REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
07/26/2019	Final	05/01/2016-4/30/2019

4. TITLE AND SUBTITLE
Ontogeny and Heritability of Cognitive and Behavioral Traits Linked to Success as a Military Working Dog

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-16-1-2603

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
MacLean, Evan, L. & Bray, Emily, E.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DUKE UNIVERSITY
2200 W MAIN ST STE 710
DURHAM, NC 27708-4677
UNITED STATES

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
875 N. Randolph Street Suite 1425
Arlington VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
By 9 weeks of age, dogs performed above chance expectation on a variety of tasks involving short-term memory, use of human social cues, and perceptual discriminations. Across development, individual difference in traits including spontaneous interest in humans, inhibitory control, and use of human communicative signals exhibited rank-order stability, suggesting consistent individual differences across time. Heritability analyses revealed that a substantial proportion of phenotypic variance was attributable to genetic factors, both in early ontogeny and young adulthood.

15. SUBJECT TERMS
Military working dog, cognition, temperament, olfaction, heritability, ontogeny, selection, breeding, screening

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Evan MacLean
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	SAR	81	520-621-0386

Final Technical Report
N00014-16-1-2603

Ontogeny and Heritability of Cognitive and Behavioral Traits Linked to Success as a Military Working Dog

Evan L. MacLean & Emily E. Bray

Contact Information:
Evan MacLean, Ph.D.
University of Arizona
Email: evanmaclean@email.arizona.edu
Phone: 520-621-0386

Table of Contents

Tables and Figures

Tables

Figures

EXECUTIVE SUMMARY

Military working dogs (MWDs) play diverse and critical roles in the United States Armed Forces, ranging from the detection of improvised explosive devices (IEDs) to the tracking and capture of enemy combatants. To be successful in these roles dogs require a unique suite of temperamental traits and cognitive skills that allow them to function effectively with a human handler under demanding environmental conditions. Previously, our research team identified cognitive and temperamental differences between the most and least successful MWDs and developed and validated a novel test battery for measuring these traits. However, the causes of individual differences in cognitive and behavioral traits have remained elusive. To address the nature and origins of individual differences we conducted a longitudinal study of the ontogeny and heritability of cognitive and temperamental traits in a working dog population. For this purpose, we developed the Dog Cognitive Development Battery (DCDB), an instrument measuring diverse cognitive and temperamental processes, which can be deployed with dogs eight weeks of age and older. We then used the DCDB with a study cohort who were tested ~9-weeks of age, and again ~18 months of age, following rearing in diverse environments by volunteer raisers throughout the western United States. We assessed (1) the early ontogeny of cognitive and behavioral traits, (2) patterns of change across development into young adulthood (3) trait stability across ontogeny, and (4) heritability of these traits. We found that by 9 weeks of age, dogs performed above chance expectation on a variety of tasks involving short-term memory, use of human social cues, and perceptual discriminations. Whereas perceptual discriminations reached levels approximating those of adults by 9 weeks, cognitive processes involving social communication, memory, and executive function developed more slowly. Across development, individual difference in traits including spontaneous interest in humans, inhibitory control, and use of human communicative signals exhibited rank-order stability, suggesting consistent individual differences across time. Heritability analyses revealed that a substantial proportion of phenotypic variance was attributable to genetic factors, both in early ontogeny and young adulthood (despite the rearing of littermates in diverse environments). Collectively, our results contribute to our understanding of the causes of individual differences in cognitive and temperamental traits and address the extent to which favorable traits can be either bred, or developmentally screened for, in working dog populations.

INTRODUCTION

Military working dogs (MWDs) play diverse and critical roles in the United States Armed Forces, ranging from the detection of improvised explosive devices (IEDs) to the tracking and capture of enemy combatants. To be successful in these roles, dogs require a unique suite of temperamental traits and cognitive skills that allow them to function effectively with a human handler under demanding environmental conditions. For example, dogs must attend to gestural and verbal communication from a handler while simultaneously exhibiting enough independence to confidently search an area of interest – potentially at great distances from the handler. MWDs must search for their target in an efficient manner showing an ability to detour obstacles, remember locations previously searched, and flexibly navigate using allocentric search rules (as opposed to inflexible egocentric biases). Within and across species, animals, including humans, can differ in such cognitive abilities and these differences can explain variance in adult problem solving skills (for review, see Shettleworth, 2010). Likewise, cognitive differences between individual dogs can potentially explain differences in their ability to learn and utilize human commands in novel contexts (Hare & Woods, 2013; Miklósi, 2014).

Although animal trainers can shape behavior so that a dog associates the completion of a goal with a social or food reward (e.g. retrieve an object), trainers cannot train animals to flexibly and spontaneously respond to barriers that might prevent the completion of a trained goal (e.g. the dog must make a detour around a novel set of obstacles to search a new area while inhibiting responses to investigate novel odors and remembering the area previously covered to assure the most efficient search path; Bray, MacLean, & Hare, 2015). In these contexts, it is cognitive flexibility and not just trained responses that allow a dog to solve the problem. A dog's cognitive abilities allow it to mentally represent space and infer the need to take a detour (Fiset, Landry, & Ouellette, 2006; Pongracz et al., 2001); to categorize odors as relevant while inhibiting a response to irrelevant odor(s) (Kaulfuß & Mills, 2008; Lazarowski & Dorman, 2014; Lazarowski et al., 2015); to maintain a mental representation in working memory of the referent of a verbal command or visual signal (Kaminski, Call, & Fischer, 2004; Pilley & Reid, 2011); and to understand the communicative intention behind a human gesture or verbal command in directing them during their search (Hare & Tomasello, 2005; MacLean, Herrmann, Suchindran, & Hare, 2017; Miklósi, 2014). Given that cognitive variance helps explain individual differences in flexibly solving novel problems, it is crucial to assess how variance in cognitive abilities relates to the success of military working dogs. Although dogs are often regarded as a ‘sensor’ technology, the successful use of dog sensing capabilities relies on other psychological and behavioral factors that allow the animal to respond adaptively in complex and dynamic environments. For example, without requisite memory systems, executive functions, social communication, or a temperament that is robust to potentially fear-inducing stimuli, dogs cannot perform the sensing and communicative tasks that we rely on them for. Therefore, it remains a central basic research problem to better understand the nature of psychological and temperamental factors that facilitate and constrain these aspects of dog performance.

Previously our team conducted research to improve the identification and training of dogs for deployment as IED detection dogs (IDDs). To quantify individual variability as it relates to success in training, we designed and administered a battery of cognitive tests that evaluates the abilities of military working dogs across a broad range of cognitive domains (i.e. communication, navigation, memory, inhibitory control, etc.). We also measured a number of temperamental variables by exposing dogs to potentially threatening stimuli and documenting behavioral and physiological arousal under these conditions. We hypothesized that by relating individual differences in cognition and temperament to variance in success as an IDD, our research would reveal the central psychological mechanism(s) that successful explosive-detection dogs rely on or are constrained by during training and deployment. Consistent with this hypothesis, analyses across a large dataset of MWDs revealed consistent cognitive differences between the most and least successful working dogs (MacLean & Hare, 2018). Across discovery and replication studies, the cognitive traits most strongly linked to MWD success included measures of short-term memory, and sensitivity to human communication.

However, the patterns of individual differences that we observed raise fundamental basic research questions regarding the underlying causes of this variation. For example, to what extent is this variance explained by genetic factors? Are these differences predictable from early in ontogeny, or are they only expressed in the adult phenotype? The answers to these types of questions will provide foundational knowledge of how and why individual differences in dog cognition and temperament arise, as well as what approaches may be most productive for manipulating these processes. Given that dogs are used in a host of defense-related activities, we expect that these data will provide a foundation for future work that may aim to exploit particular genetic, developmental, or environmental processes in order to influence adult dog phenotypes.

For example, our previous research described above demonstrated that individual differences in adult dog cognition can be used to accurately predict aptitude for working roles. If these same types of predictive cognitive traits could be identified in puppies, training efforts could be focused on the subset of dogs with the highest potential for success, and training protocols could be implemented early in ontogeny when key cognitive and behavioral processes exhibit their greatest plasticity. Additionally, if these traits are strongly heritable (i.e. genetically influenced) these aspects of dog phenotypes could be further cultivated through selective breeding programs to create canine lines optimized for particular working roles. In contrast, if individual differences are largely explained by non-genetic factors, then specialized rearing and training protocols (e.g., Slabbert & Rasa, 1997) may provide the best opportunity to enhance MWD capabilities.

To address the ontogeny and heritability of cognitive traits in working dogs, we first developed the Dog Cognitive Development Battery, a series of cognitive tests that is deployable with both young puppies (~8-10 weeks) and adult dogs. We then implemented this test in a longitudinal study with a working dog population to address the following questions: (1) how does dog cognition change across development? (2) Do individual differences measured in puppies predict subsequent individual differences in adult phenotypes? (3) To what extent do genetic factors explain individual differences at both time points?

BEHAVIORAL METHODS

Subjects

Participating puppies were recruited through Canine Companions for Independence (Santa Rosa, CA, USA), 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 at the Canine Early Development Center (*n* = 21), a state-of-the-art facility with fulltime staff dedicated to monitoring and caring for the mothers and their litters. Puppies are weaned at 6 weeks of age, and remain housed socially with their littermates until around 8 weeks. At this point, all puppies spend time at Canine Companion's national headquarters where they undergo veterinary examinations prior to being sent to their volunteer puppy raisers for the next 14-22 months. Canine Companions granted informed consent to all aspects of the study. All testing procedures adhered to regulations set forth by the University of Arizona Institutional Animal Care and Use Committee (IACUC # 16-175).

We tested 168 puppies (97 females and 71 males) from February-July of 2017 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 x Golden crosses, 40 Labrador Retrievers, and 6 Golden Retrievers from 65 different litters.

Of the original 168 puppies, we were unable to re-test 8 as adults (*n* = 5 released for medical reasons prior to turn-in, *n* = 1 released for behavioral reasons prior to turn-in, *n* = 2 did not meet participation criteria at turn-in).We tested the remaining 160 dogs (93 females and 67 males) for a second time from January 2018-March 2019 when they were just under 2 years old (range 0.99 to 2.01 years, mean = 1.79 years), within a month of each dog returning to Canine Companions' Northwest and Southwest regional campuses for advanced training. The dogs that participated in this second round of testing included 118 Labrador x Golden crosses, 37 Labrador Retrievers, and 5 Golden Retrievers.

In addition to our longitudinal sample and in an effort to better inform our heritability analyses, we also tested 163 adult dogs (118 females and 45 males) in the Canine Companions breeding population from August 2017-May 2019, 90 of whom were the dams and sires of the dogs from the longitudinal sample. These dogs were slightly older at the time of testing (range 1.96 – 10.8 years, mean = 3.79 years), and included 71 Labrador x Golden crosses, 70 Labrador Retrievers, and 22 Golden Retrievers.

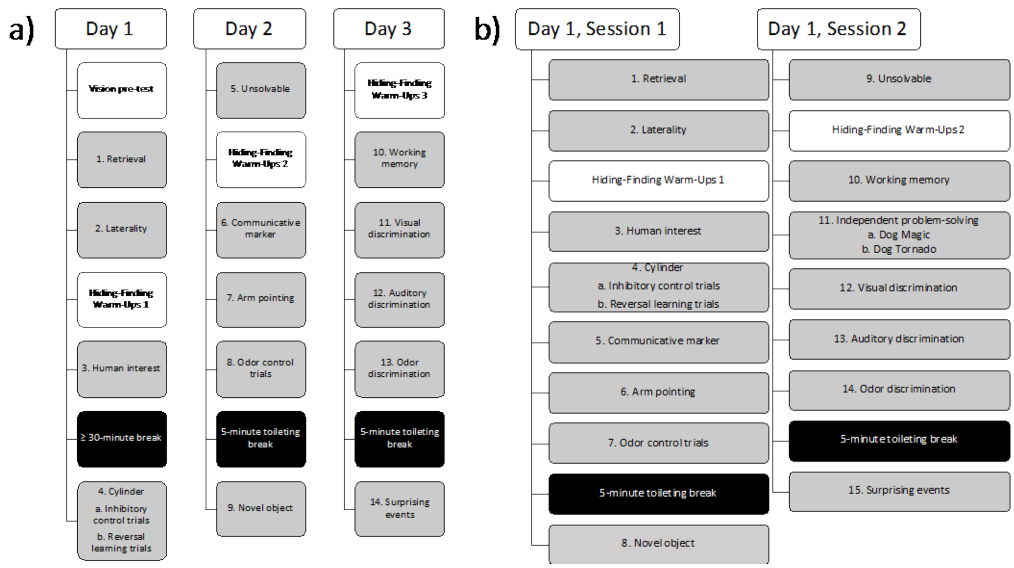
Procedure

Preliminary research methods (based on initial work with Paws4People) were piloted with a sample of 20 puppies between the ages of 7-12 weeks at CCI in December 2016. Pilot studies suggested that the majority of tasks were not feasible with CCI puppies younger than 8 weeks of age. Pilot subjects between 9-12 weeks 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 9-12 weeks, and thus we selected 9 weeks as our target age for the first assessment in longitudinal studies.

Dog Cognitive Development Battery (early ontogeny)

The developmental battery methods were finalized in late December 2016. Tasks retained in the final battery are briefly described below and detailed experimental methods are provided in the *Protocols* section. When working with 2-month-old puppies, the battery was implemented across three consecutive days with each subject, each session lasting ~45 minutes (Figure 1).

Figure 1. Design and implementation of Dog Cognitive Development Battery



Vision pre-test – a cotton ball is vertically dropped 100cm in front of the puppy, or silently flicked horizontally across the floor in view of the subject. This test confirms that subjects can visually track objects at the typical distance in which they are presented in subsequent measures.

(1) Retrieval – The experimenter throws a small ball for the puppy and vocally encourages the dog to bring the ball back. The number of times that the puppy will (a) chase the ball, (b) retrieve it in the mouth, and (c) return it to the experimenter are scored.

(2) Laterality – This task indexes behavioral measures of laterality, believed to reflect lateralization in the brain (previously linked to temperamental reactivity in adults). Puppies are held by the experimenter and then called to step onto and off of a platform. The forelimb used to initiate this motion is recorded across 30 trials, and subsequently used to compute a laterality index.

Hiding-finding warm-ups – Two opaque containers are placed in front of the puppy. The experimenter shows the puppy a food reward and places it underneath one of the containers. Puppies are required to choose correctly on 4/5 trials to advance to subsequent measures. Puppies complete this task once per session.

(3) Human interest – The experimenter stands outside the testing pen, looks at the puppy, and recites a predetermined script with a fluctuating, high-pitched intonation. The duration of the puppy's gaze to the human's face is recorded across trials.

(4a) Cylinder Inhibitory control and (4b) Cylinder Reversal learning – Puppies are familiarized with detouring around an opaque cylinder to retrieve a reward from within. In phase a test trials (inhibitory control), a transparent cylinder is used and subjects must resist the prepotent response to move directly toward the visible food, instead detouring around the obstacle. Detouring behavior is recorded across 8 trials. In phase b test trials (reversal learning), the puppy's preferred entrance to the cylinder is obstructed by transparent Plexiglas and subjects are required to switch their response, detouring to the other opening of the apparatus in order to retrieve the treat. The side of the apparatus that subjects first approach, and the latency to obtain the reward are recorded as measures of response flexibility.

(5) Unsolvable task – The puppy is familiarized with displacing the lid from a transparent container to obtain a food reward inside. In test trials, the lid to the container is affixed, and the duration of subjects' attempts to retrieve the food, and gazing/help-seeking behavior toward the human experimenter, is recorded.

Gesture use – The experimenter shows the puppy a food reward, and behind an occluder, places the reward inside one of two possible hiding locations. The experimenter then provides one of the following cues before subjects can search:

> ***(6) Marker cue –*** The experimenter places an arbitrary marker (small yellow block) next to the hidden food.
>
> ***(7) Pointing –*** The experimenter points with the contralateral arm, index finger extended toward the hidden food.
>
> ***(8) Odor control –*** The experimenter remains still and does not provide any social information.

(9) Novel object – A stuffed mechanical cat is placed inside the pen with the puppy. The experimenter and handler leave the room and the motion-activated cat begins to locomote and vocalize. Later, the experimenter and handler re-enter the room. The subject's behavior is scored from video. The ethogram used for behavioral coding in this task is described in Table 1.

Table 1. Ethogram used for the novel object test.

Phase	Variable	Type	Measure	Description
Alone	Time to first approach	Latency	# of seconds	Amount of time to first approach the cat, defined as all four paws within 50 cm. Dogs that never approached were given the maximum score of 120s
	Proximity	Duration	# of seconds	Amount of time that the dog spends within 50 cm of the cat
	Number of approaches	Count	0 to 9	Number of times that the dog approaches the cat
	Contact	Duration	# of seconds	Amount of time that dog spends sniffing and/or with any part of its body in contact with the cat
	Time to contact	Latency	# of seconds	Amount of time to first contact the cat. Dogs that never made contact were given the maximum score of 120s
	Orient	Duration	# of seconds	Amount of time that dog spends with face oriented toward cat
	Grab cat	Binary	Yes/No	Dog bites/mouths cat
	Crouch	Binary	Yes/No	Crouch/ducking: downward movement of body/head to crouched position or walking backwards away from cat to sitting
	Play bow	Binary	Yes/No	Front legs out in front and chest low to the ground as if about to lie down, but rear end remains up
	Time to first vocalize	Latency	# of seconds	Amount of time until the dog first vocalizes (groan/growl, whine, yelp/bark, and/or howl). Dogs that never vocalized were given the maximum score of 120s
	Vocal intensity	Score	0 to 4	0 – dog never vocalizes 1 – dog reaches a groan/growl 2 – dog reaches a whine 3 – dog reaches a yelp/bark 4 – dog reaches a howl
	Cat vocal	Binary	Yes/No	Dog vocalizes while orienting toward the cat
	Door vocal	Binary	Yes/No	Dog vocalizes while orienting toward the door through which the experimenter left
Social phase 1 (experimenter re-enters the room)	Time to approach	Latency	# of seconds	Amount of time from when E is standing by cat to when dog approaches cat (using same approach definition as during alone period). If the dog is already at object or follows E to object immediately, the latency is 0.

	Approach score	Score	1 to 5	1 – dog already at object or follows E to object 2 – dog approaches with no food 3 – dog approaches with food, but no luring 4 – dog approaches once lured (E must leave cat to show dog food) 5 – dog never approaches the cat
Social phase 2 (handler re-enters the room)	Time to approach	Latency	# of seconds	Amount of time from when H releases dog to when dog approaches object (using same approach definition as during alone period)
	Approach path	Score	1 to 4	1 – dog approaches E on a straight path and over the cat 2 – dog approaches E on a straight path and then goes around the cat 3 – dog approaches E on a curved path, but doesn't leave the testing mat 4 – dog approaches E on a curved path, leaving the testing mat

(10) Working memory – This task is identical to hiding-finding warmups with the exception that the delay before the subject can search increases across trials (5s, 10s, 15s, 20s).

Perceptual discriminations – The subject must choose between two search locations based on a perceptual cue regarding which contains the reward:

(11) Visual discrimination – One plate contains five pieces of visible kibble and the other is empty. The plates are presented in front of the puppy and then pulled backward to 75cm in front of the puppy, equidistant to the left and right sides. The first plate that the puppy approaches is recorded.

(12) Auditory discrimination – Two metal bowls are used as the hiding locations. The experimenter sequentially places her hand into each container, audibly dropping the food into only one of the containers. The location of the subjects' first search is recorded.

(13) Odor discrimination – Two sections of rubber tubing ("elbows") are presented, one of which contains 10 pieces of dry kibble. The ends of the elbows are filled with cotton to prevent the contents from being visible or audible. The experimenter allows the subject to sniff each elbow individually for 3s, and then the elbows are presented side by side for an additional 3s before being pulled backward 75 cm from the puppy. Puppies are released and allow to move freely for 20s. On each trial, we score the first elbow that the subject approaches, as well as the cumulative time spent within a marked 25-cm radius around the elbows.

(14) Surprising events – The puppy is presented with a sequence of three potentially startling stimuli and the puppy's behavioral reactivity as well as subsequent recovery is scored from video using the ethogram described in Table 2:

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

b. Looming object – The experimenter holds a closed umbrella facing toward the subject. The experimenter then opens the umbrella such that it rapidly expands toward the subject.

c. Loud noise – The experimenter shakes a piece of sheet metal in front of the subject which displaces air and creates a loud oscillating sound with corresponding pulses of air.

After each event in the temperament test, the experimenter sets the stimulus on the ground and the puppy can explore on his/her own for 15 seconds. The experimenter then returns and vocally encourages the subject to approach and eat kibble near the previously startling stimulus.

Table 2. Ethogram for the surprising events test.

