the black dots. Chance performance is indicated by the dashed horizontal line.

In Study 4 when dogs were required to switch between searching the top and bottom of the fully occluded apparatus, we found that they did not search as though they accounted for the changing presence or absence of the shelf, and overall, across test trials, dogs did not search correctly significantly above chance, *B* = -0.13, *SE* = 0.14, *z* = -0.87, *p* = .39 (Fig. 2). At an individual level, none of the dogs searched correctly significantly above chance across test trials (highest score 9/12 correct), further, looking at the group we did not see evidence of improvement across test trials, *B* = -0.02, *SE* = 0.04, *z* = -0.38, *p* = .71 (Fig. 3). Interestingly, this failure was not due to a location bias (like the toddlers in Experiments 1-3 of Hood et al., 2000). Even though familiarization and training data suggested that dogs were *a priori* biased to the bottom location, over test trials dogs searched randomly, but flexibly, without showing a location preference by repeatedly searching one place first. While there was not an overall location bias, looking at the performance of dogs on the first test trial we observed that a majority (14/16) were significantly more likely to search the bottom location first, exact binomial test $p = .004$. Only two dogs searched the top on the first trial, once when the shelf was present, and once when it was absent.

This performance difference between task variations could indicate a limit of dogs' ability to reason about solidity in a goal-directed way, or it could suggest that dogs' success in Studies 1-3, where the correct search location did not change across trials, was primarily the result of reinforcement learning rather than physical reasoning. However, it should be observed that procedural features of Study 4 were also more demanding than in the previous three studies. Visual access was also restricted more here than in the previous studies, and, given their performance, it is possible that dogs may not have visually registered the presence of the shelf (just the ends were visible) and searched without noticing its presence or absence. Notably, toddlers tested with a similar design in which the shelf was either present or absent between blocks of trials also showed chance performance, and we are unaware of other species tested on a version of the task where the shelf changes location across trials. By removing and re-inserting the shelf between trials, dogs were required to track multiple moving pieces (shelf, occluder, hand waving). Thus, it is possible that dogs may not have remembered the configuration of the apparatus interior by the time a treat was dropped on a given trial.

Interestingly, the two previous search-based solidity tasks carried out with dogs that tested responses with either horizontal motion (Kundey at al., 2010) or a tube pulling task respectively (Müller et al., 2014), also reported that dogs searched flexibly between locations. Together with results of the present study, this underscores that dogs are not relying on simple heuristics like proximity (which would bias them towards the top location) or a gravity (which would bias them towards the bottom) to guide behavior, even when the task demands are high, and despite an *a priori* bias towards the bottom location evident in the familiarization phase and trial 1 of test trials of this study. Additional investigation is required to understand why dogs failed this version of the shelf task and pinpoint whether features of the apparatus, procedure, or limits of reasoning may have impeded their ability to flexibly factor solidity and continuity

