

Cognitive characteristics of 8- to 10-week-old assistance dog puppies

Emily E. Bray^{a, b, *}, Margaret E. Gruen^c, Gitanjali E. Gnanadesikan^{a, d},
Daniel J. Horschler^{a, d}, Kerinne M. Levy^b, Brenda S. Kennedy^b, Brian A. Hare^{e, f},
Evan L. MacLean^{a, d, g, h}

^a Arizona Canine Cognition Center, School of Anthropology, University of Arizona, Tucson, AZ, U.S.A.

^b Canine Companions for Independence, National Headquarters, Santa Rosa, CA, U.S.A.

^c Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, U.S.A.

^d Cognitive Science Program, University of Arizona, Tucson, AZ, U.S.A.

^e Duke Canine Cognition Center, Department of Evolutionary Anthropology, Duke University, Durham, NC, U.S.A.

^f Center for Cognitive Neuroscience, Duke University, Durham, NC, U.S.A.

^g Department of Psychology, University of Arizona, Tucson, AZ, U.S.A.

^h College of Veterinary Medicine, University of Arizona, Tucson, AZ, U.S.A.

ARTICLE INFO

Article history:

Received 14 November 2019

Initial acceptance 19 February 2020

Final acceptance 15 April 2020

MS. number: A19-00758R

Keywords:

assistance dog
behaviour
canine
cognition
development
temperament

To characterize the early ontogeny of dog cognition, we tested 168 domestic dog, *Canis familiaris*, puppies (97 females, 71 males; mean age = 9.2 weeks) in a novel test battery based on previous tasks developed and employed with adolescent and adult dogs. Our sample consisted of Labrador retrievers, golden retrievers and Labrador × golden retriever crosses from 65 different litters at Canine Companions for Independence, an organization that breeds, trains and places assistance dogs for people with disabilities. Puppies participated in a 3-day cognitive battery that consisted of 14 tasks measuring different cognitive abilities and temperament traits such as executive function (e.g. inhibitory control, reversal learning, working memory), use of social cues, sensory discriminations and reactivity to and recovery from novel situations. At 8–10 weeks of age, and despite minimal experience with humans, puppies reliably used a variety of cooperative-communicative gestures from humans. Puppies accurately remembered the location of hidden food for delays of up to 20 s, and succeeded in a variety of visual, olfactory and auditory discrimination problems. They also showed some skill at executive function tasks requiring inhibitory control and reversal learning, although they scored lower on these tasks than is typical in adulthood. Taken together, our results confirm the early emergence of sensitivity to human communication in dogs and contextualize these skills within a broad array of other cognitive abilities measured at the same stage of ontogeny.

© 2020 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

A central idea in comparative cognition is that individual and species differences in cognitive phenotypes reflect variation across multiple domains of cognition (Cheney & Seyfarth, 1990; Shettleworth, 2013; Tomasello & Call, 1997). Put simply, it is unlikely that a single construct such as general intelligence will be sufficient to account for variation in cognitive phenotypes, either within or across species (Hare & Wrangham, 2002; MacLean et al., 2012; MacLean & Nunn, 2017). In an effort to develop measures that can broadly characterize variance across multiple cognitive domains, researchers have developed large-scale cognitive test

batteries, allowing them to broadly survey and describe patterns of variation both within and across species (e.g. Herrmann, Call, Hernández-Lloreda, Hare, & Tomasello, 2007; MacLean, Herrmann, Suchindran, & Hare, 2017; Schmitt, Pankau, & Fischer, 2012; Shaw & Schmelz, 2017).

Recent work with domestic dogs, *Canis familiaris*, has supported the hypothesis that individual differences reflect variance across multiple cognitive domains. Both citizen science and conventional laboratory approaches have produced evidence of multiple cognitive factors by testing hundreds of adult dogs across a diverse range of cognitive tasks (Gnanadesikan, Hare, Snyder-Mackler, & MacLean, 2020; MacLean et al., 2017; Stewart et al., 2015; see also Horschler et al., 2019). MacLean et al. (2017) recently tested more than 550 adult dogs using over two dozen cognitive tasks designed to assay a range of social and nonsocial cognitive abilities. Analysis

* Correspondence: E. E. Bray, Arizona Canine Cognition Center, University of Arizona, School of Anthropology, 1009 East South Campus Drive, Tucson, AZ, 85721-0030, U.S.A.

E-mail address: ebray@email.arizona.edu (E. E. Bray).

of this data set revealed variance in several underlying psychological constructs such as memory, inhibitory control and cooperative communication. A subset of the tasks in MacLean et al.'s (2017) battery was derived from earlier work with human infants and great apes (Herrmann, Hare, Call, & Tomasello, 2010; Wobber, Herrmann, Hare, Wrangham, & Tomasello, 2014), allowing quantitative comparisons of individual differences across species. In these analyses, dogs and human children – but not chimpanzees, *Pan troglodytes* – performed similarly on tasks measuring sensitivity to cooperative communication, and humans and dogs (but again not chimpanzees) demonstrated correlated variance in skills for cooperative communication (MacLean et al., 2017).

An important question raised by the work above concerns the developmental origins of dog cognition. Although several previous studies have characterized puppy performance on a few specific cognitive tasks, we still know little about the early ontogeny of most cognitive processes. Consequently, many questions about the developmental foundations of adult dog cognition remain poorly understood. For example, it has been proposed that dogs' sensitivity to cooperative-communicative gestures may be biologically prepared, emerging early in development prior to extensive socialization with humans (Hare, Brown, Williamson, & Tomasello, 2002). Although these predictions have been supported in some studies with young puppies, typical sample sizes – particularly for dogs younger than 10 weeks – have been modest, and questions about the early ontogeny of dog social cognition have remained controversial (Dorey, Udell, & Wynne, 2010; Hare et al., 2002, 2010; Riedel, Schumann, Kaminski, Call, & Tomasello, 2008; Wynne, Udell, & Lord, 2008). In addition to questions about the emergence of gesture following, we still know very little about the development of other cognitive processes in dogs, including additional aspects of social cognition, executive functions (such as memory and inhibitory control) and sensory processes (e.g. vision, audition, olfaction), despite substantial research on these topics in adulthood (Bensky, Gosling, & Sinn, 2013; Hare & Woods, 2013; Miklósi, 2014). Thus, research on the development of these processes will be critical to understanding the ontogenetic foundations of adult cognitive phenotypes (Tinbergen, 1963).

In addition to these basic research questions, knowledge about dog cognitive development may also inform the processes through which dogs are trained or evaluated for working roles (Lazarowski, Waggoner, & Katz, 2019; MacNamara & MacLean, 2017). For example, recent work suggests that individual differences in neural, cognitive and behavioural processes predict aptitude for a variety of working dog roles, including placement as an explosive detection, assistance or guide dog (Berns, Brooks, Spivak, & Levy, 2017; Bray et al., 2019; Bray, Sammel, Cheney, Serpell, & Seyfarth, 2017; MacLean & Hare, 2018). Therefore, understanding the development of these processes may inform when these traits can first be meaningfully measured, as well as how they may constrain or facilitate training at various points in ontogeny.

As a first step towards broadly characterizing the early ontogeny of dog cognition, we developed the Dog Cognitive Development Battery (DCDB), a series of behavioural tests that can be implemented with both young puppies and adult dogs (Fig. 1). This battery was based on previous test batteries that have been employed with adult dogs (Bray, Sammel, Seyfarth, Serpell, & Cheney, 2017; MacLean et al., 2017), with modifications to facilitate testing of dogs as young as 8 weeks. Broadly, the measures were intended to assess variation in aspects of social motivation, communication, executive function, sensory perception and temperament (Fig. 1). Our social motivation and communication tasks measured puppies' interest and engagement with humans, as well as their ability to utilize human-given cues to find hidden food. We measured processes related to executive function with tasks

requiring impulse control, reversal learning and working memory (Tapp, Head et al., 2003; Tapp, Siwak, Estrada, Holowachuk, & Milgram, 2003). Our sensory perception tasks consisted of object choice tasks in which the two response options looked, sounded or smelled different. Our temperament tasks assessed behavioural lateralization (believed to reflect lateralization in the brain) via a first-step laterality task (e.g. Batt, Batt, Baguley, & McGreevy, 2009) as well as temperamental reactivity to startling stimuli and unfamiliar situations via the novel object and surprising events tasks, based on previous studies in young puppies and adolescent dogs (e.g. Asher et al., 2013; Bray, Seyfarth et al., 2017; McGarrity, Sinn, & Gosling, 2015; Riemer, Müller, Virányi, Huber, & Range, 2014). In total, the DCDB included 14 different tasks that could be completed across three ~45 min sessions per puppy.

Based on previously published work with smaller samples as well as our pilot data (see Supplementary Material), we hypothesized that by 2 months of age, puppies would (1) perform at above-chance expectation on sensory perception tasks involving vision, olfaction and audition, (2) successfully use executive functions in contexts requiring memory, inhibitory control (i.e. the cylinder task inhibitory control trials) and a shift in response strategies (i.e. the cylinder task reversal learning trials), (3) display a motivation to interact with and look towards humans, reflected in their performance on the human interest and unsolvable tasks, and (4) perform at above-chance levels on social cognitive tasks requiring use of cooperative-communicative gestures (i.e. arm pointing, communicative marker).

GENERAL METHODS

Subjects

Participating puppies were recruited through Canine Companions for Independence (CCI; Santa Rosa, CA, U.S.A.), a national nonprofit organization that breeds, raises and trains assistance dogs for people with disabilities. Puppies were either whelped in the homes of local volunteer breeder caretakers ($N = 147$) or whelped at the Canine Early Development Center ($N = 21$), a state-of-the-art facility with full-time staff dedicated to monitoring and caring for the mothers and their litters. Puppies were kept in a whelping pool or box – to which the mother had constant access – for the first 4 weeks, and then transitioned to living and sleeping together in a larger enclosed area. Puppies were with their mothers and littermates until they were weaned at 6 weeks of age, at which point puppy food was introduced and provided to the litter by the human caretaker three times per day. Puppies remained housed socially with their littermates until around 8 weeks of age. At this point, all puppies spent time at CCI's national headquarters where they underwent veterinary examinations prior to being sent to individual volunteer puppy raisers for the next 14–22 months. CCI granted informed consent to all aspects of the study.

We tested 168 puppies (97 females and 71 males) during February–July of 2017 (Supplementary Table S1) when they were approximately 2 months old (range 7.86–10.43 weeks; mean = 9.20 weeks), prior to being placed with their puppy raiser. Our sample consisted of 122 Labrador × golden crosses, 40 Labrador retrievers and 6 golden retrievers from 65 different litters.