Phase	Variable	Type	Measure	Description
Sudden appearance	Initial reaction	Rating	1	No detectable reaction (attention orienting, such as turning head or perking ears, is fine)
			2	Flinch or startle without lowering of the body (some movement, including a small step back, is fine)
			3	Crouch or ducking (downward movement of body and/or head) without major displacement and maintaining general body orientation
			4	Rapid avoidance response away from the trash bag (can be paired with a crouch and/or change in general body orientation)
	Initial time to approach	Latency	seconds	Amount of time to first approach the trash bag, defined as all four paws within 50 cm
	Initial approach score	Rating	1	Dog initially approaches the trash bag within 15 seconds
			2	Dog initially approaches the trash bag within 16-30 seconds, after receiving verbal encouragement
			3	Dog initially approaches the trash bag after 30 seconds, after being led to it on leash
			4	Dog never approaches the trash bag over the 45-second trial, despite verbal and physical coaxing
	Initial approach path	Rating	1	Dog approaches the trash bag on a straight path and goes right up to or over the trash bag
			2	Dog approaches the trash bag on a curved path, but doesn't leave the testing mat
			3	Dog approaches the trash bag on a curved path, leaving the testing mat

Phase	Variable	Type	Measure	Description
			4	Dog does not approach the trash bag within the first 15 seconds
	Initial minimum distance	Rating	1	Dog on top of trash bag with all four paws within the first 45 seconds of the trash bag hitting the ground
			2	Dog on top of trash bag with two paws within the first 45 seconds of the trash bag hitting the ground
			3	Dog touching trash bag but not on top within the first 45 seconds of the trash bag hitting the ground
			4	Dog not within 50 cm of the trash bag within the first 45 seconds of the trash bag hitting the ground
			5	Dog never moves closer than the start line within the first 45 seconds of the trash bag hitting the ground
	Time to first contact	Latency	seconds	Amount of time to first contact or closely sniff the trash bag, at any point prior to the dog being brought back to the start line for re-approach.
	Initial contact	Duration	seconds	Amount of time that the dog spends closely sniffing and/or in contact with the trash bag
	Time to re-approach	Latency	seconds	Amount of time to re-approach the trash bag when called by E, defined as all four paws within 50 cm
	Re-approach path	Rating	1	Dog approaches E on a straight path and goes straight over the trash bag
			2	Dog approaches E on a straight path and then goes around the trash bag
			3	Dog approaches E on a curved path, but doesn't leave the testing mat
			4	Dog approaches E on a curved path, leaving the testing mat
	Re-approach minimum distance	Rating	1	Dog on top of trash bag with all four paws within the first 45 seconds of being released by H
			2	Dog on top of trash bag with two paws within the first 45 seconds of being released by H
			3	Dog touching trash bag but not on top within the first 45 seconds of being released by H
			4	Dog not within 50 cm of the trash bag within the first 45 seconds of being released by H
			5	Dog never moves closer than the start line within the first 45 seconds of being released by H
Looming object	Initial reaction	Rating	1	No detectable reaction (attention orienting, such as turning head or perking ears, is fine)
			2	Flinch or startle without lowering of the body (some movement, including a small step back, is fine)
			3	Crouch or ducking (downward movement of body and/or head) without major displacement and maintaining general body orientation
			4	Rapid avoidance response away from the umbrella (can be paired with a crouch and/or change in general body orientation)
	Initial time to approach	Latency	seconds	Amount of time to first approach the umbrella, defined as all four paws within 50 cm
	Initial approach score	Rating	1	Dog initially approaches the umbrella within 15 seconds
			2	Dog initially approaches the umbrella within 16-30 seconds, after receiving verbal encouragement
			3	Dog initially approaches the umbrella after 30 seconds, after being led to it on leash
			4	Dog never approaches the umbrella over the 45-second trial, despite verbal and physical coaxing
	Initial approach path	Rating	1	Dog approaches the umbrella on a straight path and goes right up to or over the umbrella
			2	Dog approaches the umbrella on a curved path, but doesn't leave the testing mat
			3	Dog approaches the umbrella on a curved path, leaving the testing mat
			4	Dog does not approach the umbrella within the first 15 seconds
		Rating	1	Dog on top of umbrella with all four paws within the first 45 seconds of the umbrella being placed on the ground

Phase	Variable	Type	Measure	Description
	Initial minimum distance		2	Dog on top of umbrella with two paws within the first 45 seconds of the umbrella being placed on the ground
			3	Dog touching umbrella but not on top within the first 45 seconds of the umbrella being placed on the ground
			4	Dog not within 50 cm of the umbrella within the first 45 seconds of the umbrella being placed on the ground
			5	Dog never moves closer than the start line within the first 45 seconds of the umbrella being placed on the ground
	Time to first contact	Latency	seconds	Amount of time to first contact or closely sniff the umbrella, at any point prior to the dog being brought back to the start line for re-approach.
	Initial contact	Duration	seconds	Amount of time that the dog spends closely sniffing and/or in contact with the umbrella
	Re-approach path	Rating	1	Dog approaches E on a straight path and goes straight over the umbrella
			2	Dog approaches E on a straight path and then goes around the umbrella
			3	Dog approaches E on a curved path, but doesn't leave the testing mat
			4	Dog approaches E on a curved path, leaving the testing mat
	Re-approach minimum distance	Rating	1	Dog on top of umbrella with all four paws within the first 45 seconds of being released by H
			2	Dog on top of umbrella with two paws within the first 45 seconds of being released by H
			3	Dog touching umbrella but not on top within the first 45 seconds of being released by H
			4	Dog not within 50 cm of the umbrella within the first 45 seconds of being released by H
			5	Dog never moves closer than the start line within the first 45 seconds of being released by H
Loud noise	Initial reaction	Rating	1	No detectable reaction (attention orienting, such as turning head or perking ears, is fine)
			2	Flinch or startle without lowering of the body (some movement, including a small step back, is fine)
			3	Crouch or ducking (downward movement of body and/or head) without major displacement and maintaining general body orientation
			4	Rapid avoidance response away from the metal sheet (can be paired with a crouch and/or change in general body orientation)
	Initial time to approach	Latency	seconds	Amount of time to first approach the metal sheet, defined as all four paws within 50 cm
	Initial approach score	Rating	1	Dog initially approaches the metal sheet within 15 seconds
			2	Dog initially approaches the metal sheet within 16-30 seconds, after receiving verbal encouragement
			3	Dog initially approaches the metal sheet after 30 seconds, after being led to it on leash
			4	Dog never approaches the metal sheet over the 45-second trial, despite verbal and physical coaxing
	Initial approach path	Rating	1	Dog approaches the metal sheet on a straight path and goes right up to or over the metal sheet
			2	Dog approaches the metal sheet on a curved path, but doesn't leave the testing mat
			3	Dog approaches the metal sheet on a curved path, leaving the testing mat
			4	Dog does not approach the metal sheet within the first 15 seconds
	Initial minimum distance	Rating	1	Dog on top of metal sheet with all four paws within the first 45 seconds of the metal sheet being placed on the ground
			2	Dog on top of metal sheet with two paws within the first 45 seconds of the metal sheet being placed on the ground
			3	Dog touching metal sheet but not on top within the first 45 seconds of the metal sheet being placed on the ground
			4	Dog not within 50 cm of the metal sheet within the first 45 seconds of the metal sheet being placed on the ground

Phase	Variable	Type	Measure	Description
			5	Dog never moves closer than the start line within the first 45 seconds of the metal sheet being placed on the ground
	Time to first contact	Latency	seconds	Amount of time to first contact or closely sniff the metal sheet, at any point prior to the dog being brought back to the start line for re-approach.
	Initial contact	Duration	seconds	Amount of time that the dog spends closely sniffing and/or in contact with the metal sheet
	Vocal intensity	Rating	0	Dog never vocalizes
			1	Dog reaches a groan/growl
			2	Dog reaches a whine
			3	Dog reaches a yelp/bark
			4	Dog reaches a howl
	Time to re-approach	Latency	seconds	Amount of time to re-approach the metal sheet when called by E, defined as all four paws within 50 cm
	Re-approach path	Rating	1	Dog approaches E on a straight path and goes straight over the metal sheet
			2	Dog approaches E on a straight path and then goes around the metal sheet
			3	Dog approaches E on a curved path, but doesn't leave the testing mat
			4	Dog approaches E on a curved path, leaving the testing mat
	Re-approach minimum distance	Rating	1	Dog on top of metal sheet with all four paws within the first 45 seconds of being released by H
			2	Dog on top of metal sheet with two paws within the first 45 seconds of being released by H
			3	Dog touching metal sheet but not on top within the first 45 seconds of being released by H
			4	Dog not within 50 cm of the metal sheet within the first 45 seconds of being released by H
			5	Dog never moves closer than the start line within the first 45 seconds of being released by H
Alone	Orient	Duration	seconds	Amount of time over the last minute that dog spends with face oriented toward the experimenters, who are standing outside of the testing pen
	Time to vocalize	Latency	seconds	Amount of time until the dog first vocalizes (groan/growl, whine, yelp/bark, and/or howl). Dogs that never vocalized were given the maximum score of 60s
	Vocal intensity	Rating	0	Dog never vocalizes
			1	Dog reaches a groan/growl
			2	Dog reaches a whine
			3	Dog reaches a yelp/bark
			4	Dog reaches a howl

Dog Cognitive Development Battery (adult measures)

Following completion of puppy testing, we spent 5 weeks piloting with 8 adult dogs to adapt some components of the DCDB to older dogs (i.e., 1+ years old), ensuring that all tasks yielded meaningful scores in this age range and population. Procedural modifications from the puppy version of these tasks are described below, and detailed experimental methods for the additional tasks are provided under *Protocols*:

1) Removal of vision pre-test

a. Because the adult dogs had either been selected for breeding (e.g., dams and sires) or to enter professional training (e.g., original study puppies), their eyesight has been thoroughly assessed by a veterinarian and deemed adequate. We therefore removed the vision pre-test for adult participants.

2) Addition of a physical problem-solving task linked to success in guide dogs

a. We added an independent problem-solving task that has recently been shown to predict success in a population of guide dogs (Bray et al., 2017). In this task, the dog is required to watch and remember where a treat is hidden within several possible locations on an apparatus, and then complete two motor tasks in a precise order to successfully retrieve the food. The dependent measures for this task are shown in Table 3.

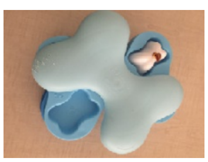
3) Minor age-appropriate modifications

a. We increased the difficulty of the working memory task. Adult subjects are required to remember where a treat is hidden while accounting for more possible hiding locations (four vs. two) as well as longer waiting periods (up to 40 s).

b. Where needed, larger stimuli are used (e.g., the ball during the retrieval task, the platform during the laterality task, the container during the unsolvable task, an extra mechanical cat during the novel object task, and the umbrella during the surprising events task).

c. With the puppies, the battery consisted of three sessions over three days. Due to the increased attention span and food motivation of adult dogs, the adult version of the DCDB can be implemented in two sessions lasting around 1 to 1.5 hours each, either on the same day with a break in between or across two different days (Figure 1B).

Table 3. Ethogram for independent problem-solving task.

Phase	Variable	Type	Measure	Description
Problem solving A 	Avg time 1	Duration	seconds	Amount of time from when the handler drops the leash to when the dog solves the first familiarization trial (by eating one visible treat out of the apparatus, capped at 45 seconds)
	Avg time 2	Duration	seconds	Amount of time from when the handler drops the leash to when the dog solves the second familiarization trial (by dislodging the bone and consuming the treat underneath, capped at 45 seconds), averaged over the number of these trials in which the dog participates
	Correct attempts	Count	0 to 4	Number of baited bones (i.e., bones with a treat underneath) that the dog successfully dislodges within the 120-second trial limit
	Incorrect attempts	Count	0 to 5	Number of non-baited bones (i.e., bones without a treat underneath) that the dog dislodges within the 120-second trial limit
	Latency to solve	Latency	seconds	The amount of time it takes the dog to complete the test trial (dislodging four baited bones and consuming the treats underneath within the 120 second trial limit), from the moment the leash drops to the moment the dog eats the last treat
Problem solving B 	Avg time 1	Duration	seconds	Amount of time from when the handler drops the leash to when the dog solves the first familiarization trial (by eating one visible treat out of the apparatus, capped at 45 seconds)
	Avg time 2	Duration	seconds	Amount of time from when the handler drops the leash to when the dog solves the second familiarization trial (by twisting the top of the apparatus and consuming the hidden treat, capped at 45 seconds), averaged over the number of these trials in which the dog participates
	Avg time 3	Duration	seconds	Amount of time from when the handler drops the leash to when the dog solves the third familiarization trial (by dislodging the bone and consuming the treat underneath, capped at 45 seconds), averaged over the number of these trials in which the dog participates
	Avg latency to solve	Latency	seconds	Amount of time from when the handler drops the leash to when the dog solves the test trial (by dislodging the bone, twisting the top of the apparatus, and consuming the treat underneath, capped at 120 seconds), averaged over the number of test trials in which the dog participates
	Successful trials	Count	0 to 3	The number of test trials that the dog successfully solves (by dislodging the bone, twisting the top of the apparatus, and consuming the treat underneath within the 120-second trial limit). Dogs only proceed to the next test trial if they are successful in the preceding test trial

Coding and Reliability

All tasks were video recorded, and most behavioral variables were scored live. The following tasks were later coded from video: novel object and the surprising events task, as well as select variables from cylinder (latency during inhibitory control and reversal trials and first side correct during reversal trials), unsolvable (average time manipulating object) and odor discrimination (time at right and left elbow, from which the variables persistence, time at incorrect elbow, and time at correct elbow were subsequently calculated).

For the live-coded data, independent coders then scored 20% of randomly selected trials, and interrater reliability was calculated using Pearson correlation for continuous variables and Cohen's Kappa for categorical variables. 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 20% of trials for reliability analyses.

Reliability was excellent for all live-coded measures and inter-rater agreement is reported in Table 4 (kappa: mean = 0.97, range 0.95-0.98; correlation: mean = 0.97, range 0.91-1.00). Reliability was also strong for the video-coded measures; inter-rater agreement for the video-coded measures of the cognitive tasks is reported in Table 5 (correlation: mean = 0.98, range = 0.97-0.98), inter-rater agreement for the novel object task is reported in Table 6 (kappa: mean = 0.90, range 0.80-0.99; correlation: mean = 0.96, range 0.80-1.00) and inter-rater agreement for the surprising events task including sudden appearance, looming object, and loud noise is reported in Table 7 (kappa: mean = 0.80, range 0.51-1.00; correlation: mean = 0.87, range 0.69-1.00)[1].

Table 4. Reliability statistics for live coded measures in the DCDB.

variable	*N*	*statistic*	*value*
Retrieval: average score	33	Correlation	0.99
Retrieval: tally count	33	Correlation	1.00
Laterality	34	Kappa	0.97
Human interest: avg look time	27	Correlation	0.91
Cylinder: inhibitory control score			*Reliability pending*
Cylinder: reversal			*Reliability pending*
Unsolvable task: avg time looking at human			*Reliability pending*
Communicative marker	34	Kappa	0.98
Arm pointing	33	Kappa	0.98
Odor control trials	34	Kappa	0.96
Memory	34	Kappa	0.97
Visual discrimination	34	Kappa	0.96
Auditory discrimination	34	Kappa	0.95
Odor discrimination: first choice	34	Kappa	0.90
Odor discrimination: final choice	34	Kappa	0.89

Table 5. Reliability statistics for video-coded measures in the DCDB.

variable	*N*	*statistic*	*value*
Cylinder: latency (inhibitory control trials)	33	Correlation	0.98
Cylinder: latency (reversal trials)	33	Correlation	0.99
Cylinder: reversal (first side correct)	34	Kappa	*Reliability pending*
Unsolvable task: avg time manipulating object			*Reliability pending*
Odor discrimination: time at right elbow	34	Correlation	0.98
Odor discrimination: time at left elbow	34	Correlation	0.98

[1] Reliability coding for adult data is ongoing and is not included in this report.

Table 6. Reliability statistics for measures from the novel object test.

variable	***N***	***statistic***	***value***
Alone: time to first approach	34	Correlation	1.00
Alone: proximity	34	Correlation	1.00
Alone: no of approaches	34	Kappa	0.99
Alone: contact	34	Correlation	1.00
Alone: time to first contact	34	Correlation	0.93
Alone: orient	34	Correlation	0.80
Alone: time to first vocalize	34	Correlation	0.93
Alone: vocal intensity	34	Kappa	0.92
Alone: grab cat	34	Kappa	0.82
Alone: crouch	34	Kappa	0.97
Alone: play bow	34	Kappa	0.92
Alone: cat vocal	34	Kappa	0.93
Alone: door vocal	34	Kappa	0.80
Social 1: time to approach	34	Correlation	1.00
Social 1: approach score	34	Kappa	0.80
Social 2: time to approach	34	Correlation	1.00
Social 2: approach path	34	Kappa	0.96

Table 7. Reliability statistics for measures from the surprising events test.

variable	*n*	*statistic*	*value*
Sudden appearance: initial reaction	34	Correlation	0.75
Sudden appearance: time to first approach	34	Correlation	0.83
Sudden appearance: initial approach score	34	Kappa	0.71
Sudden appearance: initial approach path	34	Kappa	0.60
Sudden appearance: initial minimum distance	34	Kappa	0.80
Sudden appearance: time to first contact	34	Correlation	0.93
Sudden appearance: initial contact	34	Correlation	0.89
Sudden appearance: time to re-approach	34	Correlation	0.90
Sudden appearance: re-approach path	34	Kappa	0.92
Sudden appearance: re-approach minimum distance	34	Kappa	0.88
Looming object: initial reaction	34	Correlation	0.94
Looming object: time to first approach	34	Correlation	0.96
Looming object: initial approach score	34	Kappa	0.91
Looming object: initial approach path	34	Kappa	0.97
Looming object: initial minimum distance	34	Kappa	0.80
Looming object: time to first contact	34	Correlation	1.00
Looming object: initial contact	34	Correlation	0.77
Looming object: re-approach path	34	Kappa	0.87
Looming object: re-approach minimum distance	34	Kappa	1.00
Loud noise: initial reaction	34	Correlation	0.72
Loud noise: time to first approach	34	Correlation	0.89
Loud noise: initial approach score	34	Kappa	0.79
Loud noise: initial approach path	34	Kappa	0.51
Loud noise: initial minimum distance	34	Kappa	0.80
Loud noise: time to first contact	34	Correlation	0.90
Loud noise: initial contact	34	Correlation	0.94
Loud noise: vocal intensity	34	Kappa	0.75
Loud noise: time to re-approach	34	Correlation	1.00

variable	n	statistic	value
Loud noise: re-approach path	34	Kappa	0.94
Loud noise: re-approach minimum distance	34	Kappa	0.61
Alone: orient to experimenters	34	Correlation	0.69
Alone: time to first vocalize	34	Correlation	0.87
Alone: vocal intensity	34	Kappa	0.78

ANALYSES & RESULTS

Descriptive Statistics

9-week old puppies

Below we report the sample size, mean, standard deviation, range, and standard error for all measures in the DCDB collected from dogs at 9 weeks of age. In order to determine whether there was a significant effect of sex and/or age at testing, we used linear mixed-effects models fit by REML with age at testing and sex as predictors of each measure, including litter as a random effect. Age and sex effects are only reported in text when these covariates were statistically significant.

Pre-test and warm-ups.

Vision pre-test. Across the six trials, puppies (n = 168) visibly moved their head to follow the moving cotton ball on 96.83% (± 7.94) of trials, indicating that their eyesight was sufficiently developed to view events occurring at least 1m in front of them, as required in the primary DCDB tasks.

Hiding-finding warm-ups. On Day 1 of the battery, 158 puppies successfully completed warm-ups. Eighty-six of them (54%) required more than the four mandatory two-cup visible displacement trials; within those puppies there was a wide range of trials needed to meet the criterion (range: 5-20 trials). The mean number of trials needed to complete familiarization was 6.83 (± 3.87) trials. There was an effect of age on number of Day 1 familiarization trials completed ($\beta = 0.24$, $t = 3.10$, $p < 0.01$); younger puppies were faster to reach criterion.

On Day 2 of the battery, 164 puppies successfully completed warm-ups. Seventy-one of them (43%) required more than the four mandatory two-cup visible displacement trials, and within those puppies there was a wide range of trials needed (range: 5-22 trials). The mean number of trials needed to complete familiarization was 6.22 (± 3.69) trials.

On Day 3 of the battery, 165 puppies successfully completed warm-ups. Seventy-six of them (46%) required more than the four mandatory two-cup visible displacement trials, and within those puppies there was a wide range of trials needed (range: 5-20 trials). The mean number of trials needed to complete familiarization was 5.97 (± 3.26) trials.

Primary DCDB measures

Of the 168 puppies that participated, 160 (95%) successfully completed every task in the battery (except for the human interest task, which was added shortly after data collection began and included for only 150 puppies). Results from each of the cognitive tasks are briefly described below with full descriptive statistics reported in Tables 8-12.

Table 8. Descriptive statistics for DCDB measures with 9-week old puppies.

variable	N	mean	SD	minimum	maximum	SEM
retrieval: average score	168	3.31	1.17	1	5	0.09
retrieval: tally	168	3.07	3.75	0	14	0.29
laterality: bias strength	168	40.83	26.69	0	93.33	2.06
laterality: laterality index	168	-7.18	48.35	-93.33	93.33	3.73
human interest: avg look time	150	6.26	4.01	0	20.91	0.33
cylinder: familiarization score	166	7.78	3.16	4	18	0.24
cylinder: inhibitory control score	166	51.2	24.24	0	100	1.88
cylinder: latency (inhibitory control trials)	166	3.93	2.41	1.44	17.02	0.19
cylinder: reversal	165	29.47	23.42	0	87.5	1.82
cylinder: latency (reversal trials)	165	6.58	4.51	2.22	30	0.35
cylinder: reversal (first side correct)	165	23.26	27.29	0	100	2.12
unsolvable task: avg time looking at human	168	1	1.03	0	4.42	0.08
unsolvable task: avg time manipulating object	168	12.71	3.27	3.25	23.61	0.25
arm pointing	164	69.41	18.88	16.67	100	1.47
communicative marker	166	76.21	17.75	25	100	1.38
odor control trials	163	49.54	15.65	12.5	87.5	1.23
memory (5s)	165	74.34	20.03	16.67	100	1.56
memory (10s)	163	70.24	22.05	0	100	1.73
memory (15s)	102	65.03	18.54	16.67	100	1.84

variable	*N*	*mean*	*SD*	*minimum*	*maximum*	*SEM*
memory (20s)	60	63.89	19.45	16.67	100	2.51
visual discrimination	168	91	12.75	50	100	0.98
auditory discrimination	167	58.91	19.79	12.5	100	1.53
odor discrimination: first choice	164	53.76	20.83	16.67	100	1.63
odor discrimination: final choice	164	72.56	19.78	0	100	1.54
odor discrimination: persistence	164	79.83	22.37	20.94	125.46	1.75
odor discrimination: time at correct response	164	61.56	22.32	13.9	116.43	1.74
odor discrimination: time at incorrect response	164	18.27	10.14	1.44	66.53	0.79

Retrieval. Across two one-minute trials, dogs averaged a score of 3.31 (± 1.17), reflecting a group-wide tendency to pursue and pick up the ball. Dogs brought the ball back to the experimenter 3.07 (± 3.75) times per trial on average.

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

$$Laterality\ Index\ (LI) = \frac{(R-L)/(R+L)}{100}$$

Where in the formula above: R = number of right paw uses, L = 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. As a group, puppies displayed a mean LI of -7.18 ± 48.35 and a mean bias strength of 40.83 ± 26.69.