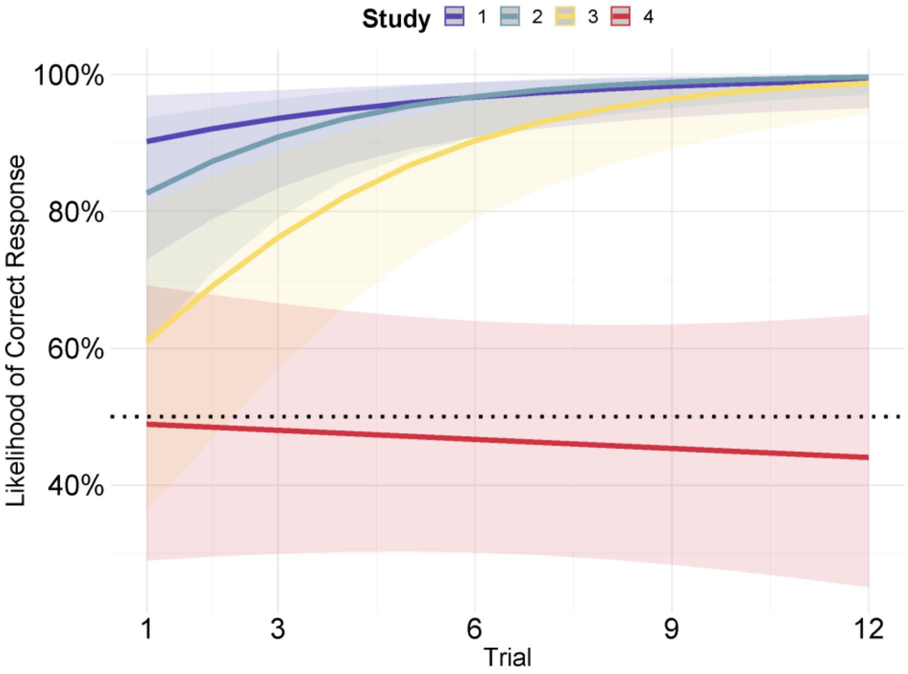


Fig. 3 Model estimated average correct responses by study, surrounded by confidence bands. Chance performance of 50% is indicated by the dashed line.

information into their search, or whether their success could be accounted for by learning the correct search location, without a broader understanding of physical properties.

Comparison Across Experiments

We compared all studies in a post-hoc analysis to see if overall performance in the first three studies differed significantly from performance in Study 4, and that performance in Studies 1-3 did not differ from each other. Performance of dogs in Studies 1-3 was significantly better than performance of dogs in Study 4, *B* = 1.45, *SE* = 0.57, z = 2.55, *p* = .01, confirming observations from each of the individual studies above (Figures 2-3): Dogs in the first study were able to search correctly from Trial 1, dogs in Studies 2-3 were able to learn the correct location within a few trials (see Supplemental Analyses), while dogs in Study 4 were at chance.

In a second post-hoc analysis we examined whether location (top or bottom) influenced performance when only solidity cues were available, contrasting Studies 2 and 3. We failed to observe a significant effect of study, *B* = 0.58, *SE* = 0.50, z = 1.15, *p* = .25, or a significant trial X study interaction, *B* = -0.02, *SE* = 0.08, z = -0.28, *p* =. 78. Overall, it appears that apparatus configurations did not influence dogs’ search behavior and a majority of dogs were able to accurately search the correct location first based on solidity information.

Analysis of initial performance across studies

To directly address the question of initial performance with appropriate statistical power, we ran an exploratory analysis on the first trial of Studies 1-3. Given the significant difference in performance between the first three studies and Study 4, the latter was not included in the mega-analysis. We pooled data across Trial 1 of the three studies and fitted a linear mixed-effects model with binomial distribution, data nested within Study. We found that on Trial 1 of the three partially occluded studies, dogs were significantly more likely to search correctly, $B = 0.99$, $SE = 0.32$, $z = 3.05$, $p = .002$, and there was no effect of Study on Trial 1 performance, $\chi^2(2) = 2.97$, $p = .23$. Taken together, dogs searched successfully on Trial 1 across the three partly occluded studies, and their performance was not significantly affected by whether the correct location was the top (Studies 1 and 2) or the bottom (Study 3). This supports the interpretation that, in addition to learning, dogs have some knowledge of physical properties that allows them to make inferences and guide behavior in a novel scenario such as the shelf task. However, this should be regarded as promising but preliminary evidence of physical inferences, as we discuss below.