Procedure

We implemented a test battery in puppies that was modelled after the dog cognition test battery (DCTB), previously developed for adult dogs (MacLean et al., 2017). To determine an appropriate testing age that would accommodate meaningful participation, as well as to adapt the tasks appropriately for puppies, preliminary

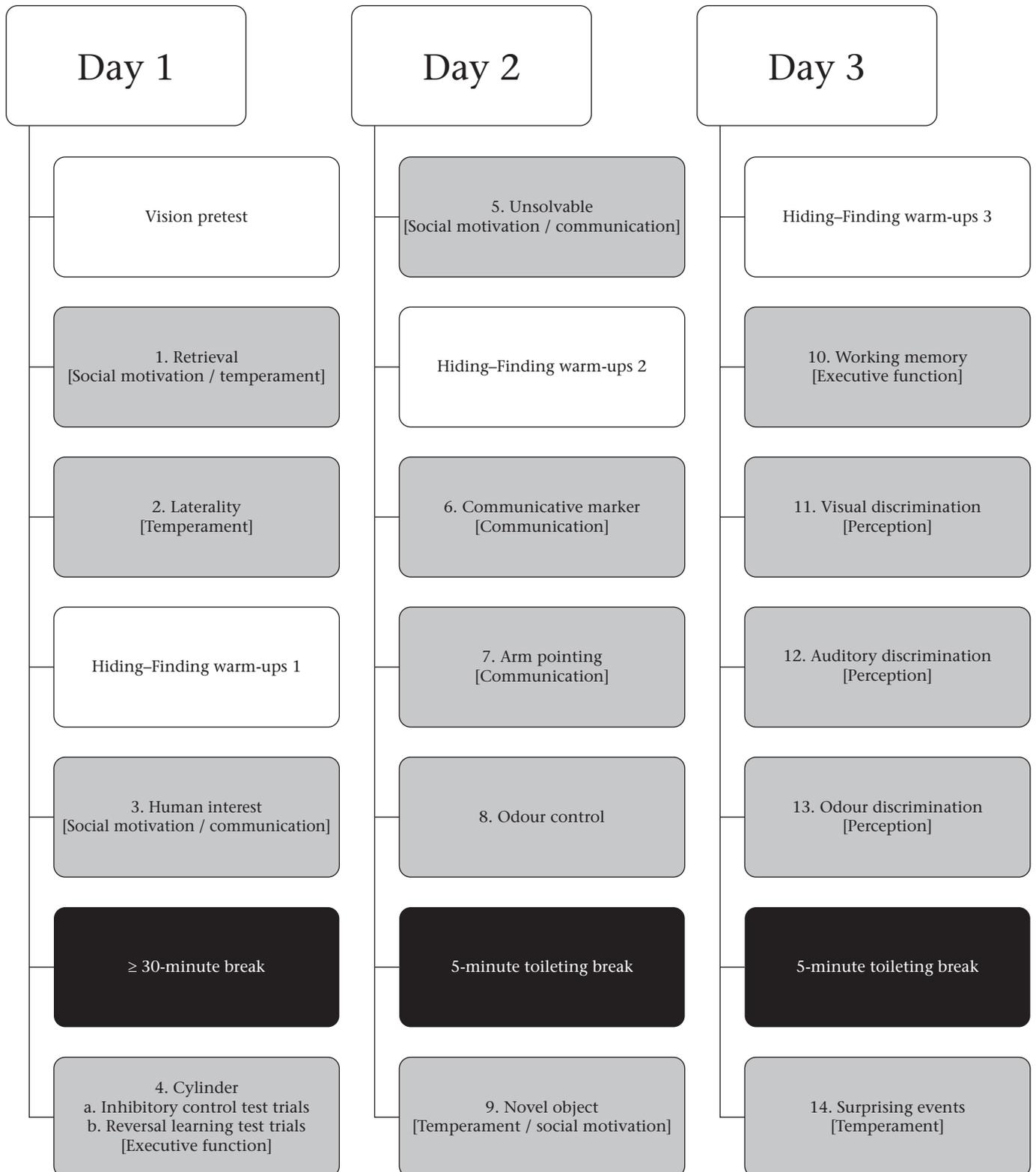


Figure 1. Dog Cognitive Development Battery (DCDB) tasks in the order in which they were administered, consisting of three sessions spread out over 3 days. Below each task, in brackets, we have indicated the primary purpose of each measure from a design perspective. However, we expected that performance on most measures would be influenced by both cognitive and temperamental factors and may reflect variation in processes beyond the target construct(s).

research methods were piloted with dogs between 7 and 12 weeks of age from two working dog populations (see Supplementary Material). Pilot studies suggested that the majority of tasks were not feasible with puppies younger than 8 weeks of age. Pilot

subjects between 8 and 12 weeks of age met basic perceptual (visual and auditory sensitivity) and cognitive (object permanence, motivation to search for hidden rewards) demands. Performance across the full range of tasks did not vary notably between 8 and 12

weeks in our pilot sample ($N = 20$), and thus we selected 9 weeks as our target age for further research.

Dog Cognitive Development Battery

Acknowledging that temperament and cognition are often interrelated (Bray, MacLean, & Hare, 2015; Bray, Seyfarth et al., 2017; Hare & Tomasello, 2005), the 14 tasks in the DCDB were explicitly designed to assess processes influenced by social motivation, communication, executive function, perception and temperament (see Fig. 1 for which tasks were specifically designed to assess each target process). All tasks in the DCDB are briefly described below; detailed experimental methods and video examples are provided in the Supplementary Material. The battery was implemented across three consecutive days with each subject, each session lasting ~45 min (Fig. 1).

To ensure consistency and allow for comparison between individuals, tasks were presented in the same order across subjects (Bray, Seyfarth et al., 2017; MacLean et al., 2017). On any given task, if a puppy failed to choose within a set amount of time or there was an experimenter error, that trial was repeated. If the subject continued to show a lack of motivation to participate, we attempted to re-engage and refamiliarize the puppy with the demands of the task, and if necessary gave the puppy a break before returning to the task (see Supplementary Material for specific refamiliarization and abort criteria for each task). In the rare instances where those attempts were unsuccessful, the task was discontinued for that puppy (Supplementary Table S2).

Vision pretest

This test ensured that puppies were capable of tracking visual stimuli at the typical distances used in subsequent tasks (based on Ollivier, Plummer, & Barrie, 2007). At a distance of 100 cm in front of the puppy, a cotton ball was dropped vertically and flicked across the ground in full view of the subject. Subjects were required to follow the motion of the cotton ball on at least three trials to advance to subsequent tasks. All puppies tested met this criterion.

Retrieval

This task measured the puppy's willingness to cooperatively engage in fetch with a human partner (based on Bray, Seyfarth et al., 2017; Slabbert & Odendaal, 1999; Wilsson & Sundgren, 1997). Following a 1 min familiarization period (see Supplementary Material), the experimenter threw a small ball for the puppy and vocally encouraged the dog to bring the ball back to her. For each of the two 1 min test trials, the puppy received a score based on the following scoring system: (1) did not interact with the ball at all, (2) only chased the ball, (3) also picked the ball up in the mouth, (4) returned the ball to the experimenter one to two times, or (5) returned the ball to the experimenter three or more times. The dependent measures were the puppy's average score across two trials and a tally of the total number of times that the puppy returned the ball to the experimenter.

Laterality

This task indexed behavioural measures of laterality by tracking the puppy's paw preference when stepping onto and off of a platform (based on Tomkins, Thomson, & McGreevy, 2010), which is believed to reflect lateralization in the brain and has been previously linked to temperamental reactivity in adult dogs (Branson & Rogers, 2006). Following a brief introduction to the platform (see Supplementary Material), puppies were held by the handler and then called by the experimenter to step onto the platform across a series of 15 trials, and then off the platform across a series of 15

trials. The forelimb used to initiate this motion on each trial was recorded and subsequently used to compute a laterality index.

Hiding–finding warm-ups

Warm-up trials ensured that puppies were motivated to search for the reward and capable of reliably choosing between two options in an object choice paradigm. After an initial familiarization to the apparatus and choice procedure (see Supplementary Material), two opaque containers were placed in front of the puppy. In this task and subsequent object choice tasks (i.e. gesture use and working memory), a piece of kibble was taped to the inside bottom of both containers as a control for odour cues. The experimenter showed the puppy a food reward and placed it underneath one of the containers. Puppies were required to choose correctly by physically touching the baited container with snout or front paw on four of five consecutive trials to advance to subsequent object choice tasks. Puppies completed this task once per session.

Human interest

This task measured the puppy's motivation to attend to a human who spoke to the puppy using dog-directed speech (Ben-Aderet, Gallego-Abenza, Reby, & Mathevon, 2017; Gergely, Faragó, Galambos, & Topál, 2017). The experimenter stood outside the testing pen, looked at the puppy and recited a predetermined script with a fluctuating, high-pitched intonation (Ben-Aderet et al., 2017). After each recitation, the experimenter entered the pen and petted the puppy if approached. This procedure was repeated three times. The duration of the puppy's gaze to the human's face during the recitation of the script and the duration of interaction with the experimenter during play breaks was recorded across trials.

Cylinder inhibitory control and cylinder reversal learning

The first part of this task measured the puppy's inhibitory control (i.e. the ability to suppress a prepotent response in favour of a choice that would ultimately be more productive) by requiring the puppy to detour to the reward location, thereby placing distance between herself and a visible reward (based on Bray, MacLean, & Hare, 2014; MacLean et al., 2014). This task is often employed in the canine literature as a measure of motor inhibition (Brucks, Marshall-Pescini, Wallis, Huber, & Range, 2017; Fagnani, Barrera, Carballo, & Bentosela, 2016; Marshall-Pescini, Virányi, & Range, 2015; but for critiques see Kabadyi, Bobrowicz, & Osvath, 2018; van Horik et al., 2018; van Horik, Beardsworth, Laker, Whiteside, & Madden, 2020). The second part of this task measured the puppy's ability to exhibit cognitive flexibility when the demands of the task changed, and the puppy's previously preferred solution was no longer available. Puppies first participated in familiarization trials by walking around the front of an opaque cylinder to retrieve a reward from one of the side openings. In inhibitory control test trials (4a), a transparent cylinder was used such that subjects had to resist the prepotent response to move directly towards the visible food, instead avoiding the transparent obstacle. Eight test trials were conducted. The dependent measures were the proportion of trials that the puppy successfully retrieved the food from either side opening of the cylinder, without first touching the exterior of the apparatus, and the average latency to obtain the reward. In reversal learning test trials (4b), the puppy's preferred side entrance to the cylinder was obstructed by a transparent plastic barrier and subjects were required to switch their response, detouring to the other opening of the apparatus to retrieve the treat. Eight test trials were conducted. The dependent measure was the proportion of trials that puppies performed the correct detour response without first touching the barrier or exterior of the cylinder. The side of the apparatus that the subject first approached (i.e. open or blocked) and the average latency to

obtain the reward were also recorded as measures of response flexibility.

Unsolvable

This task measured the puppy's inclination to persist at an unsolvable task independently versus looking at a nearby human experimenter to potentially solicit help (based on Miklósi et al., 2003; for alternative explanations of what this task measures see Lazzaroni et al., 2020). The puppy was familiarized with displacing the lid from a transparent container to obtain a visible food reward inside. Then, across four 30 s test trials, the lid to the container was affixed, and the dependent measures were the duration of time gazing at the experimenter's face and duration of time physically manipulating the container.

Gesture use

The experimenter showed the puppy a food reward, then used a foamboard occluder to block the puppy's view while placing the reward inside one of two possible hiding locations. The experimenter then removed the occluder, provided one of three cues (communicative marker, arm pointing, odour control; see below) before subjects could search and recorded the subject's first choice.

Communicative marker. This task measured the puppy's ability to use an arbitrary marker, placed in a communicative manner, to find a hidden reward (based on Agnetta, Hare, & Tomasello, 2000; Riedel, Buttelmann, Call, & Tomasello, 2006). The experimenter ostensibly (preceded by verbally addressing and making eye contact with the puppy) placed a small yellow block that the puppy had never seen before next to the baited location. Twelve test trials were conducted.