Human interest. Over three 30-second trials, dogs spent an average of 6.26 seconds (± 4.01) per trial looking to the face of the human experimenter.

Cylinder. One hundred fifty-nine of the 166 dogs who completed the cylinder task (96%) required more than the four mandatory familiarization trials, and within those dogs there was a wide range of familiarization trials needed (range: 5-18 trials). The mean number of trials needed to complete familiarization was 7.78 trials (± 3.16). There was no effect of number of familiarization trials required on dogs' subsequent performance on test trials (β = -0.01, t = -1.09, p = 0.28). However, there was an effect of number of familiarization trials on dog's reversal learning performance (β =

-0.01, t = -2.85, $p < 0.01$). Therefore, the lower the number of familiarization trials required, the more reversal learning trials that the puppies later succeeded on.

In the inhibitory control trials, 9-week-old dogs made correct choices on 51.20 ± 24.24% of test trials. As we would expect if the inhibitory control manipulation affected task difficulty, the percentage of correct responses during the first 4 familiarization trials was significantly greater than the percentage of correct responses during the first 4 inhibitory control trials (mean first 4 familiarization trials: 49.10 ± 24.60 %; mean first 4 inhibitory control trials: 34.79 ± 27.08 %; Paired t-test: $t_{165} = -5.19$, $p < 0.001$). 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. However, 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, puppies made correct choices on 29.47 ± 23.42% of trials. On average, dogs required 6.58 seconds (± 4.51) to retrieve the treat on each trial. We also tracked the pathway of the participant 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 plexiglass barrier, we considered it to be a "first side correct" trial. Puppies averaged 23.26 ± 27.29% first side correct trials 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 4 inhibitory control trials was significantly greater than the percentage of correct responses during the first 4 reversal trials (mean first 4 reversal trials: 13.18 ± 19.82%; Paired t-test: $t_{164} = 8.67$, $p < 0.001$). Latency to solve the reversal trials 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 stated above, it is not clear whether this effect reflects cognitive or locomotor differences within the sample.

Table 9. Age and sex effects on DCDB measures with 9-week old puppies.

variable	***β (age)***	***t (age)***	***p (age)***	***β (sex)***	***t (sex)***	***p (sex)***
retrieval: average score	-0.00	-0.10	0.92	0.07	0.40	0.69
retrieval: tally	0.04	0.45	0.66	0.51	0.89	0.38
laterality: bias strength	0.01	0.02	0.99	-2.21	-0.56	0.57
laterality: laterality index	0.85	0.83	0.41	-1.06	-0.14	0.89
human interest: avg look time	0.11	1.06	0.29	-0.09	-0.15	0.88

variable	***β (age)***	***t (age)***	***p (age)***	***β (sex)***	***t (sex)***	***p (sex)***
cylinder: familiarization score	-0.04	-0.54	0.59	0.72	1.46	0.15
cylinder: inhibitory control score	0.65	1.16	0.25	-2.46	-0.67	0.50
cylinder: latency (inhibitory control trials)	-0.16	-3.19	0.00	0.11	0.31	0.76
cylinder: reversal	0.08	0.15	0.88	-0.35	-0.10	0.92
cylinder: latency (reversal trials)	-0.30	-2.87	0.01	-0.72	-1.09	0.28
cylinder: reversal (first side correct)	-0.35	-0.61	0.55	-1.71	-0.40	0.69
unsolvable task: avg time looking at human	0.02	0.78	0.44	-0.23	-1.43	0.16
unsolvable task: avg time manipulating object	0.03	0.34	0.73	-0.52	-1.06	0.29
arm pointing	0.05	0.12	0.91	-0.45	-0.16	0.87
communicative marker	0.45	1.22	0.23	-3.23	-1.16	0.25
odor control trials	-0.39	-1.19	0.24	3.85	1.56	0.12
memory (5s)	-0.27	-0.53	0.60	1.31	0.45	0.65
memory (10s)	-0.38	-0.82	0.41	-1.64	-0.47	0.64
memory (15s)	0.73	1.55	0.12	-3.31	-0.89	0.37
memory (20s)	0.22	0.35	0.73	-1.92	-0.37	0.71
visual discrimination	0.15	0.48	0.63	-1.48	-0.80	0.42
auditory discrimination	-0.21	-0.51	0.61	1.31	0.42	0.67
odor discrimination: first choice	-0.02	-0.05	0.96	0.27	0.09	0.93
odor discrimination: final choice	-0.40	-0.89	0.38	-0.40	-0.13	0.89
odor discrimination: persistence	0.07	0.11	0.91	-4.36	-1.37	0.17
odor discrimination: time at correct response	0.01	0.02	0.98	-3.81	-1.15	0.25
odor discrimination: time at incorrect response	0.06	0.23	0.82	-0.75	-0.52	0.61

Unsolvable task. Over four 30-second trials, puppies looked to the experimenter's face for an average of 1.00 second (± 1.03) per trial. They manipulated the locked container for an average of 12.71 seconds (± 3.27) per trial.

Communicative marker. Over 12 trials, puppies achieved a mean score of 76.21% correct (± 17.75).

Arm pointing. Over 12 trials, puppies achieved a mean score of 69.41% correct (± 18.88).

Memory. Over the six five-second delay trials, puppies achieved a mean score of 74.34% correct (± 20.03). Over the six ten-second delay trials, puppies achieved a mean score of 70.24 % correct (± 22.05). Puppies only participated in 15-second delay trials if they chose correctly on at least 4 out of 6 10-second trials. Out of the 163 puppies who completed ten-second delay trials, 102 puppies (63%) chose correctly on over half of the trials and therefore moved on to 15-second delay. Over the six 15-second delay trials, puppies achieved a mean score of 65.03% correct (± 18.54). Puppies only participated in 20-second delay trials if they chose correctly on at least 4 out of 6 15-second trials. Out of the 102 puppies who completed 15-second delay trials, 60 puppies (59%) chose correctly on over half of the trials and therefore moved on to twenty-second delay. Over the six 20-second delay trials, those puppies achieved a mean score of 63.89% correct (± 19.45). See Table 10 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).

Table 10. Performance across delays in the memory task with 9-week old puppies who completed all four delays.

variable	***N***	***mean***	***SD***	***minimum***	***maximum***	***SEM***
memory (5s)	60	83.05	14.22	50	100	1.84
memory (10s)	60	85	12.54	66.67	100	1.62
memory (15s)	60	76.94	11.92	66.67	100	1.54
memory (20s)	60	63.89	19.45	16.67	100	2.51

Visual discrimination. Over 8 trials, puppies achieved a mean score of 91.00% correct (± 12.75).

Auditory discrimination. Over 8 trials, puppies achieved a mean score of 58.91% correct (± 19.79).

Odor discrimination. Over 6 trials, puppies achieved a mean score of 53.76% (± 20.83) correct first choices and 72.56% (± 19.78) correct final choices. Overall, they spent 61.56 seconds (± 22.32) at the correct location and 18.27 seconds (± 10.14) at the incorrect location.

Novel object. All raw variables from this task were coded from video. We were particularly interested in the initial behavioral responses and vocalizations displayed toward the mechanical cat when the puppy was alone in the room with the novel object, 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 Table 1, and the descriptive statistics for those variables are reported in Table 11. To summarize behavioral responses in the novel object task, we conducted a principal components analysis on measures scored by ethogram. Parallel analysis (Horn, 1965) suggested extraction of a 3-component solution, which was extracted using a direct oblimin rotation (Osborne & Costello, 2009). The first component was loaded highly by measures of approach, proximity, orienting, and physical interaction with the novel object and was 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, and was interpreted as 'vocal intensity'. The third component was loaded positively by latency to approach and contact the object, time orienting at the object, and negatively by the number of approaches to the object. This component was interpreted as 'caution'.

Table 11. Descriptive statistics for raw measures in the novel object task with 9-week old puppies.

variable	***N***	***mean***	***SD***	***minimum***	***maximum***	***SEM***
alone: time to first approach	168	35.23	49.75	0	120	3.84
alone: proximity	168	37.8	36.22	0	118	2.79
alone: no of approaches	168	1.73	1.66	0	9	0.13
alone: contact	168	29.26	30.34	0	114	2.34
alone: time to first contact	168	26.36	45.4	0	120	3.50
alone: orient	168	79.81	25.28	20	120	1.95
alone: grab cat	168	0.45	0.5	0	1	0.04
alone: crouch	168	0.22	0.42	0	1	0.03
alone: play bow	168	0.32	0.47	0	1	0.04
alone: time to first vocalize	168	58.68	34.29	3	120	2.64
alone: vocal intensity	168	2.43	1.12	0	4	0.09
alone: cat vocal	168	0.78	0.42	0	1	0.03
alone: door vocal	168	0.65	0.48	0	1	0.04
social 1: time to approach	168	2.22	6.8	0	60.38	0.52

variable	*N*	*mean*	*SD*	*minimum*	*maximum*	*SEM*
social 1: approach score	168	1.38	0.58	1	4	0.04
social 2: time to approach	168	3.32	6.59	0.62	70	0.51
social 2: approach path	168	1.7	1.03	1	4	0.08

Surprising Events: sudden appearance, looming object, and loud noise. All raw variables from this task were coded from video. Again, we were interested in the immediate behavioral response of the puppy to each startling occurrence, as well as metrics of how well the puppy recovered on its own and with the encouragement of a human. The ethogram for the coded variables is presented in Table 2, and the descriptive statistics for those variables are reported in Table 12. As with the novel object test, we used principal components analysis to develop a set of component measures describing behavioral variation in this test. Parallel analysis suggested a five-factor solution which was 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. This component was interpreted 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), vocal response to this event, and the latency to, and intensity of vocalizations when left alone for one minute at the end of the task. This component was interpreted as reflecting "reactivity to sound and isolation" (following the stressor). The third component was loaded by reaction 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, and was interpreted as "reactivity and recovery to the suddenly appearing object". The fourth component was loaded by measures involving proximity to the suddenly appearing object and was 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), and was interpreted as "recovery from startling sound".

Table 12. Descriptive statistics for raw measures in the surprising events task with 9-week old puppies.

variable	***n***	***mean***	***sd***	***minimum***	***maximum***	***SEM***
sudden appearance: initial reaction	168	1.33	0.69	1	4	0.05
sudden appearance: time to first approach	168	9.18	9.63	0	52	0.74
sudden appearance: initial approach score	168	1.25	0.6	1	4	0.05
sudden appearance: initial approach path	168	1.82	1.15	1	4	0.09
sudden appearance: initial minimum distance	168	2.12	0.78	1	5	0.06
sudden appearance: time to first contact	168	8.04	10.66	1	61	0.82
sudden appearance: initial contact	168	7.52	4.89	0	15	0.38
sudden appearance: time to re-approach	168	1.58	1.07	0	10	0.08
sudden appearance: re-approach path	168	1.79	0.96	1	4	0.07
sudden appearance: re-approach minimum distance	168	1.93	0.74	1	4	0.06
looming object: initial reaction	168	3.09	0.53	1	4	0.04
looming object: time to first approach	168	3.87	7.38	0	45	0.57
looming object: initial approach score	168	1.1	0.43	1	4	0.03
looming object: initial approach path	168	1.54	0.89	1	4	0.07
looming object: initial minimum distance	168	1.4	0.66	1	4	0.05
looming object: time to first contact	168	4.32	7.75	0	51	0.60
looming object: initial contact	168	8.93	5.03	0	15	0.39
looming object: re-approach path	168	2.18	1.32	1	4	0.10
looming object: re-approach minimum distance	168	1.37	0.66	1	3	0.05
loud noise: initial reaction	168	3.35	1.07	1	4	0.08
loud noise: time to first approach	168	8.92	10.87	0	48	0.84
loud noise: initial approach score	168	1.36	0.6	1	4	0.05
loud noise: initial approach path	168	2.54	1.15	1	4	0.09
loud noise: initial minimum distance	168	1.25	0.58	1	4	0.04
loud noise: time to first contact	168	12.8	13.21	0	63	1.02

variable	n	mean	sd	minimum	maximum	SEM
loud noise: initial contact	168	3.19	4.01	0	16	0.31
loud noise: vocal intensity	168	0.64	1.04	0	3	0.08
loud noise: time to re-approach	165	3.26	6.73	0	55	0.52
loud noise: re-approach path	165	2.19	1.35	1	4	0.10
loud noise: re-approach minimum distance	165	1.34	0.71	1	4	0.06
alone: orient to experimenters	167	31.75	14.51	0	60	1.12
alone: time to first vocalize	168	31.84	21.59	2	60	1.66
alone: vocal intensity	168	2.12	1.5	0	4	0.12

Young Adults

Below we report the sample size, mean, standard deviation, range, and standard error for all measures in the DCDB collected from dogs in young adulthood. In order to determine whether there was a significant effect of sex and/or age at testing, we used linear mixed-effects models fit by REML with age at testing and sex as predictors of each measure, including litter as a random effect. Age and sex effects are only reported in text when these covariates were statistically significant, but all model results are shown in Table 13.

Table 13. Age and sex effects on adult DCDB measures.

variable	β (age)	t (age)	p (age)	β (sex)	t (sex)	p (sex)
retrieval: average score	-0.58	-0.71	0.48	-0.18	-0.99	0.32
retrieval: tally	-5.24	-1.27	0.21	-0.28	-0.30	0.77
laterality: bias strength	6.51	0.33	0.74	-4.32	-0.85	0.40
laterality: laterality index	3.12	0.08	0.94	-21.15	-2.00	0.05
human interest: avg look time	-10.69	-2.36	0.02	-1.97	-1.73	0.09
human interest: avg interaction time	1.99	0.40	0.69	0.77	0.62	0.54
cylinder: familiarization score	2.20	2.03	0.04	0.09	0.34	0.73
cylinder: inhibitory control score	20.91	1.60	0.11	-5.18	-1.66	0.10
cylinder: reversal	8.10	0.50	0.61	-12.66	-3.32	0.00

variable	*β (age)*	*t (age)*	*p (age)*	*β (sex)*	*t (sex)*	*p (sex)*
communicative marker	-4.32	-0.50	0.62	-0.46	-0.21	0.84
arm pointing	26.40	2.28	0.02	-0.59	-0.20	0.84
unsolvable task: avg time looking at human	2.75	1.20	0.23	0.03	0.05	0.96
memory (20s)	-14.90	-1.04	0.30	-2.26	-0.64	0.52
memory (40s)	4.78	0.32	0.75	-3.46	-0.94	0.35
problem solving A: avg time 1	9.31	1.58	0.12	0.55	0.36	0.72
problem solving A: avg time 2	1.49	0.21	0.83	-1.58	-1.02	0.31
problem solving A: correct attempts	-1.41	-1.85	0.07	-0.19	-1.02	0.31
problem solving A: incorrect attempts	0.34	0.37	0.71	0.09	0.42	0.67
problem solving A: latency to solve	35.85	1.74	0.08	2.47	0.51	0.61
problem solving B: avg time 1	4.75	1.66	0.10	0.53	0.73	0.47
problem solving B: avg time 2	3.26	0.51	0.61	-0.59	-0.38	0.71
problem solving B: avg time 3	8.71	1.16	0.25	-0.64	-0.33	0.74
problem solving B: avg latency to solve	2.26	0.10	0.92	-4.98	-0.88	0.38
problem solving B: successful trials	0.11	0.15	0.88	0.08	0.40	0.69
visual discrimination	5.41	0.65	0.52	-2.15	-1.03	0.30
auditory discrimination	-29.73	-2.47	0.01	-1.98	-0.69	0.49
odor discrimination: first choice	18.19	1.35	0.18	2.29	0.67	0.51
odor discrimination: final choice	-8.63	-0.64	0.53	0.57	0.17	0.87

Pre-test and warm-ups.

Hiding-finding warm-ups. In session 1, all 160 dogs successfully completed warm-ups. Thirty-six of them (23%) required more than the four mandatory two-cup visible displacement trials, and within those dogs there was a wide range of trials needed (range: 5-20 trials). The mean number of trials required to complete familiarization was 4.78 (± 1.96) trials.

In session 2, all 160 dogs again successfully completed warm-ups. Forty-five of them (28%) required more than the four mandatory two-cup visible displacement trials, and within those dogs there was a wide range of trials needed (range: 5-13 trials). The mean number of trials needed to complete familiarization was 4.73 (± 1.42) trials.

<u>*Primary DCDB measures.*</u>

Of the 160 dogs that participated, 156 (98%) successfully completed every task in the battery. Results of each of the cognitive tasks are briefly described below with key descriptive statistics presented in Table 14. Results from the two temperament tasks (novel object and surprising events), from which raw measures are currently being coded from video, are not included in this report.

Retrieval. Across two one-minute trials, dogs averaged a score of 3.70 (± 1.28), reflecting a group-wide tendency to approach and retrieve the ball. Dogs brought the ball back to the experimenter 6.17 (± 6.42) times per trial on average.

Laterality. As with the puppies, we calculated a laterality index (LI) and bias strength for each adult dog. As a group, dogs displayed a mean LI of -10.7 ± 66.32 and a mean bias strength of 59.06 ± 31.67. There was an effect of sex on direction of laterality (estimate = -21.15, $t = -2.00$, $p < 0.05$): males exhibited a more left lateral biased response, replicating previous findings regarding the sex difference in laterality (McGreevy, Brueckner, Thomson, & Branson, 2010; Wells, 2003).

Table 14. Descriptive statistics for young adult performance in the DCDB.

variable	***mean***	***SD***	***min***	***max***	***SEM***
retrieval: average score	3.7	1.28	1	5	0.10
retrieval: tally	6.17	6.42	0	24	0.51
laterality: bias strength	59.06	31.67	0	100	2.50
laterality: laterality index	-10.7	66.32	-100	100	5.24
human interest: avg look time	15.55	7.31	1.95	30.81	0.58
human interest: avg interaction time	19.32	7.82	0	28.23	0.62
cylinder: familiarization score	5.06	1.72	4	13	0.14
cylinder: inhibitory control score	75.94	20.53	0	100	1.62
cylinder: reversal	59.59	25.67	0	100	2.04
communicative marker	89.32	13.81	33.33	100	1.09

arm pointing	77.14	18.65	25	100	1.48
odor control trials	50.7	15.88	12.5	87.5	1.25
unsolvable task: avg time looking at human	3.3	3.62	0	16.84	0.29
memory (20s)	73.13	22.31	0	100	1.78
memory (40s)	56.76	23.16	0	100	1.85
visual discrimination	90.08	13.16	37.5	100	1.04
auditory discrimination	65.47	18.99	25	100	1.50
odor discrimination: first choice	60.83	21.49	16.67	100	1.70
odor discrimination: final choice	71.67	21.38	0	100	1.69

Human interest. Over three 30-second trials, dogs spent an average of 15.55 seconds (± 7.31) per trial looking to the face of the human experimenter. During the 30-second play breaks, they spent an average of 19.32 seconds (± 7.82) interacting with the human experimenter. Together, the results indicate that most dogs preferred to spend much of the task engaging with an interactive human. Looking time decreased significantly as a function of age (β = -10.69, t = -2.36, p = 0.02).

Cylinder. In familiarization trials, 71 of the 160 adults required more than the four mandatory familiarization trials, and within those dogs there was a wide range of familiarization trials needed (range: 5-13 trials). The mean number of trials required to complete familiarization was 5.06 trials (± 1.72). There was a positive association between age (β = 2.20, t = 2.03, p = 0.04) and the number of familiarization trials required. In the inhibitory control trials, on average dogs made correct choices on 75.94 ± 20.53% of test trials. As we would expect if the inhibitory control manipulation affected dogs' response strategies, the percentage of correct responses during the first 4 familiarization trials was significantly greater than the percentage of correct responses during the first 4 inhibitory control trials (mean first 4 familiarization trials: 83.33 ± 22.97 %; mean first 4 inhibitory control trials: 63.89 ± 27.30 %; Paired t-test: t_{158} = 7.05, p <0.001). In the reversal trials, dogs made correct choices on 59.59 ± 25.67% of trials. As we would expect if blocking their preferred pathway affected dogs' response strategies, the percentage of correct responses during the first 4 inhibitory control trials was significantly greater than the percentage of correct responses during the first 4 reversal trials (mean first 4 reversal trials: 39.94 ± 29.65%; Paired t-test: t_{158} =

8.24, $p < 0.001$). Performance on reversal trials was affected by sex ($\beta = -12.66$, $t = -3.32$, $p = 0.00$), with females scoring higher on reversal trials.

Communicative marker. Over 12 trials, dogs achieved a mean score of 89.32% correct (± 13.81). *Arm pointing.* Over 12 trials, dogs achieved a mean score of 77.14% correct (± 18.65). Performance in the pointing task varied as a function of age, with older dogs scoring higher in this task ($\beta = 26.40$, $t = 2.28$, $p = 0.02$): older dogs were more accurate at following points.

Unsolvable task. Over four 30-second trials, dogs looked to the experimenter's face for an average of 3.30 seconds (± 3.62) per trial.

Memory. In the 20-second delay trials, dogs achieved a mean score of 73.13% correct (± 22.31). In the 40-second delay trials, dogs achieved a mean score of 56.76% correct (± 23.16).

Independent problem-solving. Descriptive statistics for the dependent measures in the independent problem-solving task are reported in Table 15.

Table 15. Descriptive statistics for the independent problem-solving task.

variable	***mean***	***SD***	***minimum***	***maximum***	***SEM***
problem solving A: avg time 1	13.27	9.39	6.37	90.27	0.74
problem solving A: avg time 2	17.96	10.79	4.9	45	0.85
problem solving A: correct attempts	3.37	1.17	0	4	0.10
problem solving A: incorrect attempts	1.97	1.37	0	5	0.11
problem solving A: latency to solve	79.4	31.48	22.5	120	2.59
problem solving B: avg time 1	3.11	4.58	0.92	34.92	0.36
problem solving B: avg time 2	11.71	10	2.62	45	0.79
problem solving B: avg time 3	22	11.82	4.52	45	0.96
problem solving B: avg latency to solve	80.23	32.92	10.28	120	2.85
problem solving B: successful trials	1.25	1.1	0	3	0.10

Visual discrimination. Over 8 trials, dogs achieved a mean score of 90.08% correct (± 13.16).

Auditory discrimination. Over 8 trials, dogs achieved a mean score of 65.47% correct (± 21.49). Auditory discrimination performance varied as a function of age, with younger dogs scoring higher on this task ($\beta = -29.73$, $t = -2.47$, $p = 0.01$).