General Discussion

Across four studies we examined solidity and continuity inferences in dogs using the shelf task. Our results from the first three studies suggest that dogs are sensitive to visual cues of solidity and continuity, supporting the theoretical claim that knowledge of physical properties is likely to be evolutionarily broadly shared, and not limited to primates (e.g., Carey & Spelke, 1996; Spelke, 1994; Spelke & Kinzler, 2007). Teasing apart the effects of evolutionary adaptation and domestication can make it challenging to ascertain whether physical reasoning is evolutionarily ancient, and traces back to a common ancestor of primates and dogs, or whether it is a by-product of more recent selection pressure on dogs to share the human environment. Though we cannot conclusively determine which of these paths is correct based on our findings, they do support that at least some non-primates have an awareness of physical object properties. Further explorations with other non-primate domesticated species, and with wolves and other canids who share a recent common ancestor with dogs, are an important direction to explore to determine evolutionary origins of physical reasoning.

To closely follow the methods used with other species and facilitate cross-species comparison, we first carried out the shelf task with dogs similar to how it was presented to toddlers, and then extended our examination of dog physical reasoning by systematically restricting visual cues and testing dogs' understanding of solidity and continuity separately. When salient cues of solidity and continuity were available and dogs were required to factor in the presence of a horizontal barrier, they searched correctly on top of the shelf from trial 1 (Study 1), with some individuals overcoming a pre-existing bias to the bottom location (evident in familiarization), and indicating some ability to account for the presence of the shelf in the context of locating dropped treats. When salient continuity cues were limited (Study 2), dogs were not different from chance on Trial 1, yet improved rapidly in the space of a few trials (see Supplemental analyses). We explored the possibility that simpler strategies such as searching in the location closest to the place they last saw the treat (a proximity bias) were responsible for dogs' success, and found that a majority of dogs learned to search the bottom correctly when the shelf was absent (Study 3). When tasked with reasoning based on the presence *and* absence of the shelf within subjects (Study 4), dogs showed flexibility in searching both locations but were overall at chance in their accuracy and did not improve in search accuracy over trials. It is entirely possible that dogs' overall success in studies 1-3 is the result of learning the correct search location over trials, without requiring physical knowledge. However, dogs' rapid success in these studies when compared to other location-based search tasks, suggests that dogs have some ability to perceive and make inferences about the properties of solidity and continuity of objects, or very rapidly learn to do so (Studies 1-3), and that their failure to solve the alternating task (Study 4) may be caused by a limit in their inferential reasoning or possibly other aspects of the procedure.

Notably, in Study 4, dogs were equally likely to search both locations (like toddlers in Exp. 4 in Hood et al., 2000), supporting that their poor performance may have resulted from the cognitive load of tracking the changing apparatus, and perhaps the reduced visibility of the shelf, rather than a gravity bias towards the bottom, or a proximity bias towards the top. Interestingly, as the occluders increased in size over studies and restricted visual access to cues of solidity, dogs' performance decreased, suggesting that salient visual cues play an important role in cognitive tasks, even for animals with strong olfactory abilities (Gadbois & Reeve, 2014). Taken together, like young children, apes and perhaps some species of monkeys, the findings suggest that dogs have some knowledge about fundamental physical properties such as solidity and continuity, but their ability to make inferences based on physical properties has limitations that should be defined in future work.

The results of our study contradict previous search-based task on solidity in dogs (Müller et al., 2014), and support that they may be sensitive to causally relevant information in a physical reasoning scenario, such as the shelf task. By using a simplified design relative to the task used by Müller and colleagues, we provided evidence of some ability to make inferences about solidity in dogs that is comparable to young children's performance on a similar task (2-2.5-year-olds, Hood, et al., 2000). This confirms the positive findings about perceptual awareness (Pattison et al., 2010) and horizontal motion tasks with dogs (Kundey et al., 2010), while also controlling for auditory confounds in the latter that could clue dogs in to the location of the reward and allowed them to succeed without necessarily making inferences using object properties.