Arm pointing. This task measured the puppy's ability to use an arm-pointing gesture to find a hidden reward (based on Hare, Call, & Tomasello, 1998; Miklósi, Polgárdi, Topál, & Csányi, 1998). The experimenter ostensibly (preceded by verbally addressing and making eye contact with the puppy) pointed with the contralateral arm, index finger extended, and gazed towards the baited location until the trial ended. Twelve test trials were conducted.

Odour control. This task acted as a control to ensure that puppies' performance on the gesture use tasks could not be attributed to olfactory cues or unintentional cuing by the experimenter (based on Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Hare et al., 2002; Miklósi et al., 1998). After baiting, the experimenter remained still and did not provide any social information. Eight test trials were conducted.

The dependent measures for the gesture use tasks were the proportion of trials that the puppy's first choice was to the baited location, where a choice was defined as the puppy physically touching the cup with the snout or a front paw (see Supplementary Material).

Novel object

This task measured the puppy's reaction to an unfamiliar object. It was not included in the original DCTB, so methods were adapted from previously published studies (Bray, Cheney et al., 2017; Goddard & Beilharz, 1984; King, Hemsworth, & Coleman, 2003; Marshall-Pescini, Virányi, Kubinyi, & Range, 2017). A stuffed mechanical, motion-activated cat (FurReal Friends Daisy Plays-With-Me Kitty Toy, Hasbro, Inc., Pawtucket, RI, U.S.A.) was placed inside the pen with the puppy. The experimenter and handler exited the room, leaving the puppy alone with the locomoting and vocalizing cat. After 2 min, the experimenter re-entered the room and encouraged the puppy to approach the cat. Finally, the handler re-

entered the room and placed the puppy approximately 1 m in front of the cat, at which point the experimenter once again encouraged the puppy to approach. The subject's behaviour was scored from video. The ethogram used for behavioural coding in this task is described in [Supplementary Table S3](#).

Working memory

This task measured the puppy's ability to recall the location of a hidden treat after temporal delays of various lengths (based on Doré, Fiset, Goulet, Dumas, & Gagnon, 1996; Fiset, Beaulieu, & Landry, 2003). It was identical to hiding–finding warm-ups with the exception that we imposed a delay before the subject was allowed to search, which increased across blocks of six trials each (5 s, 10 s, 15 s, 20 s). Only individuals who chose correctly on at least four of six trials at 10 s moved on to delays of 15 s, and only those who chose correctly on at least four of six trials at 15 s moved on to delays of 20 s. The proportion of trials that the subject first searched in the baited location was used as the dependent measure.

Perceptual discriminations

The subject had to choose between two search locations based on a perceptual cue (visual, auditory, olfactory; see below) regarding which location contained the reward.

Visual discrimination. This task measured the puppy's ability to choose a baited location versus an unbaited location based on visual cues. One plate contained five pieces of visible kibble and the other was empty. The experimenter presented the plates directly in front of the puppy before pulling them backward to 50 cm in front of the puppy, equidistant to the left and right sides. Eight test trials were conducted. The proportion of trials that the puppy first approached the baited plate (i.e. the puppy's snout extended over the plate) was used as the dependent measure.

Auditory discrimination. This task measured the puppy's ability to choose a baited location versus an unbaited location based on auditory cues (based on Bräuer et al., 2006). Two metal bowls, placed approximately 50 cm away from the puppy, were used as the hiding locations. The experimenter sequentially placed her hand into each container, audibly dropping the food into only one of the containers. Eight test trials were conducted. The dependent measure was the proportion of trials that the subject's first search was to the baited location.

Odour discrimination. This task measured the puppy's ability to choose a baited location versus an unbaited location based on olfactory cues. Two sections of rubber tubing with a 90° bend ('elbows') were presented, one of which contained 10 pieces of dry kibble. The ends of the elbows were filled with cotton to prevent the contents from being visible or audible. The experimenter allowed the subject to sniff the opening of each elbow individually for 3 s, and then the elbows were presented side by side for an additional 3 s before being pulled backward 50 cm in front of the puppy, equidistant to the left and right sides. Puppies were released and allowed to move freely for 20 s. On each trial, the first and last elbow that the subject approached was recorded as well as the cumulative time spent within a marked 10 cm radius around the elbows. Eight test trials were conducted. The dependent measures were the proportion of trials that the subject's first and last responses were directed to the baited location as well as the proportion of time that the puppy spent within each of the marked radii around the elbows.

Surprising events

This task measured the puppy's reaction to a series of unexpected and potentially startling stimuli. It was not included in the original DCTB, so methods were adapted from previously published studies (Bray, Cheney et al., 2017; Sherman et al., 2015; van der Borg, Netto, & Planta, 1991). The puppy was presented with a sequence of three potentially startling stimuli (sudden appearance, looming object, loud noise; see below) and the puppy's behavioural reactivity as well as subsequent recovery was scored from video using the ethogram described in [Supplementary Table S4](#).

Sudden appearance. From behind the puppy, the experimenter threw a large trash bag stuffed with shredded paper to the centre of the enclosure.

Looming object. The experimenter held a closed umbrella facing towards the subject. The experimenter then opened the umbrella such that it rapidly expanded towards the subject.

Loud noise. The experimenter shook a piece of sheet metal in front of the subject, which created a loud oscillating sound with corresponding pulses of air.

After each surprising event, the experimenter set the stimulus on the ground and the puppy was free to explore alone for 15 s. The experimenter then returned and vocally encouraged the subject to approach and eat kibble near the previously startling stimulus. After the third event (loud noise), the handler and experimenter left the immediate testing area and stood in the corner, leaving the puppy alone in the testing area for a brief 1 min period.

Ethical Note

All testing procedures were reviewed and adhered to regulations set forth by the Institutional Animal Care and Use Committees of the University of Arizona (IACUC No. 16–175) and Duke University (IACUC No. A182-17-08). Behavioural testing was designed to be as nondisruptive as possible, and we piloted all tasks to ensure that age-appropriate methods were used. Puppies ate their regular meals over the course of the testing session, ensuring that they were fed their usual amount. We incorporated plenty of play and bathroom breaks to ensure the testing experience was a positive one, and every task involved food rewards, play and/or praise. We also adhered to strict abort criteria that allowed puppies to opt out of a task if they were no longer engaged.

Coding and Reliability

Most behavioural variables were scored live, but all tasks were videorecorded for reliability assessment and additional analyses. The following tasks were later coded from video: novel object and surprising events, as well as select variables from human interest (duration of interaction with the experimenter during the play break), cylinder (latency during inhibitory control and reversal learning trials and first side correct during reversal learning trials), unsolvable (average time manipulating object) and odour discrimination (time at right and left elbow, from which the variables persistence, time at correct response and time at incorrect response were subsequently calculated).

For the live-coded data, independent coders scored from video all trials for 20% of randomly selected subjects, and interrater reliability was calculated using Pearson correlation for continuous variables and Cohen's κ for categorical variables. All statistical analyses were carried out in R v.3.6.0 (R Development Core Team, 2016). For the measures that were not possible to score live, two coders independently scored data from video. The primary coder

scored all data for analysis, and a reliability coder scored all trials for 20% of randomly selected subjects for reliability analyses.

Reliability was excellent for all live-coded measures and inter-rater agreement is reported in [Supplementary Table S5](#) (Cohen's κ : mean = 0.94; Pearson's r : mean = 0.96). Reliability was also excellent for the video-coded measures; inter-rater agreement for the video-coded measures of the cognitive tasks is reported in [Supplementary Table S6](#) (Cohen's κ = 0.93; Pearson's r : mean = 0.97), inter-rater agreement for the novel object task is reported in [Supplementary Table S7](#) (Cohen's κ : mean = 0.90; Pearson's r : mean = 0.96) and inter-rater agreement for the surprising events task, including sudden appearance, looming object and loud noise, is reported in [Supplementary Table S8](#) (Cohen's κ : mean = 0.82; Pearson's r : mean = 0.87).

RESULTS AND DISCUSSION

The sample size, units of measure, mean, standard deviation, range and standard error for all cognitive measures in the DCDB collected from our sample of puppies are reported in [Table 1](#).

To determine whether there was a significant effect of sex or age at testing, we used linear mixed-effects models fitted by restricted maximum likelihood (REML) with age at testing and sex as predictors of each measure, including litter as a random effect ([Supplementary Table S9](#)). Age and sex effects are only reported in the text when these covariates were statistically significant. For tasks in which chance expectation could be defined, we conducted one-sample t tests to determine whether performance deviated from the null expectation ([Table 2](#)). To determine whether there was an effect of learning in our gesture use tasks, we conducted logistic regressions with trial number as a predictor of puppy performance, as well as a binomial test on first trial performance across subjects. We also conducted a one-way repeated measures ANOVA and accompanying post hoc analyses using a Bonferroni adjustment to compare within-subject performance measures (% trials correct) for each type of our gesture use task (i.e. communicative marker, arm pointing, odour control).

We also report the sample size, units of measure, mean, standard deviation, range and standard error for the two temperament tasks (surprising events and novel object) in the DCDB ([Supplementary Tables S10–S11](#)). Due to the large number of measures in each temperament task, we used principal components analysis (PCA) to describe patterns of variation on these measures. For these analyses we used an oblique rotation, direct oblimin (Osborne & Costello, 2005), which allows the components to be correlated. We determined the number of components to retain by using parallel analysis (Horn, 1965), fitted using the 'fa.parallel' function from the R package 'psych' (Revelle, 2019). As above, all statistical analyses were carried out in R v.3.6.0 (R Development Core Team, 2016).

