

**Dog cognitive development: a longitudinal study across the first 2 years of life**

Emily E. Bray<sup>a, b, \*</sup>, Margaret E. Gruen<sup>c</sup>, Gitanjali E. Gnanadesikan<sup>a, d</sup>, Daniel J. Horschler<sup>a, d</sup>,  
Kerinne M. Levy<sup>b</sup>, Brenda S. Kennedy<sup>b</sup>, Brian A. Hare<sup>e, f</sup>, and Evan L. MacLean<sup>a, d, g, h</sup>

<sup>a</sup>Arizona Canine Cognition Center, School of Anthropology, University of Arizona, Tucson, AZ 85719, U.S.A.

<sup>b</sup>Canine Companions for Independence, National Headquarters, Santa Rosa, CA 95407, U.S.A.

<sup>c</sup>Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607, U.S.A.

<sup>d</sup>Cognitive Science Program, University of Arizona, Tucson, AZ 85719, U.S.A.

<sup>e</sup>Duke Canine Cognition Center, Department of Evolutionary Anthropology, Duke University, Durham, NC 27708, U.S.A.

<sup>f</sup>Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, U.S.A

<sup>g</sup>Department of Psychology, University of Arizona, Tucson, AZ 85719, U.S.A.

<sup>h</sup>College of Veterinary Medicine, University of Arizona, Tucson, AZ 85719, U.S.A.

\*Corresponding author email: ebray@email.arizona.edu

ORCID: 0000-0002-3230-0636

Conflict of Interest: The authors declare that they have no conflicts of interest.

47 **Abstract**

48           While our understanding of adult dog cognition has grown considerably over the past 20  
49 years, relatively little is known about the ontogeny of dog cognition. To assess the development  
50 and longitudinal stability of cognitive traits in dogs, we administered a battery of tasks to 160  
51 candidate assistance dogs at two timepoints. The tasks were designed to measure diverse aspects  
52 of cognition, ranging from executive function (e.g., inhibitory control, reversal learning, memory)  
53 to sensory discrimination (e.g., vision, audition, olfaction) to social interaction with humans.  
54 Subjects first participated as 8-to-10-week old puppies, and then were retested on the same tasks  
55 at ~21 months of age. With few exceptions, task performance improved with age, with the largest  
56 effects observed for measures of executive function and social gaze. Results also indicated that  
57 individual differences were both early emerging and enduring; for example, social attention to  
58 humans, use of human communicative signals, independent persistence at a problem, odor  
59 discrimination, and inhibitory control all exhibited moderate levels of rank-order stability between  
60 the two  
61 timepoints. Using multiple regression, we found that young adult performance on many cognitive  
62 tasks could be predicted from a set of cognitive measures collected in early development. Our  
63 findings contribute to knowledge about changes in dog cognition across early development as well  
64 as the origins and developmental stability of individual differences.

65

66

67

68

69 **Keywords:** Assistance dog, Behavior, Cognition, Development, Longitudinal, Individual  
70 differences

## 71 **Introduction**

72         Ontogeny provides an important window into the nature of any complex trait, as  
73 emphasized by Tinbergen (1963) when he dedicated one of his four questions to development.  
74 However, studies of animal cognition often focus on the cognitive phenotypes of adult animals,  
75 with limited knowledge about their developmental bases (Rosati, Wobber, Hughes, & Santos,  
76 2014). Domestic dogs present rich opportunities for comparative studies of cognitive development  
77 given their ubiquity in human societies, employment in diverse working dog applications, and  
78 highly variable environments throughout development. Although dog cognition has been  
79 extensively studied throughout the last two decades, the majority of studies have focused on adult  
80 animals, often aiming to characterize species-typical performance on cognitive tasks, with less  
81 emphasis on individual differences or their development (Arden, Bensky, & Adams, 2016;  
82 Bensky, Gosling, & Sinn, 2013; MacLean, Herrmann, Suchindran, & Hare, 2017).