We observed an improvement in search accuracy over trials in Studies 2 and 3. With a consistently rewarded location, dogs could have learned to associate the top or the bottom respectively with treats and searched there based purely on a learned reward-location association. However, dogs' rapid improvement rate in this task could also suggest a form of reinforcement learning supported by underlying understanding of physical properties. This alternative explanation is supported by evidence from another physical reasoning task that also probed dogs' use of solidity and containment when searching for dropped treats (the tubes task). Dogs failed to learn to search the correct location for a reward dropped down an opaque diagonal tube more often that expected by chance across 12 trials (Tecwyn & Buchsbaum, 2019). This was despite the fact that, as in Studies 1-3 of the present study, the correct search location remained fixed across trials and so was consistently reinforced. In other species, the process of learning to associate purely arbitrary stimuli (i.e., without any underlying physical or causal knowledge) typically takes many trials (Civelek, Call, & Seed, 2020; Seed, Hanus, & Call, 2011). In a similar vein, evidence from several spatial search tasks in dogs suggests that a substantially greater number of trials than 3 or 4 is required to learn object-reward associations (Fiset, Gagnon, & Beaulieu, 2000; Mongillo et al., 2013; Tecwyn & Buchsbaum, 2019). Consequently, evidence of successful performance on the shelf task within the first few trials (see supplementary analyses) suggests that dogs' searching behavior may be supported by some underlying knowledge of physical properties, instead of an entirely learned reward-location association.

To minimize the opt-out rate in the task, we did not penalize dogs for incorrect choices and they received rewards on each trial by being allowed to search exhaustively (similar to studies with apes, Cacchione et al., 2009; and dogs, Tecwyn & Buchsbaum, 2019). One could speculate that, under associative accounts, exhaustive searching could potentially lead to superstitious pattern formation, resulting in searching the apparatus in a fixed pattern (e.g., top-bottom with bottom rewarded), rather than refining search behaviour to focus on the rewarded location (we observed this superstitious behavior pattern in one dog in Study 3). Further, to look directly at whether a gradual improvement across trials (consistent with associative learning) is driving dogs' overall success in Studies 1-3 as opposed to rapid problem solving supported by an underlying knowledge of physical objects in the world, we ran an exploratory analysis of dogs' performance on pooled first trial data to generate sufficient power for statistical analysis. This analysis showed above chance performance by dogs from trial 1 without a driving effect from any particular Study. While we cannot rule out associative learning of the search location without additional control experiments, taken together, we believe that this comparatively high rate of success in initial searching suggests that some ability to make inferences about physical properties supports dogs' problem solving in the shelf task, prior to gaining experience with the apparatus and learning reward-location associations. Nonetheless, particularly given dogs' performance on Trial 1 of Studies 2 and 3, these findings should be regarded as promising but preliminary evidence that requires further investigation before making strong conclusions about dogs' ability to reason about physical properties of objects.

To address this conundrum of learning versus inferential ability, future work could present dogs with a variation of the search-based shelf task involving a “trick” shelf that would allow the reward to fall through the barrier to the bottom, seemingly defying physical properties. If dogs show comparatively greater difficulty learning to search the physically impossible location, or fail to learn it over the course of 12 trials, this would provide support for their use of solidity cues in determining reward location. In contrast, if they learned just as rapidly to search the rewarded location as in the present Studies 2-3 when it appears to be physically impossible, then this may imply that learning of the reinforced location is the primary mechanism by which dogs succeed and learn equally quickly regardless of the cues of solidity and continuity.

Related to the associative learning points discussed above, a limitation of our findings is that we only saw evidence of physical reasoning when the correct location was held constant and we systematically restricted visual cues the dogs may have used. This approach—breaking down the components of a task to test features of interest in isolation—has been used in infant studies where it has proved effective at identifying the limits of their reasoning (e.g., Hood, 1995). Given that older children and primate species, who like dogs succeed at search tasks similar to our Study 2 (Hood et al., 2000; Cacchione et al., 2009), but have not been tested in an alternating scenario where shelf presence/absence changed between trials, we do not know whether they would prove more successful under similar cognitive load. Future work investigating physical reasoning in dogs as well as other species could build on the foundation we have established by systematically reintroducing features of the shelf task, such as varying the rewarded location in partially occluded conditions (e.g., using the occluder from Study 3), in order to test within-subject reasoning about the presence *or* absence of solid barriers while avoiding potentially confusing aspects of the procedure that were introduced by Study 4.