Odor discrimination. Over 6 trials, dogs achieved a mean score of 60.83% (± 21.49) correct first choices and 71.67% (± 21.38) correct final choices.

One-sample tests

For the tasks in which chance expectation could be defined, we conducted one-sample t-tests to determine whether performance deviated from the null expectation. These tests were performed separately for measures collected in early ontogeny and young adulthood.

Results: Early ontogeny

The results from one-sample t-tests for puppies are shown in Table 16 and summarized below.

Laterality. 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. As a group, the mean bias strength differed from 0 ($t_{167} = 7.69$, $p < 0.001$), and although the population tended toward a left bias, the direction of laterality was not significantly different from 0 at the group level ($t_{167} = -1.93$, $p = 0.06$).

Communicative marker. Subjects' marker cue performance was significantly higher than chance expectation ($t_{165} = 19.02$, $p < 0.001$), indicating that as a group the puppies were able to successfully use the social information in this task.

Arm pointing. Subjects' point-following performance was significantly higher than chance expectation, ($t_{163} = 13.17$, $p < 0.001$) indicating that by 9 weeks of age, dogs are already highly skillful at using human gestural communication.

Odor control trials. Subjects' odor control performance did not differ significantly from chance expectation ($t_{162} = -0.38$, $p = 0.71$) confirming that puppies were unable to reliably find the treat in the absence of social cues, using scent alone.

Memory. At delays of 5s and 10s, puppies' performance in the memory task was significantly higher than chance expectation (5s: $t_{164} = 15.61$, $p < 0.001$; 10s: $t_{164} = 11.72$, $p < 0.001$). For the dogs that met the criterion to advance to longer delays, performance also exceeded chance expectation following 15s and 20s delays (15s: $t_{101} = 8.19$, $p < 0.001$; 20s: $t_{59} = 5.53$, $p < 0.001$).

Table 16. One-sample null hypothesis tests for puppy DCDB measures.

variable	*null hypothesis*	*mean*	*t*	*df*	*p*
laterality: bias strength	25	40.83	7.69	167	0.00
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
odor control trials	50	49.54	-0.38	162	0.71
memory (5s)	50	74.34	15.61	164	0.00
memory (10s)	50	70.24	11.72	162	0.00
memory (15s)	50	65.03	8.19	101	0.00
memory (20s)	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
odor discrimination: first choice	50	53.76	2.31	163	0.02
odor discrimination: final choice	50	72.56	14.61	163	0.00

Visual discrimination. Subjects' performance on the visual discrimination task was significantly greater than chance expectation, ($t_{167} = 41.66$, $p < 0.001$), indicating that as a group, puppies could successfully use visual cues to discriminate between the response options.

Auditory discrimination. Subjects' performance on the auditory discrimination task was significantly greater than chance expectation, ($t_{166} = 5.82$, $p < 0.001$), indicating that as a group, puppies could successfully use auditory cues to discriminate between the response options.

Odor discrimination. Subjects' performance on the odor discrimination task significantly exceeded chance expectation, both in terms of subjects' first choices ($t_{163} = 2.31$, $p < 0.05$), and their final choices ($t_{163} = 14.61$, $p < 0.001$), indicating that as a group, puppies could successfully use odor cues to discriminate between the response options.

Results: Young Adulthood

The results from one-sample t-tests for adult dogs are reported in Table 17 and summarized below.

Laterality. When considering the composite scores from all 32 trials (16 step-up trials and 16 step-down trials), 118 out of 160 dogs (74%) exhibited significant paw preferences. As a group, the strength of mean adult dog bias differed from 0 ($t_{159} = 13.61$, $p < 0.001$), and at the population level, the adult dogs were left-biased ($t_{159} = -2.04$, $p = 0.04$).

Communicative marker. Subjects' marker cue performance was significantly greater than chance expectation ($t_{159} = 36.03$, $p < 0.001$), indicating reliable use of the social information provided by the marker cue.

Arm pointing. Subjects' point-following performance was significantly higher than chance levels, ($t_{159} = 36.03$ $p < 0.001$), indicating that as a group the dogs were able to successfully use the social information provided by human gestural communication.

Odor control trials. Subjects' odor control performance did not significantly differ from chance levels ($t_{159} = 0.56$, $p = 0.58$), confirming that adult dogs were unable to reliably find the treat in the absence of social cues, using scent alone.

Memory. At both 20s and 40s delays, adult dogs remembered the location of the hidden reward at levels exceeding chance expectation (20s: $t_{156} = 27.03$, $p < 0.001$; 40s: $t_{155} = 17.13$, $p < 0.001$), reflecting substantial short-term memory.

Visual discrimination. Subjects' performance on the visual discrimination task was significantly higher than chance levels ($t_{159} = 38.53$, $p < 0.001$), indicating that as a group, young adult dogs could successfully use visual cues to discriminate between the response options.

Auditory discrimination. Subjects' performance on the auditory discrimination task was significantly higher than chance levels ($t_{159} = 10.30$, $p < 0.001$), indicating that as a group, adult dogs could successfully use auditory cues to discriminate between the response options.

Odor discrimination. Subjects' performance on the odor discrimination task was significantly higher than chance expectation, both in terms of subjects' first choices ($t_{159} = 6.38$, $p < 0.05$), and final choices ($t_{159} = 12.82$, $p < 0.001$), indicating that as a group, dogs could successfully use odor cues to discriminate between the two response options.

Table 17. One-sample null hypothesis tests for adult DCDB measures.

variable	null hypothesis	mean	t	df	p
laterality: bias strength	25	59.06	13.61	159	0.00
laterality: laterality index	0	-10.70	-2.04	159	0.04
communicative marker	50	89.32	36.03	159	0.00
arm pointing	50	77.14	18.40	159	0.00
odor control trials	50	50.7	0.56	159	0.58
memory (20s)	25	73.13	27.03	156	0.00
memory (40s)	25	56.76	17.13	155	0.00
visual discrimination	50	90.08	38.53	159	0.00
auditory discrimination	50	65.47	10.30	159	0.00
odor discrimination: first choice	50	60.83	6.38	159	0.00
odor discrimination: final choice	50	71.67	12.82	159	0.00

Changes across ontogeny

Analyses

To assess changes across ontogeny, we conducted paired samples t-tests on DCDB measures collected from dogs at ~9 weeks of age, and again in young adulthood (18-20 months).

Results

The results from paired-sample t-tests are shown in Table 18. As expected, subject performance improved with age across most of the tasks. However, there were a handful of exceptions in which dogs performed similarly in early ontogeny and young adulthood. In the laterality task, the mean laterality bias (as reflected by the laterality index) did not differ between the two time points, ($mean_{puppy}$ = -7.71; $mean_{adult}$ = -10.70; t_{159} = -0.48, p = 0.63), but the strength of this bias significantly increased with age ($mean_{puppy}$ = 40.88; $mean_{adult}$ = 59.06; t_{159} = 5.40, $p < 0.001$). Additionally, no differences were observed between the two age groups on the visual discrimination task ($mean_{puppy}$ = 91.33; $mean_{adult}$ = 90.08; $t(159)$ = -0.90, p = 0.37) or the final choice measure on the odor discrimination task ($mean_{puppy}$ = 72.22; $mean_{adult}$ = 71.67; t_{155} = -0.29, p = 0.77), suggesting that sensory and discriminative capabilities required for these measures have reached adult-like states within the first 2 months of life. Lastly, there were no age differences in performance on the odor control trials ($mean_{puppy} = 49.92$; $mean_{adult} = 50.70$; $t_{154} = 0.42$, $p = 0.68$), with both groups performing at chance expectation.

Table 18. Results from paired t-tests comparing DCDB measures across timepoints.

variable	***puppy mean***	***adult mean***	***mean differenc e***	***t***	***df***	***p***
retrieval: average score	3.3	3.7	0.40	2.96	159	0.00
retrieval: tally	3.01	6.17	3.16	5.46	159	0.00
laterality: bias strength	40.88	59.06	18.19	5.40	159	0.00
laterality: laterality index	-7.71	-10.7	-2.99	-0.48	159	0.63
human interest: total look time	19.32	46.65	26.55	13.62	141	0.00
cylinder: familiarization score	7.85	5.06	-2.81	-10.04	157	0.00
cylinder: inhibitory control score	51.19	75.94	24.60	9.75	157	0.00
cylinder: reversal learning score	29.7	59.59	29.97	11.53	155	0.00
unsolvable task: avg time looking at human	0.98	3.3	2.32	7.79	158	0.00
arm pointing	69.5	77.14	7.53	3.75	155	0.00
communicative marker	76.11	89.32	13.45	8.25	157	0.00
odor control trials	49.92	50.7	0.81	0.42	154	0.68
memory (20s)	63.22	73.13	10.06	2.44	57	0.02
visual discrimination	91.33	90.08	-1.25	-0.90	159	0.37
auditory discrimination	59.2	65.47	6.37	2.87	158	0.00
odor discrimination: first choice	53.31	60.83	7.91	3.30	155	0.00
odor discrimination: final choice	72.22	71.67	-0.64	-0.29	155	0.77

Early life predictors of adult phenotypes

Analyses

To assess longitudinal stability of traits measured by the DCDB, we used two analytical approaches. First, following traditional approaches for assessing the consistency of individual differences across time, we performed rank-order stability analyses by assessing the spearman correlation between phenotypes at the two time points (Caspi, Roberts, & Shiner, 2005). To test the directional prediction that phenotypes at time 1 would be positively related to phenotypes at time 2, we used a directional hypothesis testing framework, following the conventions (δ=0.01, Υ=0.04) recommended by Rice & Gaines (1994). Second, we fit Bayesian linear mixed-models (Stan Development Team, 2018) to assess the relationship between phenotypes at time 1 and time 2, controlling for relatedness between individuals, breed, sex, and (adult) testing location. For these models we converted phenotypic variables to z scores to facilitate interpretation and comparison of beta coefficients.

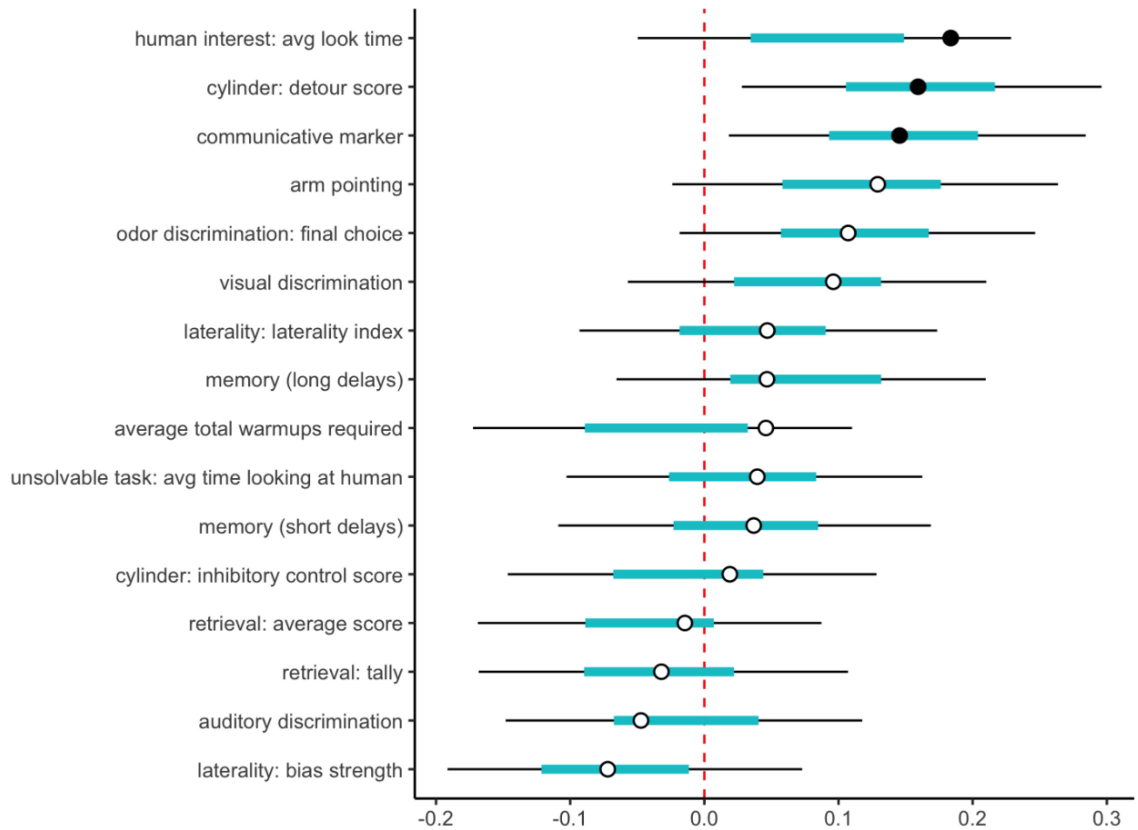
In addition to modeling the stability of individual DCDB measures across time, we also fit exploratory models using all primary phenotypic measures collected at 2 months of age as predictors of each adult measure. Thus, rather than focusing on stability in a given measure across time, these analyses investigated whether any of the phenotypic measures collected from puppies predicted variance in adult phenotypes. For these analyses, we first used lasso regression (Friedman, Hastie, & Tibshirani, 2010) for variable selection, which imposes a penalty (λ) on the beta coefficients, favoring sparse models by shrinking many beta coefficients to 0. To determine the optimal value for λ in these analyses, we used leave-one-out cross validation to obtain the λ value that yielded the minimum cross-validated error. We then fit unrestricted linear models using the subset of variables with non-zero beta coefficients in the lasso model (Hastie, Tibshirani, Friedman, & Franklin, 2005).

Results

Longitudinal stability

The main results from the analyses of longitudinal stability are shown in Figure 2.

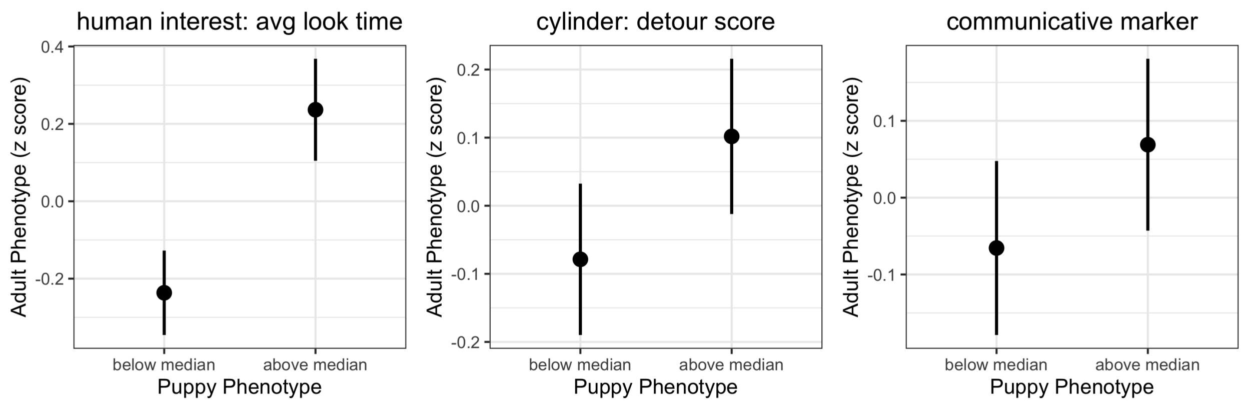
Figure 2. Longitudinal stability of DCDB traits.



Circles reflect the rank-order correlation coefficient between phenotypic measures collected from puppies and adults. Filled circles reflect significant correlations, and open circles reflect correlations with p values > .05. For mixed model analyses, the blue bars span the interquartile range of the posterior probability distribution for the beta coefficient relating puppy phenotypes to adult phenotypes; black lines span the 90% credible interval of the posterior distribution. The dashed red line corresponds to $\beta = 0$, the null hypothesis.

Across traits, rank-order stability analyses yielded spearman correlations ranging from -0.07 to 0.18 (Figure 2). Twelve of the correlation coefficients were positive, and only 4 were negative. A one-sample t-test on the rank-order correlation coefficients indicated that the mean correlation coefficient was significantly greater than 0 (mean = 0.06, $t_{15} = 2.93$, $p = 0.01$), suggesting overall positive relationships between the same traits measured at the two time points. Three individual traits had significant rank-order correlations, all of which were positive (Figure 3). These traits included a measure of attention to a human face during communication (human interest: average looking time), performance in the detour phase of cylinder task (cylinder: detour score), and sensitivity to human communication using an arbitrary communicative marker (communicative marker).

Figure 3. Traits with significant longitudinal stability.



Points and error bars reflect the mean and standard error of the adult phenotype.

The results from linear mixed models controlling for breed, sex, testing location, and relatedness between individuals supported very similar conclusions. The beta coefficients for puppy phenotype as a predictor of the adult phenotype ranged between -0.06 and 0.16 (Figure 2). Ten of these beta coefficients were positive, and 6 were negative. A one-sample t-test indicated that the mean beta coefficient was significantly greater than 0 (mean = 0.04, $t_{15} = 2.39$, $p = 0.03$). For two tasks with positive associations between the puppy and adult phenotype (cylinder: detour scores & communicative marker) the 90% credible interval for the beta coefficient did not contain 0. Therefore, while individual phenotypes changed substantially across development, for a subset of traits involving interest in, and communication with humans, as well as inhibitory control, individual differences in puppies were modestly predictive of adult phenotypes (Figure 3).

Multiple regression models

The results of lasso regressions using all puppy phenotypic measures to predict adult phenotypes are shown in Table 19. For 10 of 16 models, all beta coefficients were shrunk to zero, leaving an intercept only model (data not shown). However, models for the remaining 6 adult measures all retained some puppy phenotypic measures as predictor variables. Unconstrained linear models using these predictor variables revealed several significant associations: (1) Adult performance in the human interest task was positively predicted by puppy performance on the arm pointing and retrieval task. Given that the outcome and predictor variables all involve communication and dyadic interaction with humans, this result may capture a developmentally stable suite of traits involving cooperative interaction with humans. (2) Adult performance on the inhibitory control trials of the cylinder task was positively predicted by puppy performance on the cylinder task (reversal trials) and short-term memory task, and negatively predicted by puppy performance on the retrieval, auditory discrimination, and odor discrimination tasks. Given the positive relationships between variables involving executive function (cylinder task, memory) and negative relationship with retrieval (possibility indexing uninhibited chase behavior), this model may reflect developmental stability in traits related to behavioral inhibition and executive function. (3) Adult performance on the auditory discrimination task was positively predicted by puppy performance on the cylinder (reversal) task, and negatively predicted by puppy performance on the memory tasks (5s delays). Both predictor variables involve components of executive function, but the mechanism linking these traits to adult auditory discrimination remains unclear. Lastly, (4) adults who were the fastest and most successful in the independent problem solving task tended to score higher on retrieval, and lower on the cylinder task (reversal trials) as puppies. Thus, a pattern of uninhibited behavior as a puppy was predictive of successful independent problem solving as an adult, possibly mediated by temperamental factors with relevance to all three of these measures (e.g. uninhibited energetic physical engagement).

Table 19. Linear models predicting adult phenotypes from puppy phenotypes.

outcome	*predictor*	*β*	*t*	*p*
human interest: avg look time $r^2 = 0.16$	arm pointing	0.2082	2.6516	0.0089
	retrieval (high engagement)	0.2014	2.6691	0.0084
cylinder: inhibitory control score $r^2 = 0.23$	cylinder: reversal	0.1513	1.9793	0.0497
	memory (10s)	0.2786	3.6342	<0.0000
	auditory discrimination	-0.1972	-2.6443	0.0091
	retrieval (high engagement)	-0.1439	-2.002	0.0472
	odor discrimination (attention to incorrect location)	-0.1434	-2.0471	0.0425
auditory discrimination $r^2 = 0.10$	cylinder: reversal	0.2771	3.5933	<0.0000
	memory (5s)	-0.1796	-2.27	0.0246

outcome	*predictor*	*β*	*t*	*p*
problem solving (fast and successful) $r^2 = 0.32$	cylinder: reversal	-0.1656	-2.2337	0.027
	retrieval (high engagement)	0.234	3.362	0.001

In summary, several cognitive traits – including spontaneous interest in communication with humans, use of human communicative signals, and inhibitory control – exhibit consistency across development. Further, we find evidence that for some traits – including human interest, inhibitory control, auditory discrimination, and independent problem solving – adult phenotypes can be predicted by leveraging multiple predictor variables collected from puppies. Collectively, these findings raise the possibility of screening for certain characteristics early in a dog's lifetime, and identify a subset of traits for which this approach may be most profitably employed.

Heritability

Analyses

To assess the heritability (*h^2;* the proportion of trait variance attributable to genetic factors) of traits measured by the DCDB, we used a Bayesian implementation of the 'animal model' (Wilson et al., 2010). The animal model is a mixed model which includes a random effect for each individual's 'breeding value' (defined as the additive effect of this individual's genotype relative to the population mean phenotype) and an (additive) genetic variance-covariance matrix describing relatedness between individuals in the analysis. The model estimates the variance attributable to genetic and environmental factors. Using these variance components, heritability is estimated as the proportion of the total phenotypic variance attributable to genetic factors.