83         One common approach for inferring how variation in developmental experiences affects  
84 adult cognition compares adult phenotypes of dogs who have (presumably) experienced different  
85 environmental conditions during development. These studies have evaluated both social cognition  
86 (i.e. communicative cues, measures of social gaze) and nonsocial cognition (i.e. inhibitory control,  
87 independent problem solving, causal inferences), and compared the performance of pet dogs living  
88 in homes to other groups with less human contact, including shelter (Duranton & Gaunet, 2016),  
89 kennel-housed (Turcsán et al., 2020), purpose-bred research (Lazarowski & Dorman, 2015), pack-  
90 raised (Lampe, Bräuer, Kaminski, & Virányi, 2017), and free-ranging (Brubaker, Dasgupta,  
91 Bhattacharjee, Bhadra, & Udell, 2017) dogs. Some of these studies find no difference in the

92 cognitive performance of dogs with different rearing histories (Brubaker et al., 2017; Fagnani,  
93 Barrera, Carballo, & Bentosela, 2016; Lampe et al., 2017), but others support the conclusion that  
94 pet dogs are more adept at using social cues and more persistent at solving problems than dogs  
95 living or reared in kennel environments (Duranton & Gaunet, 2016; Lazarowski & Dorman, 2015).  
96 While informative with respect to adult phenotypes, these studies lack measures of early-life  
97 cognition and thus cannot directly characterize cognitive development.

98         Other methods for studying changes across the lifespan involve cross-sectional or  
99 longitudinal research. These approaches have been common in studies of canine temperament and  
100 personality (e.g., Goddard & Beilharz, 1984b; Head et al., 1997; Jones & Gosling, 2005; Marshall-  
101 Pescini, Virányi, Kubinyi, & Range, 2017; Riemer, Müller, Virányi, Huber, & Range, 2016; Scott  
102 & Fuller, 1965; Sforzini et al., 2009; Starling, Branson, Thomson, & McGreevy, 2013; Wallis,  
103 Szabó, & Kubinyi, 2020), but are less commonly employed in cognitive research. Bensky et al.  
104 (2013) reported that of 222 canine cognition studies published through 2012, only 12.6% employed  
105 a cross-sectional or longitudinal design. Many of those studies investigated the development of  
106 socio-cognitive skills, as have several papers published subsequently (e.g., Bhattacharjee et al.,  
107 2017; Gácsi, Györi, et al., 2009; Lazarowski, Rogers, Waggoner, & Katz, 2019; Lazarowski,  
108 Strassberg, Waggoner, & Katz, 2019; Rossano, Nitzschner, & Tomasello, 2014; Zaine,  
109 Domeniconi, & Wynne, 2015). In contrast, cross-sectional or longitudinal studies of nonsocial  
110 cognition have tended to focus on older dogs and cognitive decline associated with aging (e.g.,  
111 Christie et al., 2005; Head, 2013; Milgram et al., 2002; Milgram, Head, Weiner, & Thomas, 1994;  
112 Piotti et al., 2018; Tapp et al., 2003; Wallis et al., 2014; Watowich et al., 2020). Of the few that  
113 have investigated the early development of nonsocial cognition, one recent study in a population  
114 of working dogs (tested at 3, 6, and 11 months of age) found that inhibitory control, attention, and

115 spatial cognition all improved with age (Lazarowski, Krichbaum, Waggoner, & Katz, 2020). These  
116 findings suggest that important cognitive changes are occurring over early development and  
117 highlight the need for further research on these processes.

118         Lastly, one important question about cognitive development—which can only be addressed  
119 through a longitudinal design—concerns the stability of individual differences across time. Again,  
120 this question has been addressed in numerous studies of canine personality and temperament (e.g.,  
121 Fratkin, Sinn, Patall, & Gosling, 2013; Goddard & Beilharz, 1984a, 1986; Harvey, Craigon,  
122 Blythe, England, & Asher, 2017; Harvey et al., 2016; Riemer et al., 2016; Svartberg, Tapper,  
123 Temrin, Radesäter, & Thorman, 2005; Tomkins, Thomson, & McGreevy, 2010; Wilsson &  
124 Sundgren, 1998). However, very few cognitive studies have implemented similar longitudinal  
125 approaches. Riemer et al. (2014) reported that cognitive impulsivity in 13 dogs, quantified via  
126 performance on a delay of gratification task as well as owner assessment on a questionnaire, was  
127 highly stable across a 6-year timespan; on the other hand, they found that a measure of motor  
128 impulsivity was not correlated across the 2 timepoints. Gácsi et al. (2009) conducted a pointing  
129 task where a subset of subjects participated at 2 timepoints: 12 puppies were retested 1 to 12 weeks  
130 after the initial test, another 12 puppies were retested 8 to 18 months later (as adults), and 12 adults  
131 were retested 1 week to 6 months later. In that study, there were no effects of age on performance  
132 and the small sample size precluded a powerful assessment of the developmental stability of  
133 individual differences.