Pretest and Warm-ups

The near-perfect performance of puppies on the vision pretest indicated that their eyesight was sufficiently developed to view and respond to events occurring at least 1 m in front of them, as required in the primary DCDB tasks. Puppies were also highly successful in completing the hiding–finding warm-ups, a prerequisite for several of the subsequent object choice tasks (i.e. communicative marker, arm pointing, odour control, working memory), with over half of the puppies completing these warm-ups in the minimum number of trials by their second testing session. These results corroborate the findings of Gagnon and Doré (1994), who reported that, by 8 weeks of age, puppies were

Table 1
Descriptive statistics for Dog Cognitive Development Battery (DCDB) measures with 8- to 10-week-old puppies

Variable	N	Units	Mean	SD	Minimum	Maximum	SEM
Vision pretest	168	% Trials correct	96.83	7.94	50	100	0.61
Hiding–finding warm-ups: day 1	158	No. of trials to criterion	6.83	3.87	4	20	0.31
Hiding–finding warm-ups: day 2	164	No. of trials to criterion	6.22	3.69	4	22	0.29
Hiding–finding warm-ups: day 3	165	No. of trials to criterion	5.97	3.26	4	20	0.25
Retrieval: average score	168	Rating system (see text)	3.31	1.17	1	5	0.09
Retrieval: tally	168	No. of tallies	3.07	3.75	0	14	0.29
Laterality: bias strength	168	Bias strength (see text)	40.83	26.69	0	93.33	2.06
Laterality: laterality index	168	Laterality index (see text)	-7.18	48.35	-93.33	93.33	3.73
Human interest: avg. look time	150	No. of seconds	6.26	4.01	0	20.91	0.33
Human interest: avg. interact time	150	No. of seconds	18.48	8.01	0	29.70	0.65
Cylinder: familiarization score	166	No. of trials to criterion	7.78	3.16	4	18	0.24
Cylinder: inhibitory control score	166	% Trials correct	51.2	24.24	0	100	1.88
Cylinder: latency (inhibitory control trials)	166	No. of seconds	3.93	2.41	1.44	17.02	0.19
Cylinder: reversal learning score	165	% Trials correct	29.47	23.42	0	87.5	1.82
Cylinder: latency (reversal learning trials)	165	No. of seconds	6.58	4.51	2.22	30	0.35
Cylinder: first side correct (reversal learning trials)	165	% Trials correct	23.26	27.29	0	100	2.12
Unsolvable task: avg. time looking at human	168	No. of seconds	1	1.03	0	4.42	0.08
Unsolvable task: avg. time manipulating object	168	No. of seconds	12.71	3.27	3.25	23.61	0.25
Arm pointing	164	% Trials correct	69.41	18.88	16.67	100	1.47
Communicative marker	166	% Trials correct	76.21	17.75	25	100	1.38
Odour control trials	163	% Trials correct	49.54	15.65	12.5	87.5	1.23
Memory (5 s)	165	% Trials correct	74.34	20.03	16.67	100	1.56
Memory (10 s)	163	% Trials correct	70.24	22.05	0	100	1.73
Memory (15 s)	102	% Trials correct	65.03	18.54	16.67	100	1.84
Memory (20 s)	60	% Trials correct	63.89	19.45	16.67	100	2.51
Visual discrimination	168	% Trials correct	91	12.75	50	100	0.98
Auditory discrimination	167	% Trials correct	58.91	19.79	12.5	100	1.53
Odour discrimination: first choice	164	% trials correct	53.76	20.83	16.67	100	1.63
Odour discrimination: final choice	164	% Trials correct	72.56	19.78	0	100	1.54
Odour discrimination: persistence	164	No. of seconds	79.83	22.37	20.94	125.46	1.75
Odour discrimination: time at correct response	164	No. of seconds	61.56	22.32	13.9	116.43	1.74
Odour discrimination: time at incorrect response	164	No. of seconds	18.27	10.14	1.44	66.53	0.79

succeeding at visual accommodation tests as well as visible displacement tests of object permanence.

Primary DCDB Measures

On the 14 primary DCDB measures, 95% of the puppies who participated (160/168) successfully completed every task in the battery (with the exception of the human interest task, which was added shortly after data collection began and was successfully completed by 150/150 puppies).

Retrieval

Across two 1 min trials, puppies exhibited a group-wide tendency to chase the ball and pick it up in their mouth. Retrieval is one of the most widely tested skills in large-scale studies of puppy behaviour (Asher et al., 2013; Goddard & Beilharz, 1986;

Pfaffenberger, Scott, Fuller, Ginsburg, & Bielfelt, 1976; Riemer et al., 2014; Scott & Bielfelt, 1976; Slabbert & Odendaal, 1999; Svobodová, Vápeník, Pinc, & Bartoš, 2008; Wilsson & Sundgren, 1998). Wilsson and Sundgren (1998), who used the same scoring system as applied in the current study, reported that their population of 167 8-week-old German shepherd puppies averaged a score of 4.1 (a score of 4 indicates retrieval of the ball one or two times), approximately 0.8 points higher than our population of 8- to 10-week-old retriever puppies. Similar to our study, they found no evidence for sex differences. Slabbert and Odendaal (1999) also tested German shepherd puppies on a retrieval task at 8 weeks of age and found that the variation in scores was significantly associated with the puppies' later efficiency as police dogs, wherein higher retrieval drive was predictive of success as a police dog.

Laterality

Per Tomkins et al. (2010), we calculated a laterality index:

$$\text{Laterality index (LI)} = \frac{R - L}{R + L} \times 100$$

where R is number of right paw uses and L is number of left paw uses. Therefore, positive LI values reflect a right-side bias, whereas negative values reflect a left-side bias. Bias strength was calculated as the absolute value of the LI. Again following Tomkins et al. (2010), significant preferences were those in which the absolute value of the Z score was > 1.96 , where $Z = (R - ((R + L)/2)) / \sqrt{((R + L)/4)}$. When considering the composite scores from all 30 trials (15 step-up trials and 15 step-down trials), 84 out of 168 puppies (50%) exhibited significant laterality biases. Although the population tended towards a left bias, the direction of laterality was not significantly different from 0 at the group level ($t_{167} = -1.93$,

Table 2
One-sample null hypothesis tests for Dog Cognitive Development Battery (DCDB) measures in puppies

Variable	Null hypothesis	Mean	t	df	P
Laterality: laterality index	0	-7.18	-1.93	167	0.06
Communicative marker	50	76.21	19.02	165	0.00
Arm pointing	50	69.41	13.17	163	0.00
Odour control trials	50	49.54	-0.38	162	0.71
Memory (5 s)	50	74.34	15.61	164	0.00
Memory (10 s)	50	70.24	11.72	162	0.00
Memory (15s)	50	65.03	8.19	101	0.00
Memory (20 s)	50	63.89	5.53	59	0.00
Visual discrimination	50	91	41.66	167	0.00
Auditory discrimination	50	58.91	5.82	166	0.00
Odour discrimination: first choice	50	53.76	2.31	163	0.02
Odour discrimination: final choice	50	72.56	14.61	163	0.00

$P = 0.06$). To our knowledge, our study is the first to assess motor laterality in puppies under 6 months of age, and it confirms the presence of significant laterality biases at this age. Given that past studies have found an association between adult motor laterality and assistance dog success in purpose-bred populations (Batt, Batt, Baguley, & McGreevy, 2008; Tomkins, Thomson, & McGreevy, 2012), future work will benefit from assessing whether laterality measured in puppies is similarly associated with adult training outcomes.

Human interest

In contrast to tasks that either presented unobtainable food (see unsolvable task below), held food near a human's face (Stewart et al., 2015) or specifically rewarded gazing behaviour with food (Bentosela, Wynne, D'Orazio, Elgier, & Udell, 2016; Lenkei, Pogány, & Fugazza, 2019), the human interest task was designed to explore spontaneous interest towards an attentive and communicative human. Over three 30 s test trials, puppies spent approximately 6 s per trial looking to the face of the human experimenter who was engaging in dog-directed speech. Over the three 30 s play breaks in between test trials, puppies spent approximately 18 s per play break in proximity to the human experimenter.

Cylinder inhibitory control and reversal learning

Puppies made correct choices on about half of the inhibitory control test trials. As expected if the inhibitory control manipulation affected task difficulty, the percentage of correct responses during the first four familiarization trials (with the opaque tube) was significantly greater than the percentage of correct responses during the first four inhibitory control trials (with the transparent tube; mean \pm SD first four familiarization trials: $49.10 \pm 24.60\%$; mean \pm SD first four inhibitory control trials: $34.79 \pm 27.08\%$; paired t test: $t_{165} = -5.19$, $P < 0.001$). On average, puppies required 3.93 ± 2.41 s to retrieve the treat on each inhibitory control trial. However, latency to solve the test trials varied as a function of age (β : -0.16 , $t = -3.19$, $P < 0.01$), with older puppies solving test trials more quickly on average. It is not known whether this difference in latency reflects cognitive or motoric factors (e.g. older puppies may simply locomote faster).

In the reversal trials of the cylinder task, the preferred side (which was then blocked) was determined as the side from which the puppy retrieved the food on the majority of the final three inhibitory control trials. On average, puppies went to their preferred side seven out of eight times during the inhibitory control trials, suggesting strong preferences. During the reversal trials, puppies made correct choices on roughly one-third of trials. Quantifying the strength of puppies' side preference during inhibitory control trials and relating it to their performance on reversal trials did show that puppies with the strongest prereversal side preference performed the worst during the reversal phase (Pearson's correlation: $r_{163} = -0.20$, $P < 0.05$), which is a limitation of this task. On average, subjects required 6.58 ± 4.51 s to retrieve the treat on each trial. We also tracked the pathway of the subject on each trial; if the puppy directly approached the side from which the treat was accessible without first walking past the side of the cylinder blocked by the barrier, we coded this as a 'first side correct' response. Puppies averaged $23.26 \pm 27.29\%$ first side correct responses over the course of the task. As expected if blocking their preferred pathway affected puppies' response strategies, the percentage of correct responses during the first four inhibitory control trials was significantly greater than the percentage of correct responses during the first four reversal trials (mean first four reversal trials: $13.18 \pm 19.82\%$; paired t test: $t_{164} = 8.67$, $P < 0.001$). Latency to solve the reversal trials also varied as a function of age (β : -0.30 , $t = -2.87$, $P < 0.01$), with older puppies solving reversal trials more

quickly. However, as noted above, it is not clear whether this effect reflects cognitive or motoric differences within the sample.

While a handful of other studies have investigated detouring behaviour in young puppies (Diederich & Giffroy, 2003; Fox & Stelzner, 1966; Wyrwicka, 1959), our study represents the largest sample to date and utilizes a well-established paradigm that has been successfully applied across adult dogs and other species (Bray et al., 2014; MacLean et al., 2014; Marshall-Pescini et al., 2015).

Unsolvable task

Over four 30 s trials, puppies tended to look at the experimenter's face for an average duration of 1 s per trial. They spent more time manipulating the locked container, for an average of about 13 s per trial. Our results are similar to another study that tested 8-week-old puppies on a version of the unsolvable task (Passalacqua et al., 2011). In that study, puppies participated in a single minute-long unsolvable trial, as compared to our four 30 s unsolvable trials, during which the puppies gazed at the human an average of 0.99 s. They also found that puppies spent more time interacting with the container than gazing towards the human: puppies interacted with the container for an average of 40.75–50.36 s, depending on breed (although differences in duration across breeds was not significant). In a similar experimental set-up involving inaccessible but visible food in the presence of a human, Gácsi et al. (2005) found that 5- and 9-week-old dog puppies increased their gazing behaviour towards the human over a span of 4 min, whereas 9-week-old identically reared wolf (*Canis lupus*) puppies did not. Two studies of gazing behaviour in adult dogs found that subjects looked to the human experimenter for closer to 2–6 s over a 2 min trial (Brubaker, Dasgupta, Bhattacharjee, Bhadra, & Udell, 2017; Miklósi et al., 2003), and a third study found adult dogs looked for an average of 5 s over a 1 min trial (Sommese, Nováková; Šebková, & Bartoš, 2019). Thus, our results corroborate other findings suggesting that young puppies do orient to the human during an unsolvable task, but less so than adult dogs.

Gesture use

Puppies performed above chance expectation on both tasks involving communicative cues from humans (Fig. 2).