To estimate heritability in our study population, we first compiled a pedigree containing >20,000 known individuals from the breeding population at Canine Companions for Independence. We then generated a kinship matrix from this pedigree to be used in all analyses. Models were fit using a 5,000 iteration burn in, followed by 10,000 iterations of sampling, with retention of every 25th value for the posterior distribution. Each model was run across 4 independent MCMC chains. To calculate the h^2 estimate for each trait, we retained the mean genetic and environmental variance component estimates across chains. Because heritability is a population-specific measure that can vary across time and environments (Visscher, Hill, & Wray, 2008), we calculated heritability estimates for three different sets of measures from our project:

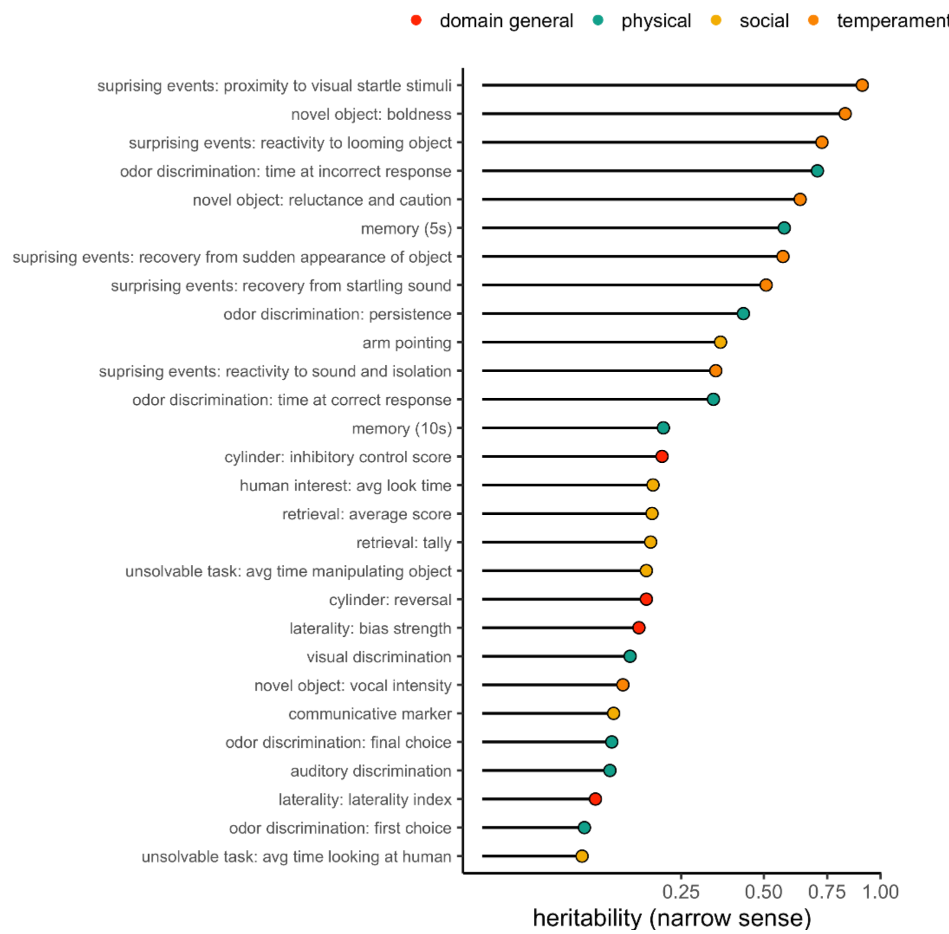
(1) puppy phenotypes ~ 2 months of age [N = 168] (2) young adult phenotypes from the sample of dogs studied longitudinally [N = 160], and (3) phenotypes for all adult dogs tested, including longitudinal study subjects and other breeders and relatives in the breeding population [N = 323] .

Results

Heritability of DCDB phenotypes in puppies

Heritability estimates for the primary measures from the DCDB implemented with 8-week old puppies are shown in Figure 4. Overall, temperamental traits were characterized by the highest heritability, with more than 70% of variation in reaction to startle stimuli, reaction to looming objects, and boldness when alone with a novel object, explained by genetic factors. Other traits with modest to high heritability included 3 measures related to olfactory discrimination (mean h^2 = 0.49), two measures related to memory (5-second delays: $h^2 = 0.54$; 10-second delays: $h^2 = 0.21$), and three social cognitive measures (response to arm pointing: $h^2 = 0.36$; human interest: $h^2 = 0.18$; retrieval: $h^2 = 0.18$). Thus, from early in ontogeny individual differences in dog cognitive and temperamental traits are shaped strongly by genetic factors.

Figure 4. Heritability of DCDB measures in early ontogeny.

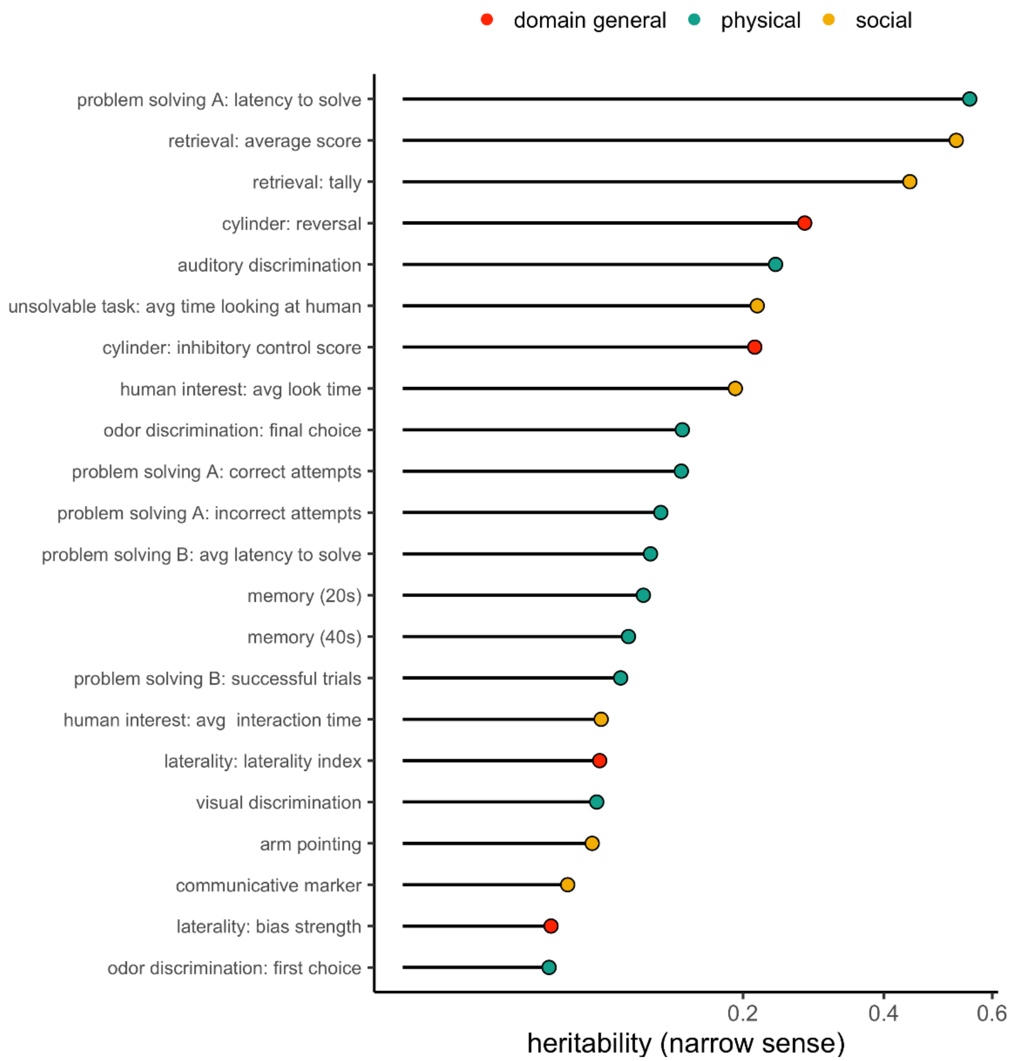


<u>*Heritability of DCDB phenotypes in young adults (longitudinal sample)*</u>

Heritability estimates for the primary measures from the DCDB implemented at the second timepoint in our longitudinal study are shown in Figure 5[2]. The traits with the highest heritability included latency to success in the independent problem-solving task ($h^2 = 0.55$), and two measures from the retrieval task (overall score: $h^2 = 0.53$; retrieval tally: $h^2 = 0.44$). Performance in the cylinder, human interest, and odor discrimination task also exhibited moderately high heritability.

[2] Temperament measures from adults are still being coding and are not included in this report.

Figure 5. Heritability of DCDB measures in young adulthood

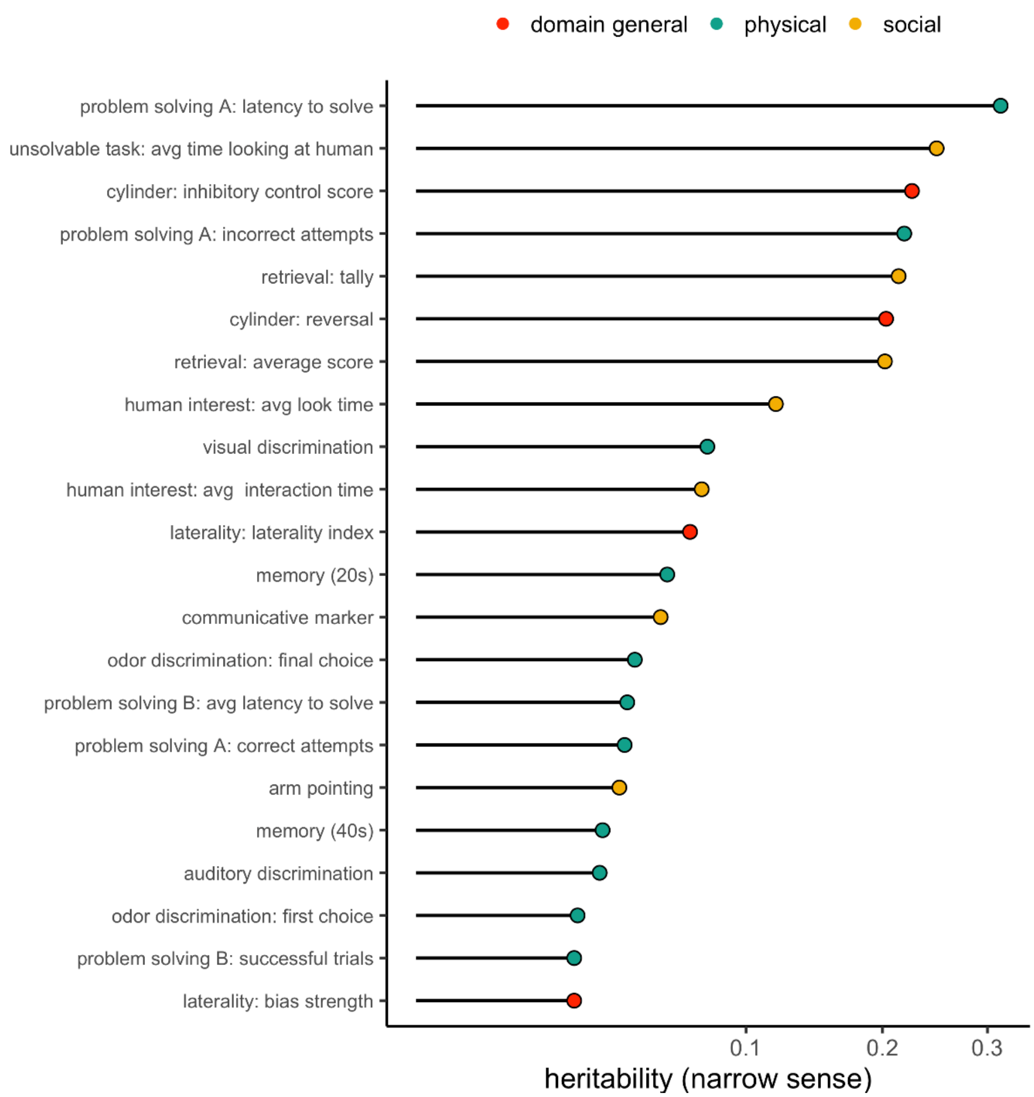


<u>*Heritability of DCDB phenotypes in all adults*</u>

Heritability estimates for the primary measures from the DCDB implemented with the full sample of adult dogs, including longitudinal study subjects and other adults in the breeding colony are shown in Figure 6. Similarly to the analyses restricted to the longitudinal cohort, the most heritable traits included latency to success in the independent problem-solving task ($h^2 = 0.31$), performance in the cylinder task (inhibitory control trials: $h^2 = 0.23$; reversal trials: $h^2 = 0.20$) and performance in the retrieval task (overall score: $h^2 = 0.20$, retrieval tally: $h^2 = 0.21$). However, across the full adult sample, time gazing at a human during the unsolvable task (a measure of help seeking) was also among the most heritable traits ($h^2 = 0.25$).

Collectively, these analyses reveal that a substantial fraction of variance in cognitive and temperamental measures is attributable to genetic factors. Given that several of these traits have been linked to success in diverse working roles (Bray, Sammel, Cheney, Serpell, & Seyfarth, 2017; MacLean & Hare, 2018), our results suggest that military working dog programs have great potential for improvement through the selective breeding of dogs with desirable characteristics.

Figure 6. Heritability of DCDB measures across the full adult sample



DISCUSSION

The principal findings from this project were that (1) many cognitive abilities critical to working dog success emerge early in ontogeny, and are quantifiable using the Dog Cognitive Development

Battery, (2) perceptual discriminations reach adult-like states by ~9 weeks of age, whereas cognitive processes involving memory, social cognition, and executive function develop more slowly across ontogeny, (3) several traits exhibit developmental stability, meaning that adult phenotypes can be predicted from measures collected early in development, and (4) most cognitive and behavioral traits exhibit substantial heritability, indicating a strong genetic basis for individual differences. Below we synthesize these primary findings and their relevance for military working dog programs.

Early Ontogeny of Dog Cognition

We found that by 9 weeks of age, dogs were highly skillful at using communicative signals from humans, including natural gestures such as arm pointing, and novel communicative cues, such as an arbitrary marker paired with ostensive signals. Thus, as suggested by previous studies with smaller samples (Gácsi et al., 2009; Gacsi, Kara, Belenyi, Topal, & Miklosi, 2009; Hare, Brown, Williamson, & Tomasello, 2002; Riedel, Schumann, Kaminski, Call, & Tomasello, 2008; Virányi et al., 2008), dogs appear to be biologically prepared for communication with humans, and these skills – which are associated with success as a MWD – emerge early in ontogeny.

9-week old dogs also performed above chance expectation on visual, auditory, and olfactory perceptual discriminations. Initial discrimination between the two response options was best for visual discriminations, followed by olfactory and auditory discriminations, respectively. Despite minimal initial preference in the olfactory discrimination task (measured as the first response option approached), 9-week old dogs spent the majority of their time investigating the baited option in the olfaction test and had chosen this option ~73% of the time by the end of the trial. Across development dogs showed no increase in accuracy on the visual discrimination task; however, performance was near ceiling at both timepoints. Performance in both the auditory and olfactory discrimination tasks increased across development, but the extent of change was relatively small (auditory discrimination: + 6%; olfactory discrimination: + 8%). Performance in social cue tasks also increased modestly, with scores on the arm pointing task increasing 8%, and scores on the communicative marker task increasing 13% on average. In contrast, measures of executive function increased markedly across development, including performance in the inhibitory control trials (+ 25%), and reversal trials (+ 30%) of the cylinder task. Similarly, performance in memory tasks improved substantially; whereas puppies averaged 70% correct on memory trials after 10s delays, adults performed at 73% accuracy at delays twice this length. Thus, as in humans (Best & Miller, 2010; De Luca et al., 2003; Luciana, Conklin, Hooper, & Yarger, 2005; Luna, Garver, Urban, Lazar, & Sweeney, 2004), executive functions are relatively late to develop in dogs.

Developmental Predictors of Adult Phenotypes

Analyses across development revealed significant rank-order stability for three traits, including interest in / attention to humans, use of an arbitrary communicative marker, and a measure of self-control. The former two findings suggest stability in emotional and communicative processes that were likely under selection during domestication (Hare & Tomasello, 2005; MacLean & Hare, 2015; MacLean et al., 2017), and indicate that these traits could be screened for from very early in a dog's lifetime. The latter finding on self-control is consistent with similar results in humans, which suggest that self-control is among the most heritable and developmentally-stable psychological traits (Moffitt et al., 2011; Shoda, Mischel, & Peake, 1990). In addition to these associations, we found that by taking a multivariate approach, it was possible to predict several adult phenotypic measures based on a linear combination of phenotypic measures collected early in ontogeny. For example, adult attention to humans was predicted by puppy performance in tasks measuring sensitivity to gestural communication and retrieval. Adult inhibitory control was positively associated with measures of memory and inhibitory control in puppies, and negatively associated with retrieval as a puppy (uninhibited chasing behavior). This approach also facilitates prediction of complex phenotypes in adults which are not easily measured in younger dogs. For example, adult performance in a multi-step independent problem-solving task (which has been linked to training success in guide dogs) was positively predicted by simple measures of executive function and retrieval when dogs were only two months of age. Thus, the test battery approach implemented here provides rich opportunities for prediction of adult phenotypes using a constellation of phenotypic measures collected early in development. Based on our results, we expect that developmental screening will be most successful when targeting traits involving sensitivity to gestural communication, attentiveness to humans, and impulse control, phenotypes which have been linked to working dog success in previous studies (Duffy & Serpell, 2012; MacLean & Hare, 2018).

Genetic contributions to individual differences

Quantitative genetic analyses of traits measured by the DCDB showed that genetic factors accounted for substantial variation in the majority of traits. In puppies, more than 50% of variance in most temperamental measures was attributable to genetic factors. Although temperament is known to have a heritable basis in dogs, heritability estimates for temperamental traits are typically between 0.1-0.4, as measured in both adult dogs and puppies (Arvelius, Strandberg, & Fikse, 2014; Ilska et al., 2017; Van der Waaij, Wilsson, & Strandberg, 2008; Wilsson & Sundgren, 1997, 1998). The heritability estimates for temperament traits in our analysis are thus markedly higher than those reported in previous studies, perhaps due to the specific test and scoring protocols we employed (which differed from previous studies and involved detailed scoring from video recordings using a standardized ethogram with high inter-rater reliability). In addition to these temperament measures, we also found that genetic factors accounted for a notable proportion of the variation in several cognitive and perceptual measures, including sensitivity to human gestural communication, memory, inhibitory control, olfactory discrimination, and attention to humans. To our knowledge, no previous studies have investigated the heritability of these cognitive traits in puppies, and thus our project contributes foundational data in this area.

Our study design enabled a powerful assessment of heritability in adulthood, because all subjects were reared in different environments throughout the western United States between 8 weeks and 18 months of age. Thus, effects of the shared environment, which are common in sibling studies, were largely eliminated in our design. In this sense, our design was analogous to classic 'twins reared apart' studies (e.g., Bouchard, Lykken, McGue, Segal, & Tellegen, 1990), with the exception that we used littermates, rather than monozygotic twins as study subjects. Despite varied rearing environments, we found that many traits were highly heritable in young adulthood. Specifically, more than half of the variation in measures related to independent problem solving and retrieval was attributable to genetic factors. Many other measures, including performance in the cylinder task, auditory discrimination, and gaze at humans when faced with an unsolvable problem, had heritability estimates between 0.2 and 0.3, suggesting substantial genetic influences. For comparative purposes, these estimates are well within the range of heritability values for other quantitative traits which have successfully been selected for in domestic animal populations, including hip distraction scores in dogs (h^2 = 0.21-0.64; Guilliard, 2014; Hamann, Kirchhoff, & Distl, 2003; Leighton, Linn, Willham, & Castleberry, 1977; Wood, Lakhani, & Rogers, 2002), milk yield in cattle (h^2 = ~0.20; Hill, Edwards, Ahmed, & Thompson, 1983; Lofgren, Vinson, Pearson, & Powell, 1985), and clutch size in chickens (h^2 = 0.18; Wolc et al., 2018). Therefore, there is strong potential for selection on cognitive and temperamental phenotypes in working dog populations.

Conclusions and future directions.

In summary, we found that (1) cognitive and temperamental traits critical to working dog success emerge early in development and are quantifiable using the Dog Cognitive Development Battery, (2) in some cases, adult phenotypes can be predicted using simple measures collected from puppies, and (3) most cognitive and temperamental traits measured by the Dog Cognitive Development battery are heritable, and thus potential targets for selection in working dog programs. These findings provide an empirical foundation for efforts to breed and select for traits linked to success in working dog programs.

Our findings raise several important questions for future research. First, although we find evidence for developmental stability and genetic influences on many traits, a considerable amount of variation in adult phenotypes is not explained by these factors. For complex traits, phenotypes most often arise through processes involving both genetic and environmental effects, as well as gene × environment interactions. Our current work focused on genetic and developmental patterns that were robust to environmental influences and did not explicitly address environmental effects or gene by environment interactions. Ultimately, our ability to predict, or cultivate desired phenotypes in working dogs will also require similar attention to the environmental factors that no doubt contribute to the remaining unexplained variation. Replication and extension of the current design would provide a powerful opportunity for this work, because genetic and environmental effects could be decoupled due to the raising of littermates in heterogenous environments. Thus, just as our current design allowed us to limit effects of shared environment when assessing genetic contributions to phenotypes, future work could similarly control for genetic factors while more systematically addressing environmental contributions to phenotypic variance. Coupled with knowledge about heritability and developmental plasticity / stability, a better understanding of how variation in rearing environments impacts adult phenotypes will further our ability to create dogs well-suited for particular working roles.

Secondly, although genetic factors explain considerable variance in the traits we measured, the underlying molecular associations remain unknown. Are individual differences driven by small effects of many variants (i.e. polygenic effects), or by a small set of loci with relatively large effects? What regions of the genome are associated with individual differences in cognitive and temperamental traits, and what are the pathways through which these variants yield their effects? Although the answers to these questions are not required for selective breeding efforts (which make use of additive variation across the genome), identifying specific genetic variants associated with these traits creates other powerful opportunities. For example, with knowledge of the underlying genetic mechanisms, it may become possible to employ gene editing techniques (e.g. CRISPR) to further enhance desirable characteristics. However, the success of this approach will be critically dependent on initial genome-wide association analyses (e.g., MacLean, Snyder-Mackler, vonHoldt, & Serpell, 2019), which can identify regions of the genome causally linked to phenotypic variation. Similarly, if key molecular variants can be identified, selection of dogs can be guided by marker-based approaches, using genomic information at the selection / procurement stage. Ultimately, this approach may allow defense agencies to assess a dog's potential simply by collecting a cheek swab, eliminating many of the complicating factors which are common to current procurement efforts. With the advent of mobile genotyping platforms (e.g., Daniels et al., 2012), this approach may become available in the near future, yet again will depend on initial basic research to identify the relevant genetic variants. The success of such efforts will also depend on mapping traits with a strong genetic basis, a key determinant of power in genome-wide association studies (Shin & Lee, 2015). The heritability estimates observed in our current work suggest that the DCDB will provide a powerful instrument for this work, which can ultimately identify genetic bases of cognitive and behavioral traits critical to MWD success.