134         The literature reviewed above has either indirectly assessed cognitive development or  
135 focused on the development of a limited subset of cognitive traits. We sought to fill these gaps in  
136 our understanding by conducting the first large-scale longitudinal study of canine social and  
137 nonsocial cognition, allowing us to assess both the early development and stability of a wide range

138 of cognitive traits. Here, we report the results of this multi-year study in which we tracked  
139 individual differences in cognition in a population of 160 candidate assistance dogs. All dogs were  
140 whelped and weaned in Northern California and participated in the Dog Cognitive Development  
141 Battery (DCDB; Bray et al., 2020) at ~9 weeks of age. This test battery – derived from the dog  
142 cognition test battery for adult dogs (MacLean et al., 2017) – was designed with the goal of  
143 measuring individual differences across a diverse range of cognitive processes. After completing  
144 testing, dogs were then raised in the homes of volunteers throughout the western United States  
145 until ~21 months of age, at which point they returned to professional training centers and  
146 completed the test battery a second time. We investigated changes in task performance across  
147 ontogeny—i.e. how performance changed from ~9 weeks of age to young adulthood—as well as  
148 the stability of individual differences across time—i.e. the extent to which task performance as a  
149 puppy predicted subsequent performance in young adulthood.

150       Regarding changes in skills related to age, we hypothesized that for tasks where there was  
151 an objectively correct response (e.g., object-choice tasks), cognitive performance would improve  
152 from ~9 weeks to ~21 months. We also expected that at least a subset of traits would exhibit  
153 consistent individual differences across time, although given the lack of previous research in this  
154 area, we had no *a priori* hypotheses regarding the relative stability of different traits.

## 155 **General Methods**

### 156 **Subjects**

157       All subjects were recruited from Canine Companions for Independence (Santa Rosa, CA,  
158 USA), a non-profit assistance dog organization in the United States. Canine Companions granted  
159 informed consent to all aspects of the study. All testing procedures were reviewed and adhered to

160 regulations set forth by the Institutional Animal Care and Use Committee at the University of  
161 Arizona (IACUC No. 16-175).

162 We aimed to test all subjects on the same tasks at two different timepoints: first in early  
163 development and later in young adulthood. To this end, we tested 168 puppies (97 females and 71  
164 males) from February to July of 2017 at approximately 9 weeks of age (mean = 9.20 weeks, range  
165 7.86 – 10.43 weeks). Our sample included 122 Labrador x golden crosses, 40 Labrador retrievers,  
166 and 6 golden retrievers from 65 different litters (Bray et al., 2020). After their initial testing, these  
167 dogs were raised by volunteer puppy raisers throughout the western United States for ~18 months  
168 before returning to Canine Companions for Independence for professional training. Of the original  
169 168 puppies, we were able to test all but eight individuals as young adults ( $n = 5$  released for  
170 medical reasons prior to turn-in,  $n = 1$  released for behavioral reasons prior to turn-in,  $n = 2$  did  
171 not meet participation criteria at turn-in). Thus, our final sample consisted of 160 dogs (93 females  
172 and 67 males). These dogs were tested for a second time from January 2018 to April 2019 when  
173 they were just under two years old (mean = 1.79 years, range 0.99 to 2.01 years), within a month  
174 of each dog returning to Canine Companions' Northwest (Santa Rosa, CA) or Southwest  
175 (Oceanside, CA) regional campuses for professional training. The dogs who participated in both  
176 rounds of testing included 118 Labrador x golden crosses, 37 Labrador retrievers, and 5 golden  
177 retrievers.

## 178 **Procedure**

### 179 *Dog Cognitive Development Battery (DCDB)*

180 The DCDB (Bray et al., 2020) consists of a series of tasks designed to assess aspects of  
181 perception, executive function, communication, social motivation, and temperament (Table S1).  
182 All dogs completed this battery once in early development (~9 weeks of age) and again in young

183 adulthood (~1.8 years of age). The general methods used with puppies and adults were identical  
184 except for one task that was only presented to adults, as well as minor procedural differences  
185 required to obtain age-appropriate measures, described below (e.g., retention intervals on memory  
186 tasks).