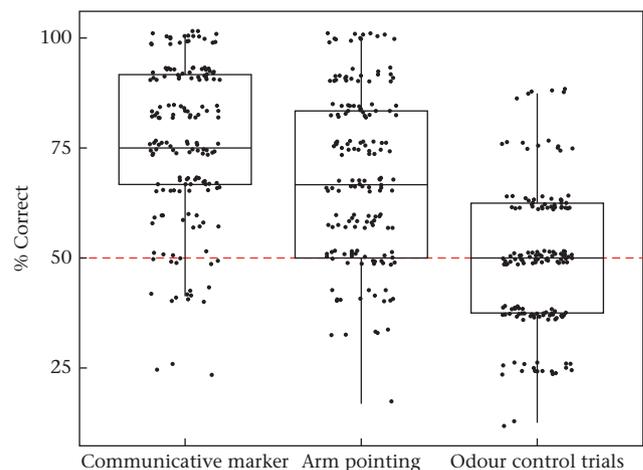


Figure 2. Performance on the gesture use (communicative marker and arm pointing) and odour control tasks. The data are presented in a box plot, overlaid with raw points. The lower and upper hinges correspond to the 25th and 75th percentiles, and the band inside the box is the median value. Chance performance (50%) is denoted by the red dashed line. Points are jittered to reduce overplotting.

Specifically, on average they made >75% correct choices based on the ostensive marker cue, and ~70% correct choices based on the ostensive pointing cue. Critically, when not provided with communicative cues, their performance was not different from chance expectation (49.54%; $t_{162} = -0.38$, $P = 0.71$). A one-way repeated measures ANOVA indicated that puppy performance was statistically different across the marker, pointing and control conditions ($F_{2, 324} = 109.6$, $P < 0.0001$, $\eta^2_g = 0.29$), with Bonferroni-adjusted post hoc tests revealing significant pairwise differences between all conditions ($P < 0.01$). These results confirm that puppies were not using olfactory cues in this context, consistent with previous findings in adult dogs (Bräuer et al., 2006; MacLean et al., 2017).

To assess the possibility of learning across test trials, we conducted logistic regressions with performance on each gesture task as the dependent measure and trial number as the predictor variable. Results revealed that for the arm-pointing task, trial number was not a significant predictor ($P = 0.052$) of performance, but there was a tendency for puppies to perform worse across time (estimate = -0.03). For the marker task, trial number was a significant predictor ($P = 0.02$), but the beta coefficient was negative (-0.04), indicating that the puppies' performance decreased over time. To assess first-trial performance, we performed a binomial test on the number of puppies who chose correctly on the first trial in the pointing and marker tasks. These tests indicated that puppies were correct on the first communicative marker trial 83% of the time (138/166), which was significantly higher than the null expectation ($P < 0.001$). Puppies were correct on the first arm pointing trial 77% of the time (126/164), also significantly higher than the null expectation ($P < 0.001$). In contrast, puppies were correct on the first odour control trial only 48% of the time (79/163), which was not significantly different than chance expectation ($P = 0.75$). Taken together, we found no evidence that overall performance in the pointing or marker tasks can be explained by learning during the task.

Past studies have found that adult dogs can successfully use a novel marker as a communicative cue to determine the location of a reward, although crucially if the dog is out of the room when the marker is placed they do not use it – the implication being that it is only a valuable cue for the dogs when used in a socially communicative manner (Agnetta et al., 2000; Riedel et al., 2006). The current study is the second to demonstrate that puppies younger than 15 weeks already possess this ability (Riedel et al., 2008). Similarly, adult dogs are known to reliably follow pointing gestures in a variety of contexts (Hare et al., 1998; Hare & Tomasello, 1999; Miklósi et al., 1998). Our results, derived from the largest sample size of puppies to date and including appropriate controls, extends previous work with puppies (Hare et al., 2002; Kaminski, Schulz, & Tomasello, 2012; Riedel et al., 2008) and confirms robust sensitivity to human communicative cues by 8–10 weeks of age.

Novel object

All raw variables from this task were coded from video. Based on previous studies using this measure (Bray, Cheney et al., 2017), we were particularly interested in the initial behavioural responses and vocalizations displayed towards the mechanical cat when the puppy was alone in the room with it, as well as the puppy's ability to recover from any fear response, either independently or with the encouragement of a human. The ethogram for the coded variables is presented in Supplementary Table S3, and the descriptive statistics for those variables are reported in Supplementary Table S10. To summarize the behavioural responses in the novel object task, we conducted a PCA on measures scored by ethogram (Supplementary Table S12). Parallel analysis suggested retention of three components, which were extracted using a direct oblimin

rotation. The first component was loaded highly by measures of approach, proximity, orienting and physical interaction with the novel object, which we interpreted as 'boldness'. The second component was loaded negatively by latency to vocalize and positively by the amount and intensity of vocalizations when alone with the novel object, which we interpreted as 'vocal intensity'. The third component was loaded positively by latency to approach and contact the object, as well as time orienting at the object, and negatively by the number of approaches to the object. We interpreted this component as 'caution'.

Other studies of early puppy behaviour commonly include arena tests in which novel objects are scattered throughout an enclosed area (Fox & Spencer, 1969; Guardini et al., 2016; Wilsson & Sundgren, 1998), or encounters with toys that move around independently and erratically (Asher et al., 2013; Goddard & Beilharz, 1984; Riemer et al., 2014). In past studies that have similarly summarized responses into components scores, factors such as 'boldness' (Riemer et al., 2014) and 'fear of object tests' (Goddard & Beilharz, 1986) emerge. In a recent study of adolescent guide dogs, vocal behaviour towards a novel object, represented in the current study as 'vocal intensity', was strongly associated with ultimate performance in the guide dog programme (Bray, Cheney et al., 2017). Our results confirm individual variation along a shy–bold axis, as well as variation in vocal responses towards novel objects.

Working memory

At delays of 5 s and 10 s, puppies' performance in the memory task was significantly higher than chance expectation (5 s: $t_{164} = 15.61$, $P < 0.001$; 10 s: $t_{164} = 11.72$, $P < 0.001$; Fig. 3).

Puppies only participated in 15 s delay trials if they chose correctly on at least four of six 10 s trials. Out of the 163 puppies who completed 10 s delay trials, 102 puppies (63%) met the criterion to advance to 15 s delays. Similarly, puppies only participated in 20 s delay trials if they chose correctly on at least four of six 15 s trials. Out of the 102 puppies who completed 15 s delay trials, 60 puppies (59%) met the criterion to advance to 20 s delays. For the puppies who met the criterion to advance to longer delays, performance also exceeded chance expectation following 15 s and 20 s delays (15 s: $t_{101} = 8.19$, $P < 0.001$; 20 s: $t_{59} = 5.53$, $P < 0.001$). While the mean scores of the puppies who advanced to the 20 s delay memory trials (mean \pm SE: 63.89 ± 19.45) were equivalent to previously published adult levels (mean \pm SD: 62.52 ± 21.21 ; MacLean et al., 2017), only 36% of puppies tested met the criterion to be tested on such trials. See Supplementary Table S13 for a secondary analysis in which we calculated mean performance at each delay length for only the puppies who completed all four delay lengths ($N = 60$, 36% of entire sample).

Perceptual discriminations

Puppies performed above chance expectation on visual, auditory and olfactory perceptual discriminations (Fig. 3). Discrimination between the two response options was best for visual discriminations (91% correct), followed by olfactory (73% correct on final choice) and auditory discriminations (59% correct), respectively. Despite minimal initial preference in the olfactory discrimination task (54% correct, measured as the first response option approached), puppies spent the majority of the time investigating the baited option (62% of the 20 s trial time at the baited option versus 18% at the nonbaited option) and had chosen this option 73% of the time by the end of the trial. Previous work has shown that dogs' behaviour is guided by olfaction by 2 weeks of age, audition by 4 weeks of age and vision by 6 weeks of age (Lord, 2013), and the current study suggests that by 8 weeks of age, dogs are able to

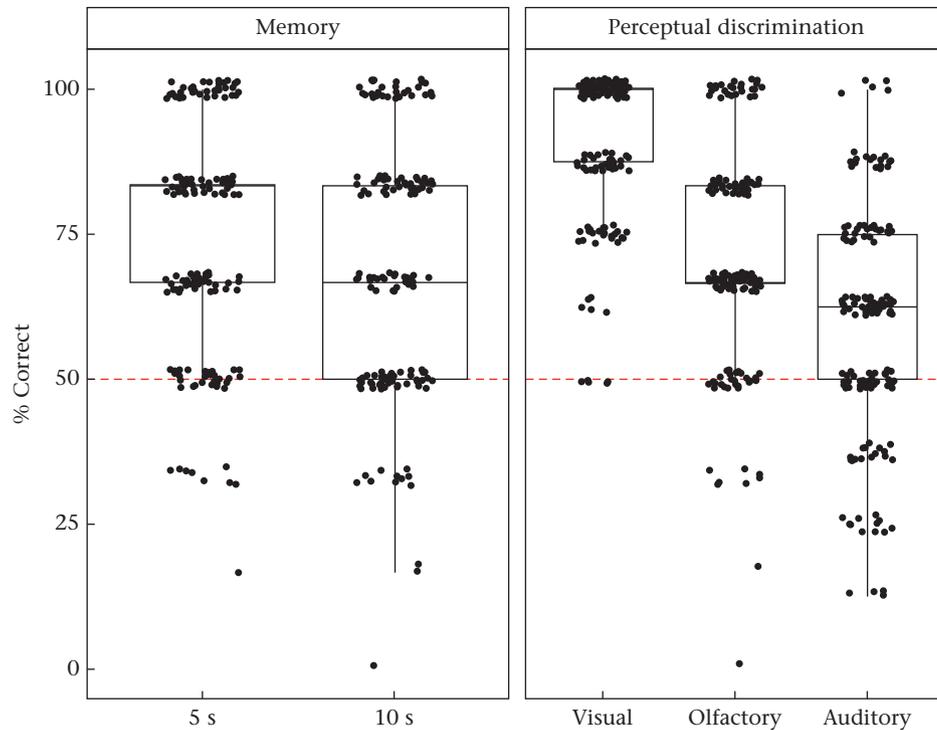


Figure 3. Performance on the memory task (at 5 s and 10 s delays) and perceptual discrimination (visual, olfactory, auditory) tasks. The data are presented in a box plot, overlaid with the original data points. The lower and upper hinges correspond to the 25th and 75th percentiles, and the band inside the box is the median value. Chance performance (50%) is denoted by the red dashed line. Points are jittered to reduce overplotting.

adequately use these senses to discriminate between multiple response options in an object choice task.

Surprising events: sudden appearance, looming object and loud noise

All raw variables from this task were coded from video. We were primarily interested in the immediate behavioural response of the puppies to each startling occurrence, as well as how well the puppies recovered on their own and with the encouragement of a human. The ethogram for the coded variables is presented in [Supplementary Table S4](#), and the descriptive statistics for those variables are reported in [Supplementary Table S11](#). As with the novel object task, we used PCA to develop a set of component measures summarizing behavioural variation in this test ([Supplementary Table S14](#)). Parallel analysis suggested retention of five components, which were extracted using a direct oblimin rotation. The first component was loaded by reactivity to the looming object (i.e. the umbrella), latency to approach the looming object and the duration of total contact with the looming object; we interpreted this component as reflecting ‘reactivity to the looming object’. The second component was loaded by reactivity to the loud noise (i.e. the shaking of the metal sheet), the vocal response to this event, the latency to vocalize and the intensity of those vocalizations when left alone for 1 min at the end of the task; we interpreted this component as reflecting ‘reactivity to sound and isolation’ (following the stressor). The third component was loaded by reactivity to the sudden appearance of the object (i.e. the trash bag), latency to approach this object after its appearance and the directness of this approach; we interpreted this component as ‘reactivity and recovery to the suddenly appearing object’. The fourth component was loaded by measures involving proximity to the suddenly appearing object, which we interpreted as ‘proximity to visual startle stimuli’. The final component was loaded by latency, distance from and path

to approach the loud noise stimulus (i.e. the metal sheet); we interpreted this component as ‘recovery from startling sound’.