REFERENCES

Arvelius, P., Strandberg, E., & Fikse, W. F. (2014). The swedish armed forces temperament test gives information on genetic differences among dogs. *Journal of Veterinary Behavior: Clinical Applications and Research, 9*(6), 281-289.

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.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child development, 81*(6), 1641-1660.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The minnesota study of twins reared apart. *Science, 250*(4978), 223-228.

Bray, E. E., MacLean, E. L., & Hare, B. A. (2015). Increasing arousal enhances inhibitory control in calm but not excitable dogs. *Animal cognition, 18*(6), 1317-1329.

Bray, E. E., Sammel, M. D., Cheney, D. L., Serpell, J. A., & Seyfarth, R. M. (2017). Effects of maternal investment, temperament, and cognition on guide dog success. *Proceedings of the National Academy of Sciences, 114*(34), 9128-9133.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annu. Rev. Psychol., 56*, 453-484.

Daniels, R., Ndiaye, D., Wall, M., McKinney, J., Séne, P. D., Sabeti, P. C., et al. (2012). Rapid, field-deployable method for genotyping and discovery of single-nucleotide polymorphisms associated with drug resistance in plasmodium falciparum. *Antimicrobial agents and chemotherapy, 56*(6), 2976-2986.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., et al. (2003). Normative data from the cantab. I: Development of executive function over the lifespan. *Journal of clinical and experimental neuropsychology, 25*(2), 242-254.

Duffy, D. L., & Serpell, J. A. (2012). Predictive validity of a method for evaluating temperament in young guide and service dogs. *Applied Animal Behaviour Science*.

Fiset, S., Landry, F., & Ouellette, M. (2006). Egocentric search for disappearing objects in domestic dogs: Evidence for a geometric hypothesis of direction. *Animal cognition, 9*(1), 1-12.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software, 33*(1), 1.

Gácsi, M., Györi, B., Gyoöri, B., Virányi, Z., Kubinyi, E., Range, F., 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), e6584.

Gacsi, M., Kara, E., Belenyi, B., Topal, J., & Miklosi, A. (2009). The effect of development and individual differences in pointing comprehension of dogs. *Animal Cognition, 12*(3), 471-479.

Guilliard, M. (2014). The pennhip method of predicting canine hip dysplasia. *In Practice, 36*(2), 66-74.

Hamann, H., Kirchhoff, T., & Distl, O. (2003). Bayesian analysis of heritability of canine hip dysplasia in german shepherd dogs. *Journal of Animal Breeding and Genetics, 120*(4), 258-268.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science, 298*(5598), 1634-1636.

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences, 9*(9), 439-444.

Hare, B., & Woods, V. (2013). *The genius of dogs: How dogs are smarter than you think*: Dutton Adult.

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: Data mining, inference and prediction. *The Mathematical Intelligencer, 27*(2), 83-85.

Hill, W. G., Edwards, M., Ahmed, M.-K., & Thompson, R. (1983). Heritability of milk yield and composition at different levels and variability of production. *Animal Science, 36*(1), 59-68.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185.

Ilska, J., Haskell, M. J., Blott, S. C., Sánchez-Molano, E., Polgar, Z., Lofgren, S. E., et al. (2017). Genetic characterization of dog personality traits. *Genetics, 206*(2), 1101-1111.

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping". *Science, 304*(5677), 1682-1683.

Kaulfuß, P., & Mills, D. (2008). Neophilia in domestic dogs (canis familiaris) and its implication for studies of dog cognition. *Animal Cognition, 11*(3), 553-556.

Lazarowski, L., & Dorman, D. C. (2014). Explosives detection by military working dogs: Olfactory generalization from components to mixtures. *Applied Animal Behaviour Science, 151*, 84-93.

Lazarowski, L., Foster, M. L., Gruen, M. E., Sherman, B. L., Fish, R. E., Milgram, N. W., et al. (2015). Olfactory discrimination and generalization of ammonium nitrate and structurally related odorants in labrador retrievers. *Animal cognition, 18*(6), 1255-1265.

Leighton, E. A., Linn, J. M., Willham, R. L., & Castleberry, M. (1977). A genetic study of canine hip dysplasia. *Am J Vet Res, 38*(2), 241-244.

Lofgren, D., Vinson, W., Pearson, R., & Powell, R. (1985). Heritability of milk yield at different herd means and variance for production. *Journal of Dairy Science, 68*(10), 2737-2739.

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child development, 76*(3), 697-712.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child development, 75*(5), 1357-1372.

MacLean, E., Snyder-Mackler, N., vonHoldt, B., & Serpell, J. (2019). Highly heritable and functionally relevant breed differences in dog behavior. *bioRxiv*.

MacLean, E. L., & Hare, B. (2015). Dogs hijack the human bonding pathway. *Science, 348*(6232), 280-281.

MacLean, E. L., & Hare, B. (2018). Enhanced selection of assistance and explosive detection dogs using cognitive measures. *Frontiers in Veterinary Science, 5*, 236.

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.

McGreevy, P. D., Brueckner, A., Thomson, P. C., & Branson, N. J. (2010). Motor laterality in 4 breeds of dog. *Journal of Veterinary Behavior-Clinical Applications and Research, 5*(6), 318-323.

Miklósi, Á. (2014). *Dog behaviour, evolution, and cognition*: Oxford University Press.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*(7), 2693-2698.

Osborne, J. W., & Costello, A. B. (2009). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pan-Pacific Management Review, 12*(2), 131-146.

Pilley, J. W., & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes, 86*(2), 184-195.

Pongracz, P., Miklosi, A., Kubinyi, E., Gurobi, K., Topal, J., & Csanyi, V. (2001). Social learning in dogs: The effect of a human demonstrator on the performance of dogs (*canis familiaris*) in a detour task. *Animal Behaviour, 62*(6), 1109-1117.

Rice, W. R., & Gaines, S. D. (1994). 'Heads i win, tails you lose': Testing directional alternative hypotheses in ecological and evolutionary research. *Trends in Ecology & Evolution, 9*(6), 235-237.

Riedel, J., Schumann, K., Kaminski, J., Call, J., & Tomasello, M. (2008). The early ontogeny of human–dog communication. *Animal Behaviour, 75*(3), 1003-1014.

Shettleworth, S. J. (2010). *Cognition, evolution, and behavior* (2nd ed.): Oxford ; New York : Oxford University Press, 2010.

Shin, J., & Lee, C. (2015). Statistical power for identifying nucleotide markers associated with quantitative traits in genome-wide association analysis using a mixed model. *Genomics, 105*(1), 1-4.

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology, 26*(6), 978.

Stan Development Team. (2018). Rstan: The r interface to stan. R package version 2.17. 3.

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.

Van der Waaij, E., Wilsson, E., & Strandberg, E. (2008). Genetic analysis of results of a swedish behavior test on german shepherd dogs and labrador retrievers. *Journal of animal science, 86*(11), 2853-2861.

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.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nature reviews genetics, 9*(4), 255.

Wells, D. L. (2003). Lateralised behaviour in the domestic dog, canis familiaris. *Behavioural processes, 61*(1-2), 27-35.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., et al. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology, 79*(1), 13-26.

Wilsson, E., & Sundgren, P.-E. (1997). The use of a behaviour test for selection of dogs for service and breeding. Ii. Heritability for tested parameters and effect of selection based on service dog characteristics. *Applied Animal Behaviour Science, 54*(2), 235-241.

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.

Wolc, A., Jankowski, T., Arango, J., Settar, P., Fulton, J., O'Sullivan, N., et al. (2018). Investigating the genetic determination of clutch traits in laying hens. *Poultry science, 98*(1), 39-45.

Wood, J., Lakhani, K., & Rogers, K. (2002). Heritability and epidemiology of canine hip-dysplasia score and its components in labrador retrievers in the united kingdom. *Preventive veterinary medicine, 55*(2), 95-108.

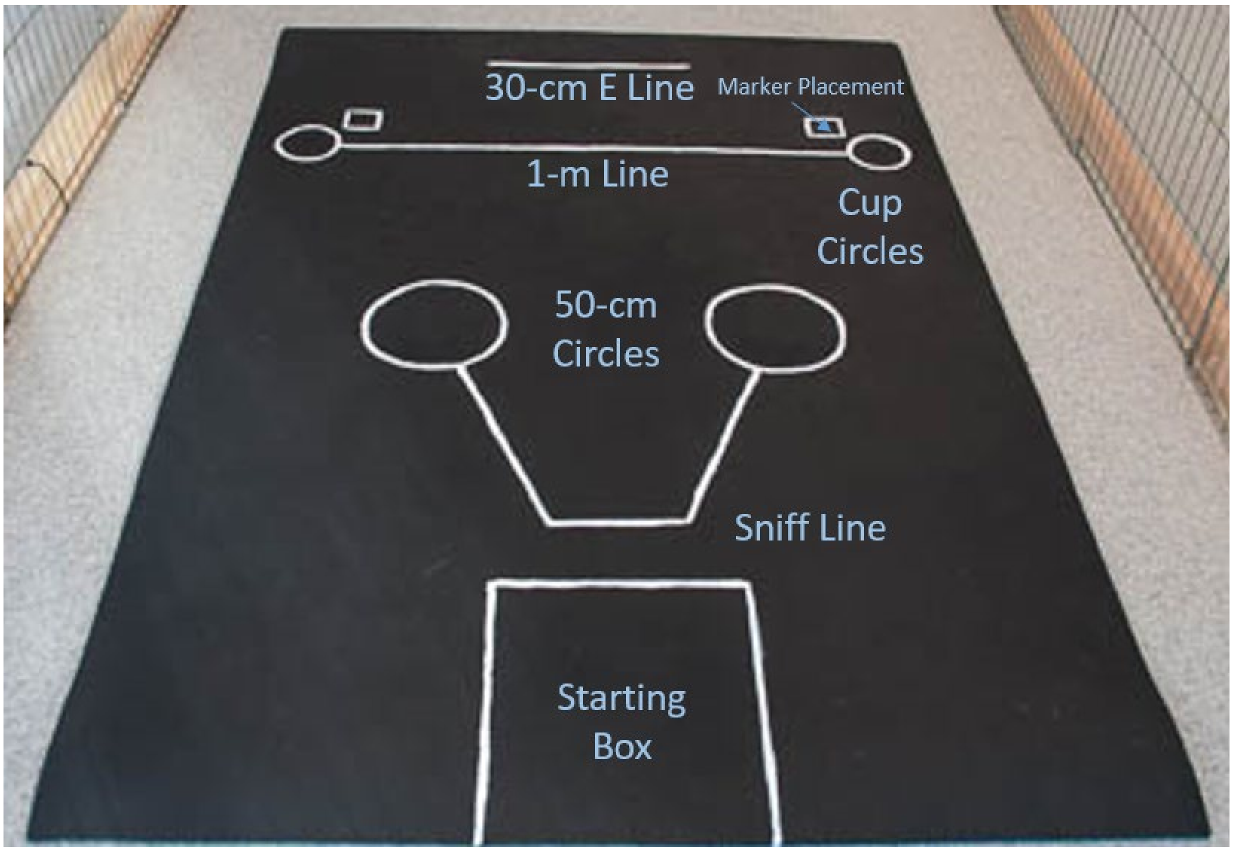
PROTOCOLS

Experimental Procedures

General Methods

A 6’ x 10’ exercise pen is placed around the testing area, which includes a centered 4’ x 6’ testing mat. The layout of the testing setup and cameras is shown in Figures 7-9.

Figure 7. General Experimental setup



<u>*Object choice task guidelines*</u>

A choice is defined as the subject physically touching the cup with their snout or a front paw. Each trial allows the dog 25s from the release to make a choice. When a choice is made, the handler (hereafter H) says, “choice”. If the experimenter (hereafter E) sees the choice first, despite looking down, E may also say “choice”. If the subject chooses the baited cup, E praises the subject before H repositions the subject in the starting box. If the subject chooses the incorrect cup, E says “wrong” in a neutral, monotone voice. E then returns the subject to H and retrieves the reward from the correct location before the next trial is administered.

E presents the reward by holding the treat in her fingers, leaning forward so the dog can sniff and lick it, and saying “Puppy, look!” If the dog is not attending, E can try to get their attention by repeating “puppy look”, waving her hand in front of the puppy, or whistling.

Handler Guidelines

At the beginning of each trial, H centers the subject in the starting box (Figure 7) and looks straight down. Once E has given the "okay" release signal, H releases the puppy. If the dog does not leave the starting box, H may reorient or lift the dog by the haunches, as described above. As soon as the subject leaves the starting box, H looks up. When the dog makes a choice, H says, "choice".

Experimenter Guidelines

E begins by kneeling, centered behind the experimenter's line (~30 cm from cup line). E leans forward, presents the reward, returns to the center position, baits the appropriate cup, and either performs the social cue or kneels in resting position. At the end of the baiting procedure, E gives an "okay" release signaling H to let go of the subject. During the trial, E looks down (or at the appropriate cup for the pointing task) and does not make eye contact with the subject. When the subject chooses a cup, E tips the cup to let the subject access the treat or see that there is no treat. E may pull the cup away completely if the dog cannot find the treat or is playing with the cup.

Refamiliarization / Abort for Social Cues and Memory

If the subject does not make a choice within 25 s, the trial is repeated. If the dog no-choices (NC) twice in a row, 2 trials of two-cup alternating warm-ups may be done. If at any point there are two more consecutive NC, increase the food reward and refamiliarize. If the dog does not engage in refamiliarization trials or makes 8 total no-choices within a single task, abort.

Figure 8. Camera placements

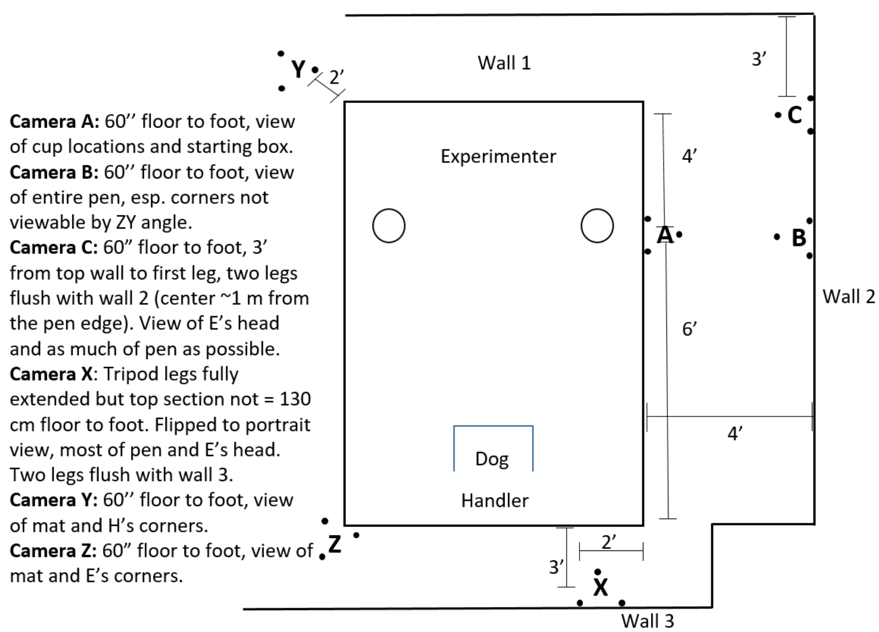
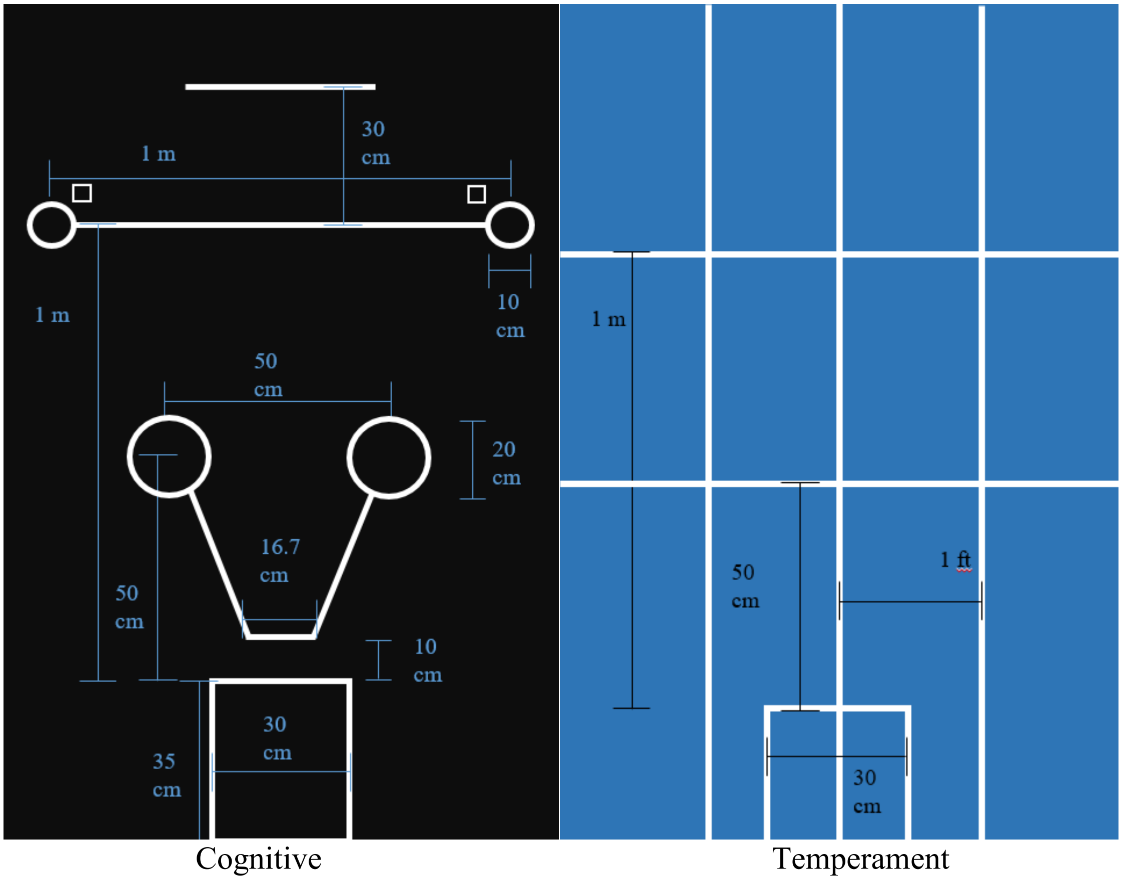


Figure 9. Mat Diagrams

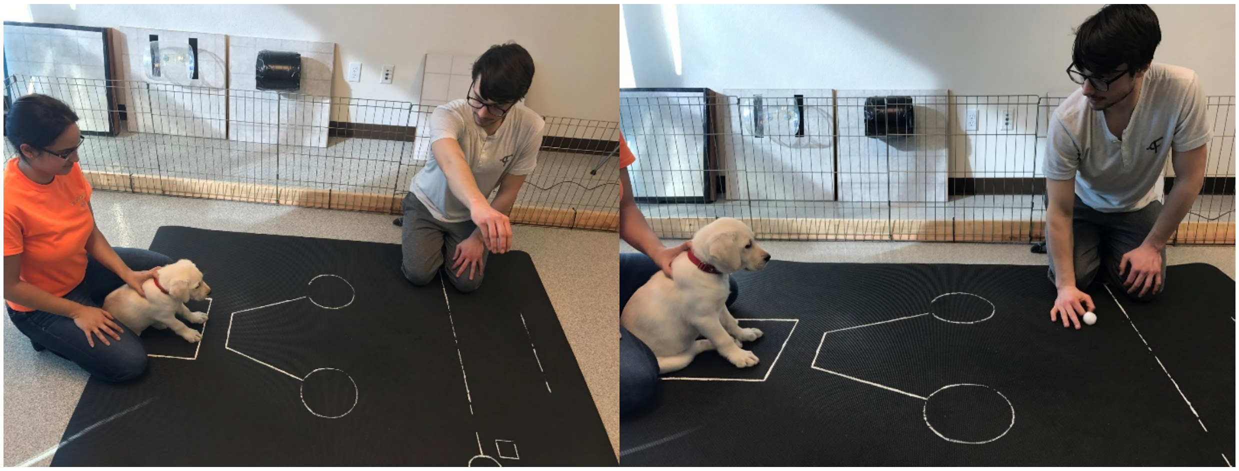


Day one; Session one

<u>*Cotton Ball Vision Test*</u>

H positions the dog in the starting box. E kneels on the floor in line with the cup marks (about 1 m from the dog) and holds the cotton ball and jingle ball in one hand with arm extended toward midline. E gets the dog's attention by jingling the jingle ball, moving her arm only as need to attract attention. When E is sure that the subject is looking at her hand, she drops the cotton ball from arm's height (Figure 10). The experimenter watches for the puppy's eyes /head to follow the cotton ball. This is repeated twice more, for a total of three trials. Then the experimenter repositions her arm near her body and flicks the cotton ball lightly horizontally across the ground and watches for the dog's eyes or head to follow the cotton ball. This is repeated twice more, for a total of three trials. If the dog looks away before the experimenter drops the ball, the trial may be repeated (without penalizing the subject). If a dog cannot follow the cotton ball at 1 m (within approximately 5 minutes), give the dog a break and try again later (without proceeding to other tests). Live coding will be a binary choice of followed/did not follow for each trial.

Figure 10. Procedure for the vision test.



Cotton ball drop test being performed at 100 cm from the puppy. Evidence that the puppy follows the cotton ball's movement is considered a positive response.