A central challenge to comparative cognition is developing species-appropriate tasks while maintaining a direct link to previous studies to allow performance to be compared across species. This is a possible factor underlying poor performance and inconclusive evidence in many studies with non-primate species. Tasks optimally designed from a human viewpoint may not adequately consider the ecology and ways in which other animals perceive and interact with the world (Andrews, 2020; e.g., elephants use scent in quantity discrimination, Plotnik et al., 2019), thus not accurately testing cognitive processes. Another layer of complexity is added by idiosyncratic methods that emerge with each adaptation to a new species, making it difficult to evaluate the reason that different species succeed or fail on tasks lacking clear comparison points. In the current design, we prioritized task similarity between primates and dogs by using the shelf task, providing a direct link to compare species' performance over, creating another new physical reasoning task, designed specifically for canines. This is an important and necessary first step to maintain the link to prior work. Future work can use this base, and the rapidly growing body of work on canine visual attention (e.g., Byosiere et al., 2017; Völter & Huber, 2021; Espinosa et al., 2021) and olfactory information gathering (e.g., Bräuer & Belger, 2018) to present problems in a way that is ecologically suited for domestic dogs.

An area requiring particular additional investigation in order to support methodological development for canines, is dogs' understanding of simple spatial relations between objects. In Studies 3 & 4, the design of the occluders transformed the apparatus into a closed container. Evidence from infants shows that reasoning about object containment emerges later than other types of physical knowledge and containment may make inferences about object interactions more challenging to represent (Hespos & Baillargeon, 2001; 2006). Training data from the current study indicates that there may be some additional challenge presented by the full occluder design, as dogs averaged higher numbers of training trials to reach test criteria in both Studies 3 and 4. We also only observed location biases (all to the bottom) in Studies 3 & 4 training phases. With scant evidence that dogs have a location preference in this task, the behaviour could indicate lack of understanding about the enclosed apparatus, potentially along the same lines displayed by infants when shown containment interactions, or simply that the larger occluders prevented adequate visual cues. Additional internal data from warm up treat-hiding games suggest that object containment may be similarly challenging for dogs as they had comparatively lower success findings treats placed inside cups than when they were occluded in other ways (i.e., underneath or behind).

Conclusions

Knowledge about the physical properties of objects is fundamental for navigating the environment. Humans demonstrate perceptual awareness of object properties from early infancy, visually detecting violations of features such as solidity (Baillargeon & Graber, 1987; Spelke et al., 1992), and later in development demonstrating a more advanced ability to make inferences about physical objects in a manner that supports goal directed actions (Hood et al., 2003). Similar perceptual and inferential abilities have also been demonstrated in some species of non-human primates which has led to the idea that the ability may be evolutionarily ancient. This has proved challenging to back up with empirical evidence from non-primate species, but that might actually be an artifact of methodological limitations or task differences that prevent direct cross-species comparisons rather than an absence of physical knowledge. By investigating dogs' reasoning about solidity and continuity, our study fills three important gaps in scientists' understanding of non-primate physical reasoning. First, by using a validated procedure from developmental psychology we can more directly compare cross-species performance. Second, our study helps to address the limited existing research on dog causal and physical reasoning. Though evidence is sparse, dogs have gained a reputation for performing poorly on physical cognition tasks. While our study only addresses one aspect of physical reasoning, the results we present here suggest that dogs may have some ability to make inferences using physical properties of objects when choosing where to search for a reward. The results of trial 1 performance pooled across Studies 1-3 suggest dogs may be able to make inferences that support appropriate action by using salient visual information about solidity and continuity, though future work is still required to tease apart the roles of physical reasoning and associative learning. Overall, these results warrant more investigation before strong conclusions can be made about dogs’ abilities to make inferences about physical objects, particularly regarding dogs’ ability to factor changes in the environment into their representation of the world, and the extent to which associative learning plays a role in successful search.