Similar to our novel object task, temperament tests involving startling stimuli are commonly used in puppy tests to assess reactivity and recovery at an early age, especially in working dog populations (e.g. [Goddard & Beilharz, 1984](#); [Slabbert & Odendaal, 1999](#)). In one such study, [Svobodová et al. \(2008\)](#) found evidence for a ‘factor for responding to noise’ in their sample of 7-week-old German shepherd puppies, which resembles our ‘recovery from startling sound’ component. Moreover, a low score on this factor (i.e. dogs that were less responsive to distracting noise) was associated with later certification as a police dog, suggesting enduring influences of temperament across ontogeny.

GENERAL DISCUSSION

We tested a large sample of puppies ($N = 168$) in a novel battery of cognitive and temperamental tests to characterize the early ontogeny of dog cognition and temperament. Past studies in young puppies have explored some of the same traits that we measured, and where applicable our results are generally consistent with those of previous studies. However, one of the novel contributions of the current work is that many of the tasks, and particularly those designed to assess cognitive skills (e.g. working memory, impulse control and reversal learning as measured by the cylinder task, gesture use and perceptual discriminations), had not previously been assessed simultaneously in large samples of puppies. Thus, the results reported here significantly contribute to our understanding of the ontogeny of cognitive and temperamental traits in dogs through broad characterization of these processes across a large sample of puppies. Below we describe the major findings from this study and synthesize these results with the existing literature.

By 8–10 weeks of age, puppies consistently performed above chance expectation on tasks requiring visual, auditory and olfactory

perceptual discrimination and memory over short delays. Overall, skills for executive function – as measured by performance on the cylinder inhibitory control task, cylinder reversal learning task and delays of 15 s and 20 s during the working memory task – were present, but less developed than previously reported adult performance on similar tasks (Bray et al., 2014; Stewart et al., 2015).

Social communicative skills were also present, albeit less pronounced than in adults. While puppies did direct their gaze to human experimenters in social situations, they did so less than is commonly documented in adult dogs. Additionally, this gazing behaviour was more apparent when a person solicited the puppy's attention by using dog-directed speech (Ben-Aderet et al., 2017) than in situations in which gaze could be used as a potential help-soliciting behaviour. This finding is consistent with developmental patterns in human social cognition: infants also comprehend speech and gestures early (e.g. word comprehension, following pointing gestures) around 4–10 months of age (e.g. Bates & Dick, 2002; Bertenthal, Boyer, & Harding, 2014), while production of such behaviours (e.g. vocal and gestural naming, declarative pointing) develops later, around 1 year of age (e.g. Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Leung & Rheingold, 1981; Shore, Bates, Bretherton, Beeghly, & O'Connell, 1990). Similarly, while young puppies attend to speech and gestures directed towards them (as evidenced through looking time during the human interest task and success in social cue tasks), they initiate few attention-getting or communicative behaviours (e.g. eye contact, gaze alternation, barking) compared to what is commonly reported in studies with adults (Marshall-Pescini, Colombo, Passalacqua, Merola, & Prato-Previde, 2013; Miklósi et al., 2003). Interestingly, in the largest sample tested to date, we found that as early as 8 weeks of age, puppies are performing at above-chance levels on object choice tasks when provided with communicative signals from humans, including natural gestures such as arm pointing and novel communicative acts such as use of an arbitrary marker paired with ostensive signals. Furthermore, their performance in these contexts cannot simply be attributed to learning over the course of the task, as performance was high from the very first trial and did not improve over time. Thus, as suggested by previous studies with smaller samples (Gácsi, Gyoöri et al., 2009; Gácsi, Kara, Belényi, Topál, & Miklósi, 2009; Hare et al., 2002; Riedel et al., 2008; Rossano, Nitzschner, & Tomasello, 2014; Virányi et al., 2008), dogs appear to be biologically prepared for communication with humans, and these skills emerge early in ontogeny. Our design also addressed several important critiques and alternative interpretations of earlier work (Dorey et al., 2010). For example, we employed a more conservative choice criterion in which puppies were required to physically touch a container to indicate their choice (Riedel et al., 2008), rather than simply approaching to within 10 cm of it (Dorey et al., 2010; Hare et al., 2002). We also controlled for unintentional odour or behavioural cuing: all cups had a food reward taped into the bottom (Dorey et al., 2010; Miklósi et al., 1998), and all puppies participated in control trials in which the parameters of the task were identical but no social cue was given. Compared to previous studies, our sample was relatively homogenous and only included dogs between 8 and 10 weeks of age, and there was no effect of age on performance within this sample. Furthermore, these puppies were among the youngest tested in the literature, and all testing took place prior to the dogs being placed in homes with puppy raisers. Thus, subjects were still spending the majority of their time (e.g. sleeping, eating, socializing) with littermates as opposed to humans, but nevertheless were highly sensitive to human communicative signals.

Many of the traits measured in this study have previously been studied in adolescent working dogs and linked to subsequent

training outcomes. For example, in military, police and detection dogs, success was predicted by boldness, memory and the ability to use human social cues (Jamieson, Baxter, & Murray, 2017; MacLean & Hare, 2018; Svartberg, 2002; Wilsson & Sundgren, 1997). In guide dogs, motor lateralization, less vocalization during testing, lower levels of anxiety and fear, and superior problem-solving abilities were predictive of placement (Batt et al., 2008; Bray, Cheney et al., 2017; Harvey et al., 2016; Tomkins et al., 2012). In assistance dogs (MacLean & Hare, 2018) and some populations of detection dogs (Lazarowski et al., 2019), those who initiated high levels of eye contact were most likely to successfully complete training. The work presented here suggests that many of these cognitive and temperamental traits emerge early in development and are quantifiable using the Dog Cognitive Development Battery. Thus, future work will benefit by investigating both the stability of these traits across development and whether phenotypic measures from puppies are associated with training outcomes in adulthood. In the future, it will also be informative to explore correlations between traits to determine the underlying structure of individual differences on these measures (MacLean et al., 2017). Collectively, this work has the potential to enhance our understanding of the processes through which adult cognitive and behavioural phenotypes arise and to inform the practices through which these traits are measured across dog development. Because our sample was restricted to retriever dogs from a working dog population, it will also be important to test whether the developmental patterns observed here generalize to more heterogenous samples of pet and free-ranging dogs. Finally, to test the impact of domestication on these patterns, it will be crucial to further compare the cognitive development of wolves and dogs, as well as to probe patterns of variation across dog breeds. Together, this work will provide important insights regarding phylogenetic and developmental influences on canine cognition.

Conflict of Interest

We declare no conflicts of interest.

Acknowledgments

We thank Ben Allen, Erika Albrecht, Kacie Bauer, Ashtyn Bernard, James Brooks, Molly Byrne, Meg Callahan, Elizabeth Carranza, Averill Cantwell, Mary Chiang, Allison Doty, Julia Kemper, Lindsey Lang, Jessica Nelson, Gianna Ossello, Ashley Ryan, Holland Smith, Paige Smith and Mia Wesselkamper for help with data collection and coding. We thank Terry Henry, Kyria Henry, Jill Rosenblum, Erica Steele and Cece McConnell of paws4people foundation for facilitating longitudinal pilot testing with their assistance dog puppies. We are also grateful to the staff of Canine Companions for Independence and their dedicated volunteer breeder caretakers for accommodating 6 months of research with their assistance dog puppies at the Canine Early Development Center. This research was supported in part by grants from the Office of Naval Research (ONR N00014-17-1-2380 to E.M. and ONR N00014-16-1-2682 to B.H.), the AKC Canine Health Foundation (Grant No. 02518 to E.B. and E.M.) and the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health, United States (Award No. R01HD097732 to B.H.). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the views of the Foundation or the National Institutes of Health.

Supplementary Material

Supplementary material associated with this article is available, in the online version, at <https://doi.org/10.1016/j.anbehav.2020.05.019>.