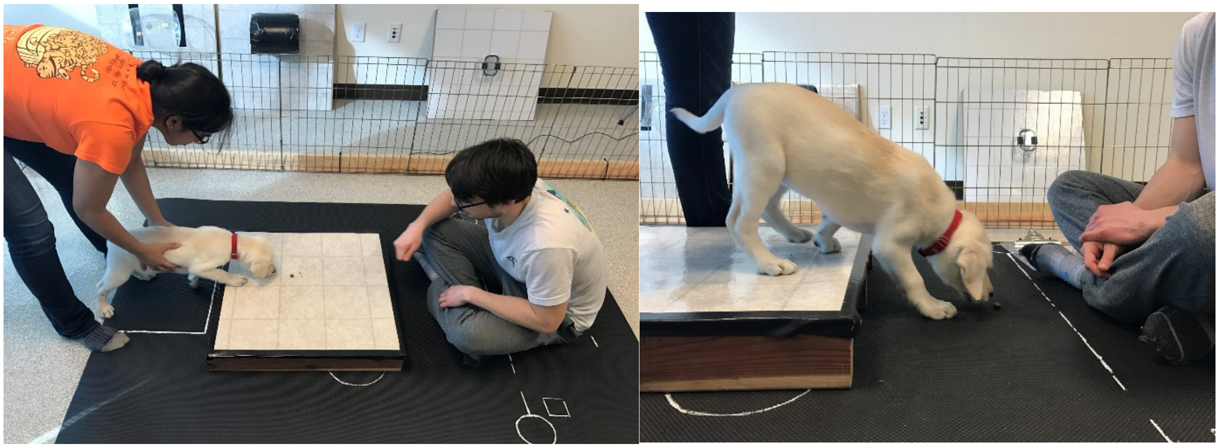
<u>Cooperative Willingness: Retrieval</u>

Positioning and setup:
H sits outside of the X-pen with the stopwatch and is responsible for all timing and recording. H will keep track of the dog's retrievals with tallies. E stands in the starting box marked on the mat. During test trials, at the time of the throw, the dog should be in front of E, facing the direction that the ball is being thrown.

Procedure:
E throws the ball for the dog for 1 minute as a warm-up, the goal of which is to get and maintain the dog's interest and initiate fetching, if possible. During the warm-up, E may move around freely. The warm-up is followed by two 60-second test trials. E does not move from the start line during these trials. At the start of each trial, E bends down and throws the jingle ball in front of the dog, as the dog is watching, then stands up with legs straight. E is allowed to bend over and lightly clap or tap thighs to encourage subject to fetch. H starts a count-up timer for 60 seconds (the length of the trial). During the test trials, E does not move their feet, but verbally encourages the dog to retrieve the ball. A retrieval is considered any behavior that returns the ball to within E's reach (without squatting).

Scoring:
The dependent measures for each trial are the tally and a retrieval score from 1 to 5: Score 1 is assigned if the dog shows no interest in the ball. Score 2 is assigned if the dog runs after the ball and touches it but doesn't pick it up in its mouth. Score 3 is assigned if the dog carries the ball but doesn't bring it back to within touching distance of the experimenter. Score 4 is given to dogs who retrieve the ball and bring it back to the experimenter 1-2 times. Score 5 is given to dogs who retrieve the ball and bring it back to the experimenter 3 or more times.

Figure 11. Procedure for the laterality task.



<u>Laterality</u>

Positioning and setup:
The step apparatus is placed approximately 5 cm in front of the starting box. Before conducting the trials, H coaxes the dog onto and off of the step a few times to ensure that she is not fearful of the apparatus. H should kneel or stand directly behind the subject, restraining the dog but touching her as little as possible, and making sure they can see the dog's front feet. The dog should be facing E, sitting if possible, and completely straight. If H needs to reposition the subject, they should lift the dog straight up and not touch the forepaws or forelegs.

Step up:
H positions the dog in the starting box, squarely in front of the step, with front paws approximately 5-10 cm from the step. E shows the dog the reward, saying "puppy, look!" E pulls the reward backward, places it in the center of the platform, and says "okay". H then releases the dog (Figure 11). If the dog is leaping, E may place the kibble closer or farther from the dog to encourage steps.

Step down:
This task is identical to the step-up task, except the dog begins on the platform, squarely positioned and sitting or standing near the edge (~5 cm). E shows the subject the reward, saying "puppy, look!" E pulls the reward backward, places it on the ground in front of the platform (~10-20 cm), directly ahead of the dog, and says "okay."

Scoring:
The first paw to touch the step or ground as the dog steps up onto or down off of the platform, respectively, is recorded. 15 trials are conducted each for step-up and step-down. If the subject leaps with both paws simultaneously, the trial is repeated. H calls out the paw used and E confirms and records; if E and H disagree about which paw was used, mark 'code' and check later from video before entering data.

Abort Criteria:
If the dog cannot be coaxed onto the step within 5 minutes, the task is aborted. If at any time, the subject is not participating, a break may be taken for play or the food reward may be increased. If the task cannot be completed within 15 min, abort and proceed to warm-ups.

Warm-Ups

The subject is required to pass a warm-up criterion prior to completing any other object choice tasks. Warm-up trials ensure that the subjects are motivated to search for the reward and to prevent side biases. These trials consist of three phases: (1) no-cup visible placement and free-form cup familiarization game, (2) one-cup alternating visible placement, and (3) two-cup visible placement.

Positioning and setup:
H holds subject in the starting box. E kneels in the center, approximately 30 cm behind the cup line or 1.3 m from the puppy. When baiting, E should lift the cup ~10 cm off the ground.

Phase 1 – no-cup visible placement and free-form cup game
E presents the reward then visibly places the food on the mat in two locations: first halfway between the dog and E and then directly in front of E. When E places the food on the mat, she says "Puppy, look!" again. E rests with her hands behind her back, looks straight down, and says "okay". The subject is then allowed to approach the food and obtain the reward. If the subject does not approach the kibble in 25 s, the trial is repeated. If the dog approaches/climbs on E or sits and waits for E, the trial may be repeated until the dog retrieves the kibble immediately on its own.

After retrieving the food successfully from each location, E plays a free-form cup game with the dog to familiarize her to finding food under cups. The dog watches as E places food under a single cup, then is rewarded when she touches the cup. As needed, E can tilt or tap the cup to draw the dog's attention to the food and encourage the dog to touch the cup. These hiding instances happen in quick succession to keep the dog fully engaged. The first time, this game may take 1-2 minutes; on subsequent days, this step may be shorter.

Phase 2- one-cup alternating
E presents the reward and visibly baits a single cup placed in either the R or L position. After baiting, E kneels at the E location in resting position. The subject is then allowed to approach the cup and obtain the reward. If the subject does not approach the cup in 25s, E may call the dog over, show the kibble if necessary, and reward for touching the cup; the trial is then repeated. This phase of warm-ups familiarizes the subject with the set-up and assures that the subject is motivated to find the reward. Repeating these trials serves as a correction procedure for spontaneous side biases and ensures that subjects gain experience finding the reward in both locations. The subject is required to successfully retrieve the reward on four trials, twice on each side, to move on (within a maximum of twelve trials).

Phase 3- two-cup visible displacement
Two cups are placed in the small circles connected to the 1-m line (Figure 7). E presents the reward to the dog and visibly baits one of the two cups according to the data sheet. After baiting, E kneels in resting position. The subject is then allowed 25s to make a choice. If the subject chooses the baited cup, the subject is allowed to have the reward, E praises the subject, and the next trial is

administered. If the subject chooses the incorrect cup, E says, “Wrong” in a neutral tone, the subject is not rewarded, and the trial is repeated until the subject chooses correctly. If the subject does not choose any cup within 25s, the trial is repeated. This phase of warm-ups assures that the subject is not choosing cups randomly and is attending to the experimenter’s actions. Subjects are required to choose correctly on their first attempt in four out of five consecutive trials to advance to test trials within a maximum of 20 trials.

Refamiliarization:
If the subject does not choose on two trials in a row during phase 2, resume cup games until the dog becomes reengaged; then resume trials. If the subject does not choose on two trials in a row during phase 3, conduct 2 trials of phase 2 (one-cup alternating). If the subject does not choose on another two trials in a row (including previous phase), increase the food reward.

Abort Criteria
If there are 8 NCs total in either phase 2 or phase 3 (including refamiliarizations), abort. When 12 trials of phase 2, or 20 trials of phase 3 have been conducted (including repeated trials, but not refamiliarizations) and the pass criterion has not been met, the task is aborted.

<u>*Human Interest*</u>

Setup
E stands inside the pen until at least 10 s after H leaves the room. E then stands 20 cm from the corner of the pen nearest the door (“NE corner”). E’s left foot should be perpendicular to the corner of the pen, and E’s right foot should be facing the “N” edge of the pen, as marked; thus, camera C has a clear view of the “NE” corner of the pen.

Phase 1:
E looks at the dog, making eye contact whenever possible, and recites the following script using dog-directed speech (Ben-Aderet, Gallego-Abenza, Reby, & Mathevon, 2017), which takes approximately 30s. E uses the silent count-up timer to measure the amount of eye contact made during this time.

“Hi pup! Are you a good puppy? Yes you are. What a good puppy. Aww, look how cute you are. Look at those big eyes and floppy ears. You’re such a cute puppy! Do you like to play? Are these experiments fun? You’re coming back tomorrow to play with me! We’ll play more cup games. We’re gonna have so much fun. I can’t wait to play with you! Are you the best puppy? Yes, you are. Of course you are. That’s a good puppy!”

Phase 2:
E starts a 30-second countdown on the vibrating stopwatch, then steps into the center of the pen and faces the wall (“N” side of pen). If the dog approaches E during this time, E says “Hi pup!”, pets the dog as long as they are within reach, and is then quiet for the rest of the time. E does not move until the timer buzzes.

E repeats phases 1 and 2 in succession three times.

Day one; Session two

Cylinder Task

Setup
E kneels behind the cylinder base while H centers the dog in the starting box. The longitudinal midline of the cylinder should be aligned with the cup line, so that the center of the cylinder is 100 cm from the starting line.

Phase 1: Warm-up
Part 1: E places kibble in bowl, shows the dog and says “look”, and then places the bowl about 50 cm from dog. E says “okay” and allows the dog to eat the kibble out of the bowl. The subject has 25s to approach the bowl.

Part 2: After the dog has successfully approached the bowl at 50 cm, repeat with the bowl directly in front of the cylinder (~90 cm from subject). Once the dog has approached the bowl and cylinder, begin leading trials (part 3).

Part 3: E shows the dog the bowl with reward, H releases puppy, and E uses the bowl to guide the dog into the side of the tube. Repeat as needed until the dog succeeds at following the bowl into the tube for two trials, one baited from the right side and one from the left.

Phase 1 Abort Criteria:
During any of these three steps, if the subject does not approach, try to make the dog comfortable (petting, encouragement, play if necessary) and repeat. If they do not approach on the repeat, try an increased food reward. If they still do not approach on the next two trials, abort.

Phase 2: Familiarization (opaque cylinder)
E says ‘look’ while showing the subject the bowl and reward and baits the opaque cylinder with the right hand, from the right side. E assumes resting position and says “okay!”. H then starts a 30-s timer and releases the dog to make a choice. Subjects are permitted (and encouraged) to retrieve the reward on all trials regardless of the accuracy of their first attempt or if they chose at all. In the case that the dog touches the cylinder but never retrieves the reward on its own after 30 seconds, E shows the dog the solution and gives the dog the reward.

On every trial, E records whether the subject made a choice without touching (score = 1) or touched first (score = 0); if the subject touched, which side they touched first (front, top, back); and which side they chose (L/R from E’s perspective). A choice consists of the dog’s snout or front paw crossing the plane at either open side of the cylinder. If the dog brushes the edge of the cylinder in the process of going in the side, it is coded as a correct response, not a touch. If the dog’s ears, but nothing else, brush the cylinder at any point, this is also not considered a touch. Subjects are required to correctly (score of 1) retrieve the reward in 4 out of a sliding window of 5 trials before advancing to test trials. The front/back/top distinctions involve the direction of motion: if the dog’s snout or paw is moving downward onto the top half of the cylinder, this is a “top” touch; if the dog touches the half of the cylinder closest to H (and it is not a top), this is “front” touch (F); if the dog touches the half of the cylinder closest to E (and it is not a top), this is “back” touch (B).

Phase 3: Test phase (transparent cylinder)

The test procedure is identical to familiarization trials (phase 2) except that the apparatus is the transparent cylinder (Figure 12). As before, E codes whether the subject's first attempt to retrieve the reward through the front/back (0-incorrect) or side (1-correct) of the apparatus; if the subject touched, which side they touched first (front, top, back); which side they chose (L/R from E's perspective), regardless of whether they touched first. Again, subjects are allowed to retrieve the reward on all trials regardless of the accuracy of their first attempt. Eight trials are conducted.

Figure 12. Cylinder task. Left: correct response; Right: reversal phase

Phase 4 Detour task (transparent cylinder + transparent side):

The detour task is identical to the test trials except that after the cylinder is baited and before the pup is called, E places a clear plexiglass barrier over one side of the cylinder. The blocked side is determined by whichever side was preferred for the last three inhibitory control trials. This is also the hand with which E will bait the cylinder for this phase. As before, the dog has 30 s to make a choice (touch or around) and is rewarded regardless of accuracy (i.e. allowed to eat the treat when found, or if the dog makes an incorrect response but never solves on its own, the dog is shown the solution after 30 seconds. To do so, the experimenter calls the dog's attention to the treat. If not sufficient, the experimenter may gently guide the dog toward the open side of the cylinder to get the treat). If the dog does not make a choice within 30s, they are not rewarded and the trial is repeated.

Note: ambiguous responses can be marked as 'code' and scored later from video. In all cases, if the dog does not approach, they are not rewarded and the trial is repeated.

Phases 2-4 Abort criteria:

In phase 2, if two consecutive no choices occur, add two more leading trials and then resume. If another two consecutive no choices occur, increase the food reward. If the dog does not meet the criterion to pass phase 2 within 20 trials, abort. In phases 3 and 4, if two consecutive no choices occur, increase the food reward. If the subject does not choose on a total of 8 trials across phases, the task is aborted.

Day Two

Unsolvable Task
Setup:

The dog is held by H at a distance of 100 cm from the center of the container, and the lid of the clear container is located directly behind the container from E's perspective. E is positioned centered and behind the apparatus, directly behind the board, to allow room for the dog to walk around the container. The reward for this task is 3 pieces of soaked kibble per trial.

Familiarization:
In familiarization trials, E presents the subject with the reward, places the reward inside the clear container, and positions the lid upside down and loosely on top of the container. Four warmup trails are conducted with the following conditions (Figure 13):

1) The lid is propped on the side, not covering the container.
2) The lid covers ~½ of the container, as marked.
3) The lid covers ~ ¾ of the container, as marked.
4) The lid will (again) cover ¾ of the container, as marked.

After baiting the apparatus, E looks at the container, and says "okay!" H releases the dog and starts a 30 s timer. The dog is allowed to approach the container, displace the lid, and obtain the reward. The trial ends when the dog retrieves the kibble.

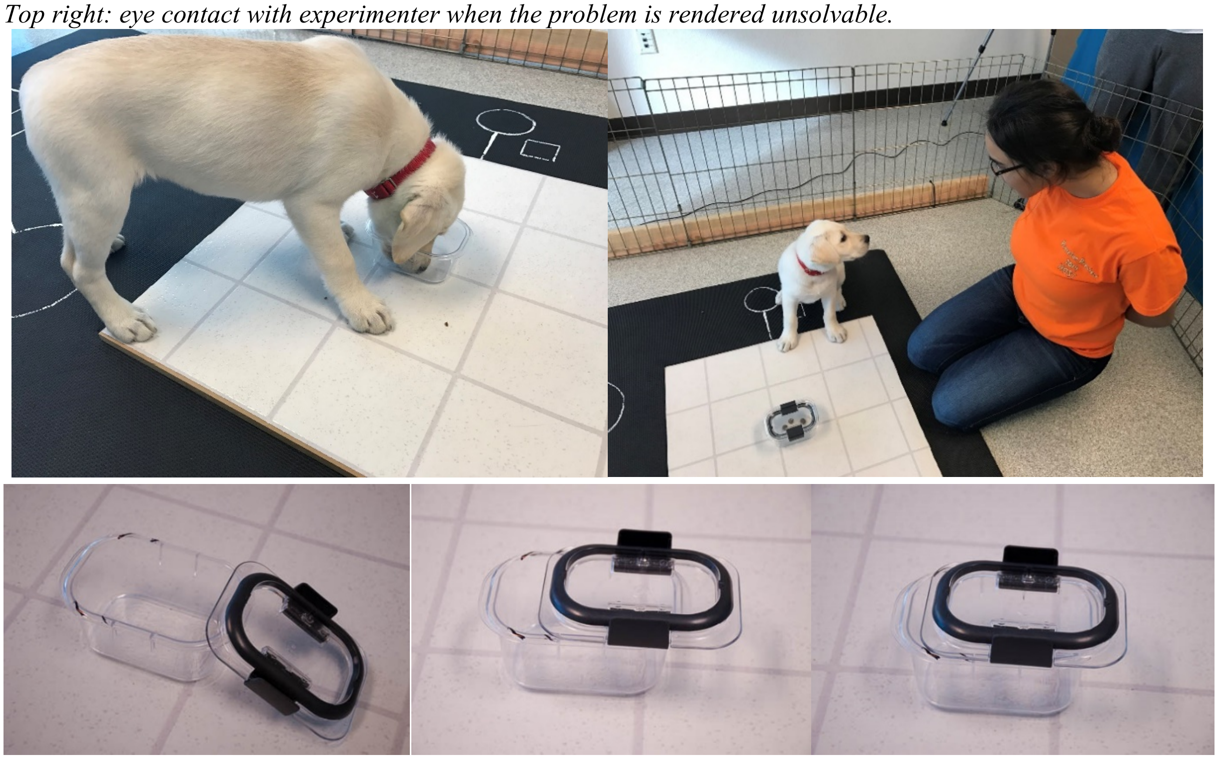
If the dog does not retrieve the reward at the end of the 30 s trial, E may call them over, draw their attention to the container, and help them solve it. That trial is then repeated. The dog is required to successfully complete 4 familiarization trials without help before advancing to the test.

Test:
Test trials are identical to familiarization trials except that E snaps the lid onto the container so that it cannot be removed. Subjects are given 30 seconds to attempt to access the reward. E remains sitting and visually follows the dog (rotating their torso if necessary) during this period. E holds a silent stopwatch behind her back and starts and stops it in count-up mode to measure the total time that the dog looks to her face. H times the 30 second trial and looks down throughout the trial, ignoring the dog. After each test trial, E praises the dog, draw their attention to the container, opens the container, and allows the dog to retrieve the contents. Four test trials are conducted.

Abort Criteria:
If the dog does not touch the container on two consecutive trials (NC), the food reward may be increased. If the subject does not finish all four trials within 12 attempts at familiarization trials, the task is aborted.

Figure 13. Unsolvable task.

Top left: Dog obtains food in familiarization trials. Bottom: Sequence of progressive covering of the food reward. Top right: eye contact with experimenter when the problem is rendered unsolvable.



<u>*Warm-ups*</u>

Methods are identical to warm-ups on Day 1.

<u>*Social Cues: General Procedures*</u>

Occluded Baiting (For marker cue, pointing, and odor control):
E places cups on the line connecting the final cup locations with the occluder ~10 cm behind, blocking the puppy's view of the cups. E presents the subject with the reward, places the treat in front of one of the cups (but near the center), lifts both cups, and places them over the treat and empty space. E then removes the occluder, placing it behind her outside of the pen, and slides the cups into the positions marked on the mat.

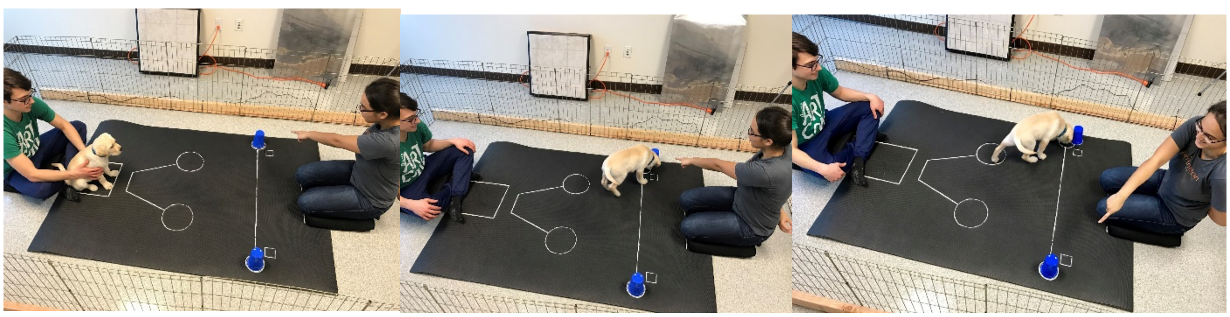
<u>Communicative Marker:</u>

The marker (yellow wooden block) is placed behind the cups while baiting. After baiting as described above, E picks up the marker in her right hand, leans forward to show it to the dog (~10 cm from the subject's nose), and says "puppy, look!". E then pulls the marker back to the central position, lifts it up and tries to make eye contact with the dog, and says "puppy, look!" before placing the marker next to the appropriate cup (in the square marked on the mat; Figure 7). E then gives an "okay!" release, signaling H to release the subject to make a choice. Until this point, H should be holding the dog but looking down so that H is unaware which side is being indicated. Once E has given the "okay!" release signal and the dog has left the starting box, H looks up; when the puppy makes a choice, H says, “choice”. E maintains resting position until the subject makes a choice or the maximum trial time of 25 s has elapsed. Twelve trials are conducted and the dependent measure for this task is the percentage of trials that a dog chooses the cup next to which E placed the wooden block.

Arm Pointing:

After baiting as above, E tries to make eye contact with the dog, says “puppy, look!”, and points toward the cup with the reward. The pointing gesture consists of the index finger of the contralateral hand extended about 20 cm from the cup and E’s head and gaze directed toward the cup (Figure 14). E then gives an “okay!” release signaling H to release the subject to make a choice. Until this point, H should be holding looking down so that H is unaware which side is being indicated. Once E has given the “okay!” release signal and the dog has left the starting box, H looks up; when the dog makes a choice, H says, “choice”. E maintains the pointing gesture and gaze until the subject makes a choice or the maximum trial time of 25s has elapsed. Twelve trials are conducted and the dependent measure is the percentage of trials that a subject chooses the cup toward which E gestured.

Figure 14. Procedure for the arm pointing task.



Odor Control

Use non-baited, clean cups for this task. No cue is administered on the eight odor control trials. After baiting as above, E kneels at the E location in resting position. E then says “okay!”, signaling H to release the subject to make a choice. Eight trials are conducted and the dependent measure for this task is the percentage of trials that a dog chooses the baited cup.