References

Albiach-Serrano, A., Bräuer, J., Cacchione, T., Zickert, N., & Amici, F. (2012). The effect of domestication and ontogeny in swine cognition (*Sus scrofa scrofa and S. s. domestica*). *Applied Animal Behaviour Science*, *141*(1-2), 25-35.

Andrews, K. (2020). *How to Study Animal Minds* (Elements in the Philosophy of Biology). Cambridge: Cambridge University Press.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. *Blackwell handbook of childhood cognitive development*, *1*, 46-83.

Baillargeon, R., & Graber, M. (1987). Where's the rabbit? 5.5-month-old infants' representation of the height of a hidden object. *Cognitive Development*, *2*(4), 375-392.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bensky, M. K., Gosling, S. D., & Sinn, D. L. (2013). The world from a dog's point of view: a review and synthesis of dog cognition research. *Advances in the Study of Behavior*, *45*, 209-406. https://doi.org/10.1016/B978-0-12-407186-5.00005-7

Bräuer, J., & Belger, J. (2018). A ball is not a Kong: Odor representation and search behavior in domestic dogs (*Canis familiaris*) of different education. *Journal of Comparative Psychology, 132*(2), 189–199. https://doi.org/10.1037/com0000115

Bräuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making inferences about the location of hidden food: social dog, causal ape. *Journal of Comparative Psychology*, *120*(1), 38-47.

Byosiere, S.-E., Feng, L. C., Woodhead, J. K., Rutter, N. J., Chouinard, P. A., Howell, T. J., & Bennett, P. C. (2017). Visual perception in domestic dogs: Susceptibility to the Ebbinghaus–Titchener and Delboeuf illusions. *Animal Cognition*, *20*(3), 435–448. https://doi.org/10.1007/s10071-016-1067-1

Cacchione, T., & Burkart, J. M. (2012). Dissociation between seeing and acting: Insights from common marmosets (*Callithrix jacchus*). *Behavioural processes*, *89*(1), 52-60. https://doi.org/10.1016/j.beproc.2011.10.010

Cacchione, T., Call, J., & Zingg, R. (2009). Gravity and solidity in four great ape species (*Gorilla gorilla*, *Pongo pygmaeus*, *Pan troglodytes*, *Pan paniscus*): Vertical and horizontal variations of the table task. *Journal of Comparative Psychology*, *123*(2), 168-180. http://doi.org/10.1037/a0013580

Cacchione, T., & Rakoczy, H. (2017). Comparative metaphysics: Thinking about objects in space and time. 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. 579–599). American Psychological Association. https://doi.org/10.1037/0000012-026

Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of science*, *63*(4), 515-533.

Christie, L. A., Studzinski, C. M., Araujo, J. A., Leung, C. S., Ikeda-Douglas, C. J., Head, E., ... & Milgram, N. W. (2005). A comparison of egocentric and allocentric age-dependent spatial learning in the beagle dog. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(3), 361-369.

Espinosa, J., Dong, L., & Buchsbaum, D. (2021, May 30). Domestic dogs' gaze and behavior in 2-alternative choice tasks.*PsyArXiv*. https://doi.org/10.31234/osf.io/pt3vh

Fiset, S., Gagnon, S., & Beaulieu, C. (2000). Spatial encoding of hidden objects in dogs (*Canis familiaris*). *Journal of Comparative Psychology, 114*(4), 315-324.

Gadbois, S., & Reeve, C. (2014). Canine olfaction: scent, sign, and situation. In *Domestic dog cognition and behavior* (pp. 3-29). Springer, Berlin, Heidelberg.