References

- Agnetta, B., Hare, B., & Tomasello, M. (2000). Cues to food location that domestic dogs (*Canis familiaris*) of different ages do and do not use. *Animal Cognition*, 3(2), 107–112. <https://doi.org/10.1007/s100710000070>.
- Asher, L., Blythe, S., Roberts, R., Toothill, L., Craigh, P. J., Evans, K. M., et al. (2013). A standardized behavior test for potential guide dog puppies: Methods and association with subsequent success in guide dog training. *Journal of Veterinary Behavior*, 8(6), 431–438. <https://doi.org/10.1016/j.jveb.2013.08.004>.
- Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, 40(3), 293–310. <https://doi.org/10.1002/dev.10034>.
- Batt, L. S., Batt, M. S., Baguley, J. A., & McGreevy, P. D. (2008). Factors associated with success in guide dog training. *Journal of Veterinary Behavior*, 3(4), 143–151. <https://doi.org/10.1016/j.jveb.2008.04.003>.
- Batt, L. S., Batt, M. S., Baguley, J. A., & McGreevy, P. D. (2009). The relationships between motor lateralization, salivary cortisol concentrations and behavior in dogs. *Journal of Veterinary Behavior*, 4(6), 216–222. <https://doi.org/10.1016/j.jveb.2009.02.001>.
- Ben-Aderet, T., Gallego-Abenza, M., Reby, D., & Mathevon, N. (2017). Dog-directed speech: Why do we use it and do dogs pay attention to it? *Proceedings of the Royal Society B: Biological Sciences*, 284(1846), 20162429. <https://doi.org/10.1098/rspb.2016.2429>.
- 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>.
- Bentosela, M., Wynne, C., D'Orazio, M., Elgier, A., & Udell, M. A. (2016). Sociability and gazing toward humans in dogs and wolves: Simple behaviors with broad implications. *Journal of the Experimental Analysis of Behavior*, 105(1), 68–75. <https://doi.org/10.1002/jeab.191>.
- Berns, G. S., Brooks, A. M., Spivak, M., & Levy, K. (2017). Functional MRI in awake dogs predicts suitability for assistance work. *Scientific Reports*, 7(43704), 1–10. <https://doi.org/10.1038/srep43704>.
- Berenthal, B., Boyer, T., & Harding, S. (2014). When do infants begin to follow a point? *Developmental Psychology*, 50(8), 2036–2048. <https://doi.org/10.1037/a0037152>.
- Branson, N., & Rogers, L. (2006). Relationship between paw preference strength and noise phobia in *Canis familiaris*. *Journal of Comparative Psychology*, 120(3), 176–183. <https://doi.org/10.1037/0735-7036.120.3.176>.
- 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. <https://doi.org/10.1037/0735-7036.120.1.38>.
- Bray, E. E., Levy, K. M., Kennedy, B. S., Duffy, D. L., Serpell, J. A., & MacLean, E. L. (2019). Predictive models of assistance dog training outcomes using the Canine Behavioral Assessment and Research Questionnaire and a standardized temperament evaluation. *Frontiers in Veterinary Science*, 6, 49. <https://doi.org/10.3389/fvets.2019.00049>.
- Bray, E. E., MacLean, E. L., & Hare, B. (2014). Context specificity of inhibitory control in dogs. *Animal Cognition*, 17(1), 15–31. <https://doi.org/10.1007/s10071-013-0633-z>.
- Bray, E. E., MacLean, E. L., & Hare, B. (2015). Increasing arousal enhances inhibitory control in calm but not excitable dogs. *Animal Cognition*, 18(6), 1317–1329. <https://doi.org/10.1007/s10071-015-0901-1>.
- Bray, E. E., Sammel, M. D., Cheney, D. L., Serpell, J. A., & Seyfarth, R. M. (2017). The effects of maternal investment, temperament, and cognition on guide dog success. *Proceedings of the National Academy of Sciences of the United States of America*, 114(34), 9128–9133. <https://doi.org/10.1073/pnas.1704303114>.
- Bray, E. E., Sammel, M. D., Seyfarth, R. M., Serpell, J. A., & Cheney, D. L. (2017b). Temperament and cognition in a population of adolescent guide dogs. *Animal Cognition*, 20(5), 923–939. <https://doi.org/10.1007/s10071-017-1112-8>.
- Brubaker, L., Dasgupta, S., Bhattacharjee, D., Bhadra, A., & Udell, M. A. (2017). Differences in problem-solving between canid populations: Do domestication and lifetime experience affect persistence? *Animal Cognition*, 20(4), 717–723. <https://doi.org/10.1007/s10071-017-1093-7>.
- Brucks, D., Marshall-Pescini, S., Wallis, L. J., Huber, L., & Range, F. (2017). Measures of dogs' inhibitory control abilities do not correlate across tasks. *Frontiers in Psychology*, 8, 849. <https://doi.org/10.3389/fpsyg.2017.00849>.
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), i–174.
- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago, IL: University of Chicago Press.
- Diederich, C., & Giffroy, J.-M. (2003). How do dogs to avoid the collision with an obstacle? *Revue des Questions Scientifiques*, 174(1–2), 106–111.
- Doré, F. Y., Fiset, S., Goulet, S., Dumas, M.-C., & Gagnon, S. (1996). Search behavior in cats and dogs: Interspecific differences in working memory and spatial cognition. *Animal Learning & Behavior*, 24(2), 142–149.
- Dorey, N. R., Udell, M. A., & Wynne, C. D. (2010). When do domestic dogs, *Canis familiaris*, start to understand human pointing? The role of ontogeny in the development of interspecies communication. *Animal Behaviour*, 79(1), 37–41. <https://doi.org/10.1016/j.anbehav.2009.09.032>.
- Fagnani, J., Barrera, G., Carballo, F., & Bentosela, M. (2016). Is previous experience important for inhibitory control? A comparison between shelter and pet dogs in A-not-B and cylinder tasks. *Animal Cognition*, 19(6), 1165–1172. <https://doi.org/10.1007/s10071-016-1024-z>.
- Fiset, S., Beaulieu, C., & Landry, F. (2003). Duration of dogs' (*Canis familiaris*) working memory in search for disappearing objects. *Animal Cognition*, 6(1), 1–10. <https://doi.org/10.1007/s10071-002-0157-4>.
- Fox, M., & Spencer, J. (1969). Exploratory behavior in the dog: Experiential or age dependent? *Developmental Psychobiology*, 2(2), 68–74.
- Fox, M., & Stelzner, D. (1966). Behavioural effects of differential early experience in the dog. *Animal Behaviour*, 14(2–3), 273–281.
- Gácsi, M., Györi, B., Virányi, Z., Kubinyi, E., Range, F., Belényi, B., et al. (2009). Explaining dog wolf differences in utilizing human pointing gestures: Selection for synergistic shifts in the development of some social skills. *PLoS One*, 4(8). <https://doi.org/10.1371/journal.pone.0006584>. e6584.
- Gácsi, M., Kara, E., Belényi, B., Topál, J., & Miklósi, Á. (2009). The effect of development and individual differences in pointing comprehension of dogs. *Animal Cognition*, 12(3), 471–479. <https://doi.org/10.1007/s10071-008-0208-6>.
- Gácsi, M., Györi, B., Miklósi, Á., Virányi, Z., Kubinyi, E., Topál, J., et al. (2005). Species-specific differences and similarities in the behavior of hand-raised dog and wolf pups in social situations with humans. *Developmental Psychobiology*, 47(2), 111–122. <https://doi.org/10.1002/dev.20082>.
- Gagnon, S., & Doré, F. Y. (1994). Cross-sectional study of object permanence in domestic puppies (*Canis familiaris*). *Journal of Comparative Psychology*, 108(3), 220. <https://doi.org/10.1037/0735-7036.108.3.220>.
- Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific Reports*, 7(1), 1–10.
- Gnanadesikan, G. E., Hare, B., Snyder-Mackler, N., & MacLean, E. L. (2020). Estimating the heritability of cognitive traits across dog breeds reveals highly heritable inhibitory control and communication factors. *Animal Cognition*. <https://doi.org/10.1007/s10071-020-01400-4> (in press).
- Goddard, M., & Beilharz, R. (1984). A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science*, 12(3), 253–265. [https://doi.org/10.1016/0168-1591\(84\)90118-7](https://doi.org/10.1016/0168-1591(84)90118-7).
- Goddard, M., & Beilharz, R. (1986). Early prediction of adult behaviour in potential guide dogs. *Applied Animal Behaviour Science*, 15(3), 247–260. [0.1016/0168-1591\(86\)90095-X](https://doi.org/10.1016/0168-1591(86)90095-X).
- Guardini, G., Mariti, C., Bowen, J., Fatjó, J., Ruzzante, S., Martorell, A., et al. (2016). Influence of morning maternal care on the behavioural responses of 8-week-old beagle puppies to new environmental and social stimuli. *Applied Animal Behaviour Science*, 181, 137–144. <https://doi.org/10.1016/j.applanim.2016.05.006>.
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, 298(5598), 1634–1636. <https://doi.org/10.1126/science.1072702>.
- Hare, B., Call, J., & Tomasello, M. (1998). Communication of food location between human and dog (*Canis familiaris*). *Evolution of Communication*, 2(1), 137–159. <https://doi.org/10.1075/eoc.2.1.06har>.
- Hare, B., Rosati, A., Kaminski, J., Bräuer, J., Call, J., & Tomasello, M. (2010). The domestication hypothesis for dogs' skills with human communication: A response to Udell et al. (2008) and Wynne et al. (2008). *Animal Behaviour*, 79(2), e1–e6. <https://doi.org/10.1016/j.anbehav.2009.06.031>.
- Hare, B., & Tomasello, M. (1999). Domestic dogs (*Canis familiaris*) use human and conspecific social cues to locate hidden food. *Journal of Comparative Psychology*, 113(2), 173–177. <https://doi.org/10.1037/0735-7036.113.2.173>.
- Hare, B., & Tomasello, M. (2005). The emotional reactivity hypothesis and cognitive evolution: Reply to Miklósi and Topál. *Trends in Cognitive Sciences*, 9(10), 464–465.
- Hare, B., & Woods, V. (2013). *The genius of dogs: How dogs are smarter than you think*. New York, NY: Penguin Group.
- Hare, B., & Wrangham, R. (2002). Integrating two evolutionary models for the study of social cognition. In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 363–369). Cambridge, MA: MIT Press.
- Harvey, N. D., Craigh, P. J., Sommerville, R., McMillan, C., Green, M., England, G. C., et al. (2016). Test–retest reliability and predictive validity of a juvenile guide dog behavior test. *Journal of Veterinary Behavior*, 11, 65–76. <https://doi.org/10.1016/j.jveb.2015.09.005>.
- Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317(5843), 1360–1366. <https://doi.org/10.1126/science.1146282>.
- Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the cognitive skills of bonobos and chimpanzees. *PLoS One*, 5(8). <https://doi.org/10.1371/journal.pone.0012438>. e12438.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185.
- Horschler, D. J., Hare, B., Call, J., Kaminski, J., Miklósi, Á., & MacLean, E. L. (2019). Absolute brain size predicts dog breed differences in executive function. *Animal Cognition*, 22(2), 187–198. <https://doi.org/10.1007/s10071-018-01234-1>.