Novel Object

Positioning and Setup:
E puts some dry kibble in their pocket. H leaves the room to watch from the outside video monitor.

Figure 15. Setup for the novel object task.

Procedure: E turns on the mechanical cat and places it in the white subject square, then walks to the opposite end of the pen and positions the dog at the edge of the mat, facing the mechanical cat (Figure 15). E immediately leaves the testing room. H starts the two-minute count-down timer when the dog is released. Variables coded: latency to approach the mechanical cat, amount of time in proximity to cat, latency to contact the mechanical cat, amount of time in physical contact with cat, amount of time oriented toward cat, amount of time until dog makes first vocalization (groan/growl, whine, yelp/bark, howl), intensity of vocalization.

After two minutes, E re-enters the testing room and performs the following procedure:

1) E enters the pen and speaks positively/encouragingly/in high tones toward the mechanical cat. E turns off the mechanical cat.
2) If the dog doesn't approach within ~5 seconds, E calls the dog toward the stimulus.
3) If the dog doesn't approach within ~ 10 seconds, E calls the dog toward the stimulus, luring her with food.
4) If the dog doesn't approach within ~15 seconds, E approaches the dog halfway and presents shows food in hand.
5) When the subject touches the stimulus, E reinforces the dog's presence by giving kibble and speaking positively to the dog. Food should always be introduced to ensure that the dog is willing to eat around the novel object.
6) Lastly, E lets H back into the room and H brings the dog to the opposite end of the X-pen and holds the dog by the collar, facing away from H and toward E and the mechanical cat. E places the cat in its original location. E then squats directly behind the mechanical cat, calls the dog over, and rewards the dog upon approach (in close proximity to the novel object).

Day Three

Warm-ups

Methods are identical to warm-ups on Day 1.

Working Memory

The dog is positioned in the starting box, 100 cm away from the cup line. Two cups are positioned in the 1m cup circles. H holds a countdown stopwatch set for 25s. E places a running stopwatch on the kneeling pad to count the delay interval for each trial.

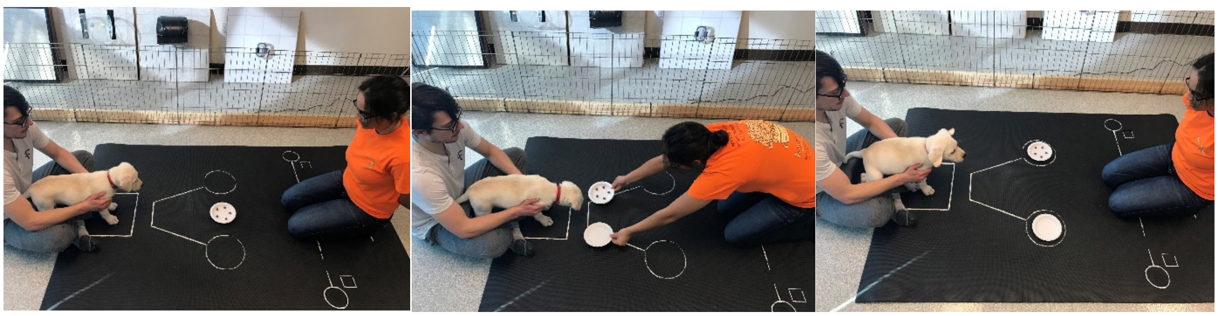
Procedure:
E presents the subject with the reward, saying "Puppy, look!" and then visibly baits one cup, lifting the cup ~10 cm off the ground. E then looks down at the running stopwatch and counts the seconds of delay for that trial. When the time is up, E says, "okay!" and H releases the puppy and starts a timer. The trial ends when the dog makes a choice or after 25s has elapsed.

There are 6 trials at each memory interval (5, 10, 15, 20 seconds). All puppies are tested at 5 and 10s delays, but then must pass at least 4 of 6 trials to proceed to the next time interval. After the first 12 trials, dogs receive a short break before both the 15 and the 20s trials (if criterion to advance to this stage was met).

Visual Discrimination

Positioning and setup:
E prepares two plates, one clean (never has food) and one that contains 5 pieces of kibble spread out on the plate (Figure 16). H holds the dog in the starting box; E is kneeling at a central position, ~75 cm directly in front of the dog.

Figure 16. Procedure for the visual discrimination task.



Familiarization:
Using her right hand, E holds the plate of kibble in front of subject, saying "Puppy, look!", places it directly in front of her, 50 cm from the starting box, and says "Okay!". H releases the dog, and the dog is allowed to retrieve at least one piece of kibble from the plate. This sequence is repeated once with E presenting the plate with her left hand as well.

Test trials:
E holds both plates in front of the dog with their farthest rim at the 10 cm line and says, "Puppy, look!". E then slides the plates along the diagonal lines painted on the mat and places them in the 50-cm circles (which are 50 cm from the dog and 50 cm apart). E then says, "Okay!" and H releases the dog to make a choice and starts the 20-second countdown timer. If the dog chooses the plate with the food, the dog is allowed to eat at least 1 piece of kibble before being returned to H. If the dog chooses the plate without the food, E says, "Wrong" in a neutral voice and returns the dog to H with no food reward. Eight trials are completed and the dependent measure is the number of trials that the dog selects the plate with the food.

Refamiliarization/Abort criteria:
If the dog makes no choice within 20 seconds on 2 consecutive trials, two refamiliarization trials are conducted to encourage participation. If a dog accumulates a total of 8 no-choice trials, the task is aborted.

Auditory Discrimination

Positioning and setup:
H holds the dog in the starting box; E kneels at a central position ~75 cm directly in front of the dog.

Figure 17. Procedure for the auditory discrimination task.

Familiarization:
E places a bowl 50 cm from the starting box, in line with the 50-cm circles. E presents the subject with dry kibble and baits the bowl. When baiting, E drops the kibble in the front of the bowl (so the dog cannot see it) and from a height of 1-2 cm (so that it makes an audible cue). E says "Okay!", H releases the dog, and the dog is allowed to retrieve kibble from the bowl. This process is repeated one time.

Test trials:
E places the bowls in the 50-cm circles. E presents the subject with dry kibble, baits and sham baits the bowls. When baiting, E drops the kibble in the front of the bowl (so the dog cannot see it) from a height of 1-2 cm (so that it makes an audible cue; Figure 17). When sham baiting, E puts her hand about 2 cm from the bottom of the bowl, pauses, and relaxes her grip to emulate the dropping motion. Bowls are baited/sham baited from right to left. The subject is given 20 s to make a choice. The dependent measure is the percentage of trials that a dog chooses the baited option.

Refamiliarization/Abort criteria:
If the dog does not choose within 20s on 2 consecutive trials, two refamiliarization trials are performed to encourage them to eat from the bowl. If there are 8 total no-choices, the task is aborted.

<u>*Odor Discrimination*</u>
<u>Set-up:</u>

Link-elbows are set up with one designated as the odor elbow and the other as the blank. A cotton ball is placed at the bottom of each elbow (wide end) to block visual access to the treats, but still allowing for the odor to dissipate (Figure 18). 10 dry kibble pieces are loaded into the odor elbow. Another cotton ball is used at the top (narrow end) of both elbows to block the top and allow a treat to be placed on top of the odor elbow for fast delivery to the dog upon a correct choice. E holds the elbows from above with palms over the upper end so that the dog cannot access the treat or the top of the elbow.

Choice criterion:
The puppy must touch the elbow or the experimenter's hand (from the wrist down) to count as a choice with its snout or front paw(s).

Familiarization:
H positions the dog in the starting box. E says "sniff it" and presents the dog with the odor elbow 10 cm in front of the start box and at the height of the puppy's nose. After 3 seconds, E pulls the elbow back to 50 cm away and holds it in place on the mat. The dog is given 20 s to touch the elbow and is rewarded immediately after doing so. On the first trial the elbow is held in the right hand and the second in the left hand. A trial is repeated if the dog does not choose.

Test:
H positions the dog in the starting box. E sits approximately 75 cm from the puppy in the center of the mat. E puts both elbows at ~60 cm from the dog, and then presents them to the dog at the edges of the 10-cm line, at the height of the dog's nose, in the following order. First the right elbow, then the left elbow, then both simultaneously, giving the subject 3s to sniff at each stage. E says "sniff it" at each stage to encourage the dog to smell the elbows. E then pulls the elbows back, following the diagonal lines on the mat and places the elbow on the ground, centered in the 50-cm circles. E says, "okay!" and H releases the dog to make a choice. The dog is given 20s to explore the elbows. Regardless of the dog's behavior, the dog is rewarded after the 20s. If the dog is not near the correct elbow, E uses that elbow to lure the dog to the location of the correct choice and rewards them there. Eight test trials are conducted.

Figure 18. Procedure for the odor discrimination task.

Scoring:
The experimenter records the elbow that the subject touched first and the elbow that the subject touched last. If E is unsure of whether the dog touched an elbow or not, the trial is marked for subsequent video coding.

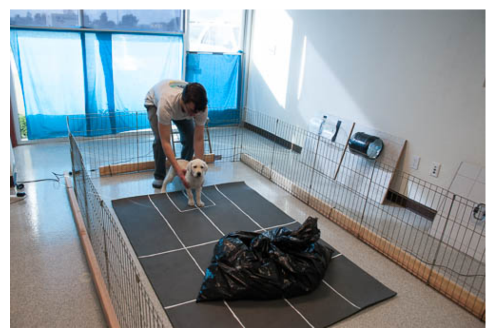
Abort Criteria:
If the subject does not choose on two consecutive trials, two refamiliarization trials are performed in which the dog is presented with a single elbow and rewarded for approaching. If the dog does not approach on either refamiliarization trial or if a dog accumulates a total of 8 no-choice trials, the task is aborted.

<u>*Surprising Events*</u>

Positioning:
At the beginning of each phase of the task, H stands 0.6 m from the back of X-pen and holds the dog by the collar, facing away from H. The dog is standing at the start of each trial.

TASK 1: SUDDEN APPEARANCE (TRASH BAG)
Phase 1
Standing outside of the X-pen, E brings a stuffed trash bag from the back of the room, stands behind H, and waits until the dog is facing forward. E then throws the trash bag such that it lands approximately 0.6 m from the opposite end of the X-pen (Figure 19). H releases the dog before the bag hits the floor and starts a timer as soon as the bag hits the floor. H then stands with feet shoulder-width apart and looks at the timer for the duration of the trial. H announces when 15 seconds have passed, at which point E enters the pen and performs the following steps:

Figure 19. Sudden appearance phase of the surprising events task.



1) If the dog is not currently exploring the object, E encourages the dog to approach by progressing through the following steps until the dog touches the object (if the dog is already exploring object, skip to step 2):
   a. E speaks positively/encouragingly/in high tones toward the object
   b. If the dog doesn't approach within ~5 seconds, E calls the dog toward the object
   c. If the dog doesn't approach within ~15 seconds, E calls the dog toward the object while presenting food
   d. If the dog doesn't approach within ~20 seconds, E approaches the dog halfway and presents the food

2) Once the dog is touching the object, E reinforces the dog's presence by providing kibble and speaking positively to the dog. Food is always introduced to ensure that the dog is willing to eat while near the startling object. After the dog is comfortable and has eaten the kibble, H approaches and leads the dog away.

If a dog does not touch the object, the trial ends 45 s after the initial event.

Phase 2
H brings the dog back to the starting position. E squats directly behind the trash bag, calls the dog, and reinforces the dog upon approach (using Step 1, a-d as outlined above). The dog's path to E is coded from video. H then holds the dog, facing away from E, while E quietly removes the trash bag from the testing area.

TASK 2: LOOMING OBJECT (UMBRELLA)
Phase 1
H brings the dog back to the starting position. E stands centered and facing the dog, directly behind the X-Pen. E holds the unopened umbrella facing into the X-pen, pointed toward the dog. E calls the dog's attention and when the dog is looking, E pushes the button to release the umbrella. E then immediately lowers the opened umbrella onto the ground, centered and between the edge of the mat and the 1-m line (Figure 20). E quickly walks away with back turned and stands facing the wall. H releases the dog when E whistles and starts the timer as soon as the umbrella touches the ground. Steps 1 and 2 described above are used to encourage and reward approach toward the stimulus.

Figure 20. Looming object component of the surprising events task.

Phase 2
H brings the dog back to the starting position. E squats directly beside the umbrella, calls the dog, and reinforces the dog upon approach (using Step 1, a-d as outlined above). The dog's path to E is coded from video. H holds the dog, facing away from E, while E quietly removes the umbrella from the testing area.

TASK 3: LOUD NOISE (METAL SHEET)
Phase 1
H brings the dog back to the starting position. E stands centered and facing the dog, standing directly behind the X-Pen and holding the metal sheet over the edge of the X-pen. E calls the dog's attention and when the dog is looking, E flicks her wrists six times to wave the metal sheet, starting toward the dog first, thereby creating a loud undulating noise. E then immediately lowers the sheet onto the ground inside of the X-pen, directly in front of the dog start position, with the far end even with the 1m line (Figure 21). E quickly walks away with back turned and stands facing the wall. H releases the dog when E whistles and starts the timer. Steps 1 and 2 described above are used to encourage and reward approach toward the stimulus.

Phase 2
H brings the dog back to the starting position. E squats directly behind the metal sheet, calls the dog, and reinforces the dog upon approach (using Step 1, a-d as outlined above). The dog's path to E is coded from video.

Figure 21. Loud noise phase of the surprising events task.

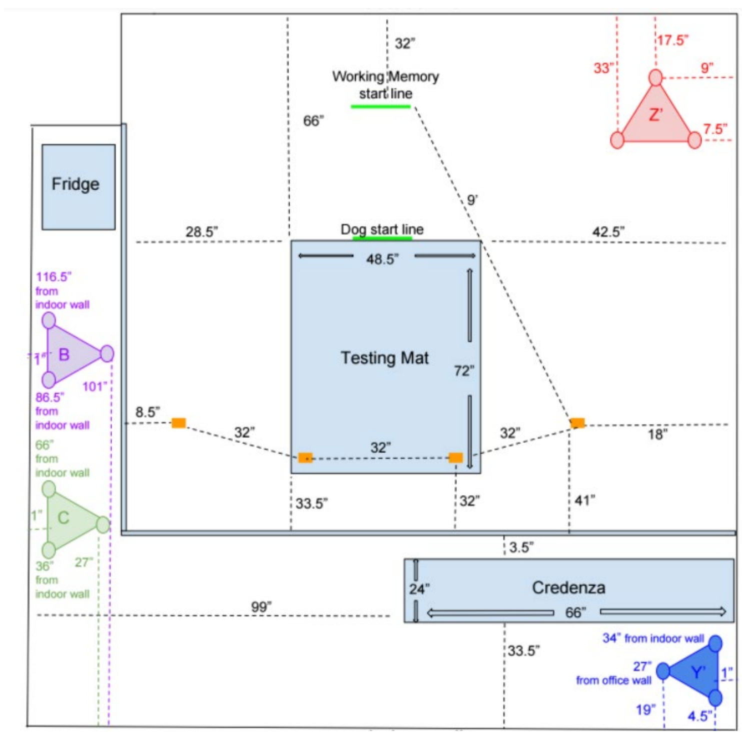
Concluding events:

Both E and H leave the pen and stand in the corner near the metal sheet. H starts the timer for one minute. E and H ignore the dog and talk in low voices, which allows the camera to record the dog's reaction to the startling object when left alone/unsupervised. After one minute, E re-enters the pen, praises the dog, and gives the dog the remainder of the food on the metal sheet.

Methodological Differences for Adult Testing

Where applicable, procedural modifications from the puppy version of the tasks described above are presented below. Figure 22 shows camera placements and the general setup for adult testing.

Figure 22. Test room setup and camera placements for experiments with adult dogs.



Laterality

Positioning and setup:
The steps are placed in the center of the mat. Before conducting the trials, H coaxes the dog onto and off of the step a few times to ensure that she is not fearful of the apparatus. H should stand behind and to the left of the subject for the first sixteen trials, then switch to behind and to the right of the subject for the last sixteen trials. The dog should be facing E, with paws aligned and sitting. If H needs to reposition the subject, they should walk the dog in a circle and have her sit again, and not touch any paws or legs. Alternatively, E can use the treat to encourage the dog to reposition the paws, again without touching any paws or legs. Up and down trials were conducted in alternating sequence, rather than performing all the up trials followed by the down trials.

Working Memory

Positioning and setup:
The dog is positioned at the start line, 9 feet away from the cup line. Four cups are positioned 32" apart from one another. H holds a countdown stopwatch set for 25 sec. E wears a watch on her wrist to track the delay interval for each trial.

Familiarization:
E visibly baits each of the four cups, starting with the outside right cup (from E's perspective). Once the dog successfully chooses that cup on its first choice, E will then bait the middle right cup (from E's perspective) until the dog chooses that cup on its first choice, and so on, ending with the dog choosing the outside left cup. The dog must successfully retrieve the treat from each of the hiding places before the test trials begin.

Abort criteria: Familiarization trials are repeated if the dog doesn't choose within 25 seconds, or if the dog chooses the incorrect cup on a given trial. If the dog reaches 8 NCs and/or incorrect attempts, the task is aborted.

Procedure:
E presents the subject with the reward, saying "Puppy, look!" and then visibly baits one cup, lifting the cup ~10 cm off the ground. E then returns to the back of the X-pen and stands centered, looking down at the wristwatch on her wrist and counting the seconds of delay for that trial. When the time is up, E says, "okay!" and H releases the dog and starts their timer. The trial ends when the dog makes a choice or after 25 s has elapsed. There are 6 trials at each memory interval (20 and 40 seconds). All dogs are tested at all intervals. If a dog doesn't choose, E records "NC" and moves on to the next trial (i.e., the trial is not repeated). The dog is given a brief break in between the 20 s and 40 s trials.

Independent Problem Solving

Methods:
Problem Solving A: Dog Magic
STAGE 1. E baits every other well (4 total) along the outside of Dog Magic apparatus while the dog watches. H stands with back to door and holds dog on leash until E says, "Okay" and then H

releases the leash. The dog is then allowed to find 4 uncovered treats to familiarize itself with the apparatus. Encouragement guidelines: If dog stops interacting with the apparatus and leaves the immediate area on which the apparatus is sitting, E repeats "Find the treats" and points toward the apparatus at 10 second intervals. Praise guidelines: Each time the dog finds and consumes a treat, E will say "good dog." E times how long the dog takes to complete the trial. Do not progress to Stage 2 until Stage 1 is complete.

STAGE 2. E baits one outside well and covers it with a bone while dog watches. H holds dog on leash until E says, "Okay, find the treats" and then H releases the leash. The dog is then allowed 45 seconds to solve the problem and retrieve the reward. The encouragement and praise guidelines are the same as in Stage 1. E verbally encourages the dog [at the same rate as in stage 1], but never physically interacts with the dog or the apparatus. If the dog doesn't recover the treat after 45 seconds, E shows the dog the hidden reward. E times how long the dog takes to complete each trial. Do not progress to Stage 3 until the dog successfully completes two trials.

STAGE 3. E baits every other well (4 total) along the outside of Dog Magic apparatus while the dog watches. E then places bones (9 total) over all wells. H holds the dog on leash until E says, "Okay, find the treats" and then H releases the leash. The dog is then allowed to find the treats. The encouragement and praise guidelines are the same as in Stage 1. E times how long the dog takes to complete the task and records the number of correctly overturned wells and the number of incorrectly overturned wells. The trial is capped at 2 minutes. At the end of 2 minutes, if the dog has not completed the task, E shows and gives the dog the rest of the treats.

Problem Solving B: Dog Tornado
STAGE 1. When dog has completed Stage 3 (by successfully uncovering all 4 treats or once 2 minutes has elapsed, whichever comes first), set out the Dog Tornado apparatus. E baits one well along the outside of Dog Tornado apparatus while the dog watches. H holds the dog on leash until E says, "Okay, find the treats" and then H releases the leash. The dog is then allowed to find 1 uncovered treat to familiarize itself with the apparatus. E times how long the dog takes to complete the trial. Do not progress to Stage 2 until Stage 1 is complete.

STAGE 2. E baits one outside well and twists the apparatus so the well is hidden while the dog watches. H holds the dog on leash until E says, "Okay, find the treats" and then H releases the leash. The dog is then allowed 45 seconds to solve the problem and retrieve the reward. E verbally encourages the dog [at the same rate as in stage 1], but never physically interacts with the dog or the apparatus. If dog doesn't recover the treat after 45 seconds, E shows the dog the reward. E times how long the dog takes to complete each trial. Do not progress to Stage 3 until the dog successfully completes two trials.

STAGE 3. E baits one outside well and covers it with a bone while the dog watches. H holds the dog on leash until E says, "Okay, find the treats" and then H releases the leash. The dog is then allowed 45 seconds to solve the problem and retrieve the reward. E verbally encourages the dog [at the same rate as in stage 1], but never physically interacts with the dog or the box. If the dog doesn't recover the treat after 45 seconds, E shows the dog the reward. E times how long the dog takes to complete each trial. Do not progress to Stage 4a until the dog successfully completes two trials.

STAGE 4a. E baits one outside well, tapes the treat into place, twists the apparatus, and covers the closest empty well with a bone while the dog watches. H holds the dog on leash until E says, "Okay, find the treats" and then H releases the leash. The dog is then allowed to find the treat. E times how long the dog takes to complete task. The trial is capped at 2 minutes.

STAGE 4b. A dog progresses to Stage 4b only if successful at Stage 4a. E baits one outside well, twists the apparatus, and covers the farthest empty well with a bone while the dog watches. All other elements are the same as in Stage 4a.

STAGE 4c. A dog progresses to Stage 4c only if successful at Stage 4b. E baits one outside well, twists the apparatus, and covers both empty wells with two bones while the dog watches. All other elements are the same as in Stage 4a.

Problem Solving A: Dog Magic abort criteria: If the dog has not solved stage 2 (removing the bone from over top of single well within 45 seconds) successfully after 5 trials, the task is aborted.

Problem Solving B: Dog Tornado abort criteria: If the dog does not eat the treat during stage 1 within 45 seconds, the task is aborted. If the dog has not solved stage 2 (twisting the bone to reveal baited well within 45 seconds) successfully after 5 trials the task is aborted. If the dog has not solved stage 3 (removing bone from over top of single well within 45 seconds) successfully after 5 trials the task is aborted.