Gopnik, A., & Meltzoff, A. (1987). The Development of Categorization in the Second Year and Its Relation to Other Cognitive and Linguistic Developments. *Child Development, 58*(6), 1523-1531. https://doi.org/10.2307/1130692

Haemmerli, S., Thill, C., Amici, F., & Cacchione, T. (2018). Domestic horses (*Equus ferus caballus*) fail to intuitively reason about object properties like solidity and weight. *Animal cognition, 21*(3), 441-446.

Hood, B. M. (1995). Gravity rules for 2-to 4-year olds?. *Cognitive Development*, *10*(4), 577-598.

Hood, B., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events: Two-year-olds fail to reveal knowledge of solidity and support. *Child Development*, *71*(6), 1540-1554.

Hood, B., Cole-Davies, V., & Dias, M. (2003). Looking and search measures of object knowledge in preschool children. *Developmental Psychology, 39*(1), 61–70. https://doi.org/10.1037/0012-1649.39.1.61

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. R. (1999). Gravity Biases in a Non-Human Primate? *Developmental Science*, *2*, 35–41. https://doi.org/10.1111/1467-7687.00051

Keen, R. (2003). Representation of objects and events: Why do infants look so smart and toddlers look so dumb?. *Current Directions in Psychological Science*, *12*(3), 79-83.

Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in Cognitive Sciences, 21*(10), 749-759. https://doi.org/10.1016/j.tics.2017.06.002

Kundey, S. M., De Los Reyes, A., Taglang, C., Baruch, A., & German, R. (2010). Domesticated dogs'(*Canis familiaris*) use of the solidity principle. *Animal Cognition*, *13*(3), 497-505.

Lampe, M., Bräuer, J., Kaminski, J., & Virányi, Z. (2017). The effects of domestication and ontogeny on cognition in dogs and wolves. *Scientific Reports, 7*(1), 1-8. https://doi.org/10.1038/s41598-017-12055-6

Mongillo, P., Araujo, J. A., Pitteri, E., Carnier, P., Adamelli, S., Regolin, L., & Marinelli, L. (2013). Spatial reversal learning is impaired by age in pet dogs. *Age*, *35*(6), 2273-2282.

Müller, C. A., Riemer, S., Range, F., & Huber, L. (2014). Dogs' use of the solidity principle: revisited. *Animal Cognition*, *17*(3), 821-825.

Osthaus, B., Lea, S. E., & Slater, A. M. (2005). Dogs (*Canis lupus familiaris*) fail to show understanding of means-end connections in a string-pulling task. *Animal Cognition*, *8*(1), 37-47. https://doi.org/10.1007/s10071-004-0230-2

Osthaus, B., Slater, A. M., & Lea, S. E. G. (2003). Can dogs defy gravity? A comparison with the human infant and a non-human primate. *Developmental Science*, *6*(5), 489–497. https://doi.org/10.1111/1467-7687.00306

Pattison, K. F., Miller, H. C., Rayburn-Reeves, R., & Zentall, T. (2010). The case of the disappearing bone: Dogs' understanding of the physical properties of objects. *Behavioural Processes*, *85*(3), 278-282.

Piotti, P., Szabó, D., Bognár, Z., Egerer, A., Hulsbosch, P., Carson, R. S., & Kubinyi, E. (2018). Effect of age on discrimination learning, reversal learning, and cognitive bias in family dogs. *Learning & behavior*, *46*(4), 537-553.

Plotnik, J. M., Brubaker, D. L., Dale, R., Tiller, L. N., Mumby, H. S., & Clayton, N. S. (2019). Elephants have a nose for quantity. *Proceedings of the National Academy of Sciences*, *116*(25), 12566-12571.