- Jamieson, L. T. J., Baxter, G. S., & Murray, P. J. (2017). Identifying suitable detection dogs. *Applied Animal Behaviour Science*, 195, 1–7. <https://doi.org/10.1016/j.applanim.2017.06.010>.
- Kabadayi, C., Bobrowicz, K., & Osvath, M. (2018). The detour paradigm in animal cognition. *Animal Cognition*, 21(1), 21.
- Kaminski, J., Schulz, L., & Tomasello, M. (2012). How dogs know when communication is intended for them. *Developmental Science*, 15(2), 222–232. <https://doi.org/10.1111/j.1467-7687.2011.01120.x>.
- King, T., Hemsworth, P., & Coleman, G. (2003). Fear of novel and startling stimuli in domestic dogs. *Applied Animal Behaviour Science*, 82(1), 45–64. [https://doi.org/10.1016/S0168-1591\(03\)00040-6](https://doi.org/10.1016/S0168-1591(03)00040-6).
- Lazarowski, L., Strassberg, L. R., Waggoner, L. P., & Katz, J. S. (2019a). Persistence and human-directed behavior in detection dogs: Ontogenetic development and relationships to working dog success. *Applied Animal Behaviour Science*, 220, 104860.
- Lazarowski, L., Waggoner, P., & Katz, J. S. (2019b). The future of detector dog research. *Comparative Cognition & Behavior Reviews*, 14, 77–80. <https://doi.org/10.3819/CCBR.2019.140008>.
- Lazzaroni, M., Marshall-Pescini, S., Manzenreiter, H., Gosch, S., Přibilová, L., Darc, L., et al. (2020). Why do dogs look back at the human in an impossible task? Looking back behaviour may be over-interpreted. *Animal Cognition*, 23(3), 427–441.
- Lenkei, R., Pogány, Á., & Fugazza, C. (2019). Social behavior in dog puppies: Breed differences and the effect of rearing conditions. *Biologia Futura*, 70(2), 134–142. <https://doi.org/10.1556/019.70.2019.17>.
- Leung, E. H., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology*, 17(2), 215–220. <https://doi.org/10.1037/0012-1649.17.2.215>.
- Lord, K. (2013). A comparison of the sensory development of wolves (*Canis lupus lupus*) and dogs (*Canis lupus familiaris*). *Ethology*, 119(2), 110–120. <https://doi.org/10.1111/eth.12044>.
- MacLean, E. L., & Hare, B. (2018). Enhanced selection of assistance and explosive detection dogs using cognitive measures. *Frontiers in Veterinary Science*, 5(236), 1–14. <https://doi.org/10.3389/fvets.2018.00236>.
- MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., et al. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences of the United States of America*, 111(20), E2140–E2148. <https://doi.org/10.1073/pnas.1323533111>.
- MacLean, E. L., Herrmann, E., Suchindran, S., & Hare, B. (2017). Individual differences in cooperative communicative skills are more similar between dogs and humans than chimpanzees. *Animal Behaviour*, 126, 41–51. <https://doi.org/10.1016/j.anbehav.2017.01.005>.
- MacLean, E. L., Matthews, L. J., Hare, B. A., Nunn, C. L., Anderson, R. C., Aureli, F., et al. (2012). How does cognition evolve? Phylogenetic comparative psychology. *Animal Cognition*, 15(2), 223–238. <https://doi.org/10.1007/s10071-011-0448-8>.
- MacLean, E. L., & Nunn, C. L. (2017). Phylogenetic approaches for research in comparative cognition. In J. Call, G. M. Burghardt, I. M. Pepperberg, C. Snowdon, & T. Zentall (Eds.), *APA handbook of comparative psychology: Basic concepts, methods, neural substrate, and behavior* (Vol. 1, pp. 201–216). Washington, D.C.: American Psychological Association.
- MacNamara, M., & MacLean, E. (2017). Selecting animals for education environments. In N. R. Gee, A. H. Fine, & P. McCordle (Eds.), *How animals help students learn: Research and practice for educators and mental-health professionals* (pp. 182–196). New York, NY: Routledge.
- Marshall-Pescini, S., Colombo, E., Passalacqua, C., Merola, I., & Prato-Previde, E. (2013). Gaze alternation in dogs and toddlers in an unsolvable task: Evidence of an audience effect. *Animal Cognition*, 16(6), 933–943. <https://doi.org/10.1007/s10071-013-0627-x>.
- Marshall-Pescini, S., Virányi, Z., Kubinyi, E., & Range, F. (2017). Motivational factors underlying problem solving: Comparing wolf and dog puppies' explorative and neophobic behaviors at 5, 6, and 8 weeks of age. *Frontiers in Psychology*, 8, 180. <https://doi.org/10.3389/fpsyg.2017.00180>.
- Marshall-Pescini, S., Virányi, Z., & Range, F. (2015). The effect of domestication on inhibitory control: Wolves and dogs compared. *PLoS One*, 10(2). <https://doi.org/10.1371/journal.pone.0118469>. e0118469.
- McGarrity, M. E., Sinn, D. L., & Gosling, S. D. (2015). Which personality dimensions do puppy tests measure? A systematic procedure for categorizing behavioral assays. *Behavioural Processes*, 110, 117–124. <https://doi.org/10.1016/j.beproc.2014.09.029>.
- Miklósi, Á. (2014). *Dog behaviour, evolution, and cognition*. Oxford, U.K.: Oxford University Press.
- Miklósi, Á., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z., & Csányi, V. (2003). A simple reason for a big difference: Wolves do not look back at humans, but dogs do. *Current Biology*, 13(9), 763–766. [https://doi.org/10.1016/S0960-9822\(03\)00263-X](https://doi.org/10.1016/S0960-9822(03)00263-X).
- Miklósi, Á., Polgárdi, R., Topál, J., & Csányi, V. (1998). Use of experimenter-given cues in dogs. *Animal Cognition*, 1(2), 113–121. <https://doi.org/10.1007/s100710050016>.
- Ollivier, F., Plummer, C., & Barrie, K. (2007). Ophthalmic examination and diagnostics. Part 1: The eye examination and diagnostic procedures. In K. N. Gelatt (Ed.), *Veterinary ophthalmology* (4th ed., pp. 438–483). Ames, IA: Blackwell.
- Osborne, J. W., & Costello, A. B. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10(7), 1–9.
- Passalacqua, C., Marshall-Pescini, S., Barnard, S., Lakatos, G., Valsecchi, P., & Previde, E. P. (2011). Human-directed gazing behaviour in puppies and adult dogs, *Canis lupus familiaris*. *Animal Behaviour*, 82(5), 1043–1050. <https://doi.org/10.1016/j.anbehav.2011.07.039>.
- Pfaffenberger, C. J., Scott, J. P., Fuller, J. L., Ginsburg, B. E., & Bielfelt, S. W. (1976). *Guide dogs for the blind: Their selection, development, and training*. Amsterdam, The Netherlands: Elsevier.
- R Development Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>.
- Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research* (R package Version 1.9.12). Retrieved from <https://CRAN.R-project.org/package=psych>.
- Riedel, J., Buttelmann, D., Call, J., & Tomasello, M. (2006). Domestic dogs (*Canis familiaris*) use a physical marker to locate hidden food. *Animal Cognition*, 9(1), 27–35. <https://doi.org/10.1007/s10071-005-0256-0>.
- 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>.
- Riemer, S., Müller, C., Virányi, Z., Huber, L., & Range, F. (2014). The predictive value of early behavioural assessments in pet dogs: A longitudinal study from neonates to adults. *PLoS One*, 9(7). <https://doi.org/10.1371/journal.pone.0101237>. e101237.
- Rossano, F., Nitzschner, M., & Tomasello, M. (2014). Domestic dogs and puppies can use human voice direction referentially. *Proceedings of the Royal Society B: Biological Sciences*, 281(1785), 20133201. <https://doi.org/10.1098/rspb.2013.3201>.
- Schmitt, V., Pankau, B., & Fischer, J. (2012). Old world monkeys compare to apes in the primate cognition test battery. *PLoS One*, 7(4). <https://doi.org/10.1371/journal.pone.0032024>. e32024.
- Scott, J. P., & Bielfelt, S. W. (1976). Analysis of the puppy-testing program. In C. J. Pfaffenberger, J. P. Scott, J. L. Fuller, B. E. Ginsburg, & S. W. Bielfelt (Eds.), *Guide dogs for the blind: Their selection, development and training* (pp. 39–76). Amsterdam, The Netherlands: Elsevier.
- Shaw, R. C., & Schmelz, M. (2017). Cognitive test batteries in animal cognition research: Evaluating the past, present and future of comparative psychometrics. *Animal Cognition*, 20(6), 1003–1018. <https://doi.org/10.1007/s10071-017-1135-1>.
- Sherman, B. L., Gruen, M. E., Case, B. C., Foster, M. L., Fish, R. E., Lazarowski, L., et al. (2015). A test for the evaluation of emotional reactivity in Labrador retrievers used for explosives detection. *Journal of Veterinary Behavior*, 10(2), 94–102. <https://doi.org/10.1016/j.jvbeh.2014.12.007>.
- Shettleworth, S. J. (2013). *Fundamentals of comparative cognition*. New York, NY: Oxford University Press.
- Shore, C., Bates, E., Bretherton, I., Beeghly, M., & O'Connell, B. (1990). Vocal and gestural symbols: Similarities and differences from 13 to 28 months. In V. Volterra, & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 79–91). Berlin, Germany: Springer.
- Slabbert, J., & Odendaal, J. (1999). Early prediction of adult police dog efficiency: A longitudinal study. *Applied Animal Behaviour Science*, 64(4), 269–288. [https://doi.org/10.1016/S0168-1591\(99\)00038-6](https://doi.org/10.1016/S0168-1591(99)00038-6).
- Sommese, A., Nováková, K., Šebková, N. F., & Bartoš, L. (2019). A wolfdog point of view on the impossible task paradigm. *Animal Cognition*, 22(6), 1073–1083. <https://doi.org/10.1007/s10071-019-01298-7>.
- Stewart, L., MacLean, E. L., Ivy, D., Woods, V., Cohen, E., Rodriguez, K., et al. (2015). Citizen science as a new tool in dog cognition research. *PLoS One*, 10(9). <https://doi.org/10.1371/journal.pone.0135176>. e0135176.
- Svartberg, K. (2002). Shyness–boldness predicts performance in working dogs. *Applied Animal Behaviour Science*, 79(2), 157–174. [https://doi.org/10.1016/S0168-1591\(02\)00120-X](https://doi.org/10.1016/S0168-1591(02)00120-X).
- Svobodová, I., Vápeník, P., Pinc, L., & Bartoš, L. (2008). Testing German shepherd puppies to assess their chances of certification. *Applied Animal Behaviour Science*, 113(1), 139–149. <https://doi.org/10.1016/j.applanim.2007.09.010>.
- Tapp, P. D., Siwak, C. T., Estrada, J., Head, E., Muggenburg, B. A., Cotman, C. W., et al. (2003a). Size and reversal learning in the beagle dog as a measure of executive function and inhibitory control in aging. *Learning & Memory*, 10(1), 64–73.
- Tapp, P. D., Siwak, C. T., Estrada, J., Holowachuk, D., & Milgram, N. W. (2003b). 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.
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20(4), 410–433.
- Tomasello, M., & Call, J. (1997). *Primate cognition*. New York, NY: Oxford University Press.
- Tomkins, L. M., Thomson, P. C., & McGreevy, P. D. (2010). First-stepping Test as a measure of motor laterality in dogs (*Canis familiaris*). *Journal of Veterinary Behavior: Clinical Applications and Research*, 5(5), 247–255. <https://doi.org/10.1016/j.jvbeh.2010.03.001>.
- Tomkins, L. M., Thomson, P. C., & McGreevy, P. D. (2012). Associations between motor, sensory and structural lateralisation and guide dog success. *Veterinary Journal*, 192(3), 359–367. <https://doi.org/10.1016/j.tvjl.2011.09.010>.
- van der Borg, J. A., Netto, W. J., & Planta, D. J. (1991). Behavioural testing of dogs in animal shelters to predict problem behaviour. *Applied Animal Behaviour Science*, 32(2), 237–251. [https://doi.org/10.1016/S0168-1591\(05\)80047-4](https://doi.org/10.1016/S0168-1591(05)80047-4).
- van Horik, J. O., Beardsworth, C. E., Laker, P. R., Whiteside, M. A., & Madden, J. R. (2020). Response learning confounds assays of inhibitory control on detour tasks. *Animal Cognition*, 23(1), 215–225.

- van Horik, J. O., Langley, E. J., Whiteside, M. A., Laker, P. R., Beardsworth, C. E., & Madden, J. R. (2018). Do detour tasks provide accurate assays of inhibitory control? *Proceedings of the Royal Society B: Biological Sciences*, 285(1875), 20180150.
- Virányi, Z., Gácsi, M., Kubinyi, E., Topál, J., Belényi, B., Ujfalussy, D., et al. (2008). Comprehension of human pointing gestures in young human-reared wolves (*Canis lupus*) and dogs (*Canis familiaris*). *Animal Cognition*, 11(3), 373–387. <https://doi.org/10.1007/s10071-007-0127-y>.
- Wilsson, E., & Sundgren, P.-E. (1997). The use of a behaviour test for the selection of dogs for service and breeding. I: Method of testing and evaluating test results in the adult dog, demands on different kinds of service dogs, sex and breed differences. *Applied Animal Behaviour Science*, 53(4), 279–295. [https://doi.org/10.1016/S0168-1591\(96\)01174-4](https://doi.org/10.1016/S0168-1591(96)01174-4).
- Wilsson, E., & Sundgren, P.-E. (1998). Behaviour test for eight-week old puppies: Heritabilities of tested behaviour traits and its correspondence to later behaviour. *Applied Animal Behaviour Science*, 58(1), 151–162. [https://doi.org/10.1016/S0168-1591\(97\)00093-2](https://doi.org/10.1016/S0168-1591(97)00093-2).
- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (2014). Differences in the early cognitive development of children and great apes. *Developmental Psychobiology*, 56(3), 547–573. <https://doi.org/10.1002/dev.21125>.
- Wynne, C. D., Udell, M. A., & Lord, K. A. (2008). Ontogeny's impacts on human–dog communication. *Animal Behaviour*, 76(4), e1–e4. <https://doi.org/10.1016/j.anbehav.2008.03.010>.
- Wyrwicka, W. (1959). Studies on detour behaviour. *Behaviour*, 14(3), 240–264.