### 187 **Implementation with puppies**

188 Puppies completed testing in a dedicated 19.5' x 14' room at Canine Companions for  
189 Independence's Canine Early Development Center. Each subject completed one ~45-minute  
190 session per day over three consecutive days (Fig. 1a). All of the cognitive tasks in the DCDB are  
191 briefly described below; detailed experimental methods and video examples are provided in Bray  
192 et al. (2020), as well as in the Supplementary Material. Because the temperament tasks (i.e. novel  
193 object and surprising events) were not the focus of the current study, detailed methods for these  
194 tasks are not presented here but are available in a separate manuscript (Bray et al., 2020). Although  
195 we primarily categorized laterality as a temperament task (Batt, Batt, Baguley, & McGreevy,  
196 2009), we include it here as there is some evidence across species that behavioral lateralization  
197 (i.e. handedness) is associated with cognition (Bibost & Brown, 2014; Güntürkün, Ströckens, &  
198 Ocklenburg, 2020; Magat & Brown, 2009; although see Whiteside et al., 2020). However, given  
199 that measures of behavioral lateralization have been shown to vary based on the task in humans  
200 (Annett, 1994) and dogs (e.g., Batt, Batt, Baguley, & McGreevy, 2008; Tomkins et al., 2010;  
201 Wells, 2003), it is a limitation of the current study that we only include only one measure of  
202 laterality. Nonetheless, the measure of laterality that we included has been associated with  
203 measures of both structural and sensory laterality and is among the most widely-used assessments  
204 of laterality in dogs (Tomkins, Williams, Thomson, & McGreevy, 2012).



205 For the sake of comparison between individuals, all subjects completed the tasks in the  
206 same order (Bray, Sammel, Seyfarth, Serpell, & Cheney, 2017; MacLean et al., 2017). For tasks  
207 requiring a choice (e.g., hiding-finding warm-ups, cylinder, gesture use, working memory, and  
208 perceptual discriminations), if a puppy did not choose within the predetermined number of seconds  
209 or if there was an experimenter error, that trial was repeated. If the subject's lack of interest in  
210 participation continued, we employed a standardized protocol for trying to re-engage and re-  
211 familiarize the puppy with the task and if necessary gave the puppy a break before returning to the  
212 task (see the Supplementary Material and Bray et al. (2020) for specific refamiliarization and abort  
213 criteria for each task). On infrequent occasions, when those attempts were ineffective, and as  
214 indicated by the predetermined abort criteria, the task was discontinued for that puppy  
215 (Supplementary Table S2).

216 ***Vision pretest*** – This test ensured that puppies were capable of tracking visual stimuli at the typical  
217 distances used in subsequent tasks (based on Ollivier, Plummer, & Barrie, 2007). At a distance of  
218 100 cm in front of the puppy, a cotton ball was dropped vertically and flicked across the ground in  
219 full view of the subject. Subjects were required to follow the motion of the cotton ball on at least  
220 three trials to advance to subsequent tasks. All puppies tested met this criterion.

221 ***Retrieval (Fig. 1a task 1, Fig. 1btask 1)*** – This task measured the puppy's willingness to  
222 cooperatively engage in fetch with a human partner (based on Bray, Sammel, Seyfarth, et al., 2017;  
223 Slabbert & Odendaal, 1999; Wilsson & Sundgren, 1997). Following a 1 min familiarization period  
224 (see Supplementary Material), the experimenter threw a small ball for the puppy and vocally  
225 encouraged the dog to bring the ball back to her. For each of the two 1 min test trials, the puppy  
226 received a score based on the following scoring system: (1) did not interact with the ball at all, (2)  
227 only chased the ball, (3) also picked the ball up in the mouth, (4) returned the ball to the

228 experimenter one to two times, or (5) returned the ball to the experimenter three or more times.  
229 The dependent measures were the puppy's average score across two trials and a tally of the total  
230 number of times that the puppy returned the ball to the experimenter.

231 ***Laterality (Fig. 1a task 2, Fig. 1b task 2)*** – This task indexed behavioral measures of laterality  
232 by tracking the puppy's paw preference when stepping onto and off of a platform (based on  
233 Tomkins et al., 2010), which is believed to reflect lateralization in the brain and has been  
234 previously linked to temperamental reactivity in adult dogs (Branson & Rogers, 2006). Following  
235 a brief introduction to the platform (see Supplementary Material), puppies were held by the handler  
236 and then called by the experimenter to step onto the platform across a series of 15 trials, and then  
237 off the platform across a series of 15 trials. The forelimb used to initiate this motion on each trial  
238 was recorded and subsequently used to compute a laterality index.