Povinelli, D. J., & Henley, T. (2020). More rope tricks reveal why more task variants will never lead to strong inferences about higher-order causal reasoning in chimpanzees. *Animal Behavior and Cognition*, *7*(3), 392-418. https://doi.org/10.26451/abc.07.03.08.2020

Range, F., Hentrup, M., & Virányi, Z. (2011). Dogs are able to solve a means-end task. *Animal Cognition*, *14*(4), 575-583. https://doi.org/10.1007/s10071-011-0394-5

Range, F., Möslinger, H., & Virányi, Z. (2012). Domestication has not affected the understanding of means-end connections in dogs. *Animal cognition*, *15*(4), 597-607. https://doi.org/10.1007/s10071-012-0488-8

Riemer, S., Müller, C., Range, F., & Huber, L. (2014). Dogs (*Canis familiaris*) can learn to attend to connectivity in string pulling tasks. *Journal of Comparative Psychology*, *128*(1), 31–39. https://doi.org/10.1037/a0033202

Rosner, B. (2015). *Fundamentals of biostatistics*. Nelson Education.

Santos, L. R., & Hauser, M. D. (2002). A non-human primate's understanding of solidity: dissociations between seeing and acting. *Developmental Science*, *5*(2), 1–7. https://doi.org/10.1111/1467-7687.t01-1-00216

Santos, L. R. (2004). "Core Knowledges": A dissociation between spatiotemporal knowledge and contact-mechanics in a non-human primate? *Developmental Science*, *7*(2), 167–174. https://doi.org/10.1111/j.1467-7687.2004.00335.x

Santos, L. R., Seelig, D., & Hauser, M. D. (2006). Cotton-top tamarins' (*Saguinus oedipus*) expectations about occluded objects: A dissociation between looking and reaching tasks. *Infancy*, *9*(2), 141–165. https://doi.org/10.1207/s15327078in0902_4

Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(1), 23–34. https://doi.org/10.1037/a0012925

Southgate, V., & Gomez, J. C. (2006). Searching beneath the shelf in macaque monkeys: Evidence for a gravity bias or a foraging bias? *Journal of Comparative Psychology*, *120*(3), 314-321.

Spelke, E. S. (1988). *The origins of physical knowledge.* In L. Weiskrantz (Ed.), *A Fyssen Foundation symposium. Thought without language* (p. 168–184). Clarendon Press/Oxford University Press.

Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*(1-3), 431-445. https://doi.org/10.1016/0010-0277(94)90039-6

Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. *Language in mind: Advances in the study of language and thought*, 277-311.

Spelke, E. S. (2000). Core knowledge. *American psychologist*, *55*(11), 1233-1243.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological review*, *99*(4), 605-632.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89-96.

Spelke, E. S., Lee, S. A., Izard, V. (2010). Beyond Core Knowledge: Natural Geometry, *Cognitive Science*, *34*(5), 863-884. https://doi.org/10.1111/j.1551-6709.2010.01110.x

Tecwyn, E.C. & Buchsbaum, D. (2018) Hood's Gravity Rules. In: Vonk, J. & Shackleford, T. (Eds) *Encyclopedia of Animal Cognition and Behavior*. Springer. https://doi.org/10.1007/978-3-319-47829-6_1535-1

Tecwyn, E.C., Buchsbaum, D. (2019) What factors really influence domestic dogs' (*Canis familiaris*) search for an item dropped down a diagonal tube? The tubes task revisited. *Journal of Comparative Psychology*, *133*(1), 4-19. http://dx.doi.org/10.1037/com0000145

Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual completion in newborn human infants. *Child Development*, *77*(6), 1810-1821.

Völter, C., & Huber, L. (2021, April 27). Expectancy Violations about Physical Properties of Animated Objects in Dogs. *PsyArXiv*. https://doi.org/10.31234/osf.io/3pr9z

Völter, C. J., Karl, S., & Huber, L. (2020). Dogs accurately track a moving object on a screen and anticipate its destination. *Scientific reports*, *10*(1), 1-10. https://doi.org/10.1038/s41598-020-72506-5

Wellman, H.M., & Gelman, S.A. (1992). Cognitive Development: Foundational Theories of Core Domains, *Annual Review of Psychology*, *43*(1), 337-375.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103-128. https://doi.org/10.1016/0010-0277(83)90004-5