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

248 ***Human interest (Fig. 1a task 3, Fig. 1b task 3)*** – This task measured the puppy's motivation to  
249 attend to a human who spoke to the puppy using dog-directed speech (Ben-Aderet, Gallego-  
250 Abenza, Reby, & Mathevon, 2017; Gergely, Faragó, Galambos, & Topál, 2017). The experimenter

251 stood outside the testing pen, looked at the puppy, and recited a predetermined script with a  
252 fluctuating, high-pitched intonation (Ben-Aderet et al., 2017). After each recitation, the  
253 experimenter entered the pen and petted the puppy if approached. This procedure was repeated  
254 three times. The duration of the puppy's gaze to the human's face during the recitation of the script  
255 and the duration of interaction with the experimenter during play breaks was recorded across trials.

256 *Cylinder inhibitory control and Cylinder reversal learning (Fig. 1a task 4, Fig. 1b task 4)* – The  
257 first part of this task measured the puppy's inhibitory control (i.e. the ability to suppress a prepotent  
258 response in favor of a choice that would ultimately be more productive) by requiring the puppy to  
259 detour to the reward location, thereby placing distance between herself and a visible reward (based  
260 on Bray, MacLean, & Hare, 2014; MacLean et al., 2014). This task is often employed in the canine  
261 literature as a measure of motor inhibition (Brucks, Marshall-Pescini, Wallis, Huber, & Range,  
262 2017; Fagnani et al., 2016; Marshall-Pescini, Virányi, & Range, 2015; but for critiques see  
263 Kabadyi, Bobrowicz, & Osvath, 2018; van Horik et al., 2018; van Horik, Beardsworth, Laker,  
264 Whiteside, & Madden, 2020). The second part of this task measured the puppy's ability to exhibit  
265 cognitive flexibility when the demands of the task changed, and the puppy's previously preferred  
266 solution was no longer available. Puppies first participated in familiarization trials by walking  
267 around the front of an opaque cylinder to retrieve a reward from one of the side openings. In a)  
268 inhibitory control test trials, a transparent cylinder was used such that subjects had to resist the  
269 prepotent response to move directly towards the visible food, instead avoiding the transparent  
270 obstacle. Eight trials were conducted. The dependent measures were the proportion of trials that  
271 the puppy successfully retrieved the food from either side opening of the cylinder, without first  
272 touching the exterior of the apparatus, and the average latency to obtain the reward. In b) reversal  
273 learning test trials, the puppy's preferred side entrance to the cylinder was obstructed by a

274 transparent plastic barrier and subjects were required to switch their response, detouring to the  
275 other opening of the apparatus to retrieve the treat. Eight test trials were conducted. The dependent  
276 measure was the proportion of trials that puppies performed the correct detour response without  
277 first touching the barrier or exterior of the cylinder. The side of the apparatus that the subject first  
278 approached (i.e. open or blocked) and the average latency to obtain the reward were also recorded  
279 as measures of response flexibility.

280 ***Unsolvable (Fig. 1a task 5, Fig. 1b task 9)*** – This task measured the puppy’s inclination to persist  
281 at an unsolvable task independently versus looking at a nearby human experimenter, potentially to  
282 solicit help (based on Miklósi et al., 2003; for alternative explanations of what this task measures  
283 see Lazzaroni et al., 2020). The puppy was familiarized with displacing the lid from a transparent  
284 container to obtain a visible food reward inside. Then, across four 30 s test trials, the lid to the  
285 container was affixed, and the dependent measures were the duration of time gazing at the  
286 experimenter’s face and duration of time physically manipulating the container.

287 ***Gesture use*** – The experimenter showed the puppy a food reward, then used a foam board occluder  
288 to block the puppy’s view while placing the reward inside one of two possible hiding locations.  
289 The experimenter then removed the occluder, provided one of three cues (communicative marker,  
290 arm pointing, odor control; see below) before subjects could search and recorded the subject’s first  
291 choice.

292 ***Communicative marker (Fig. 1a task 6, Fig. 1b task 5)*** – This task measured the puppy’s  
293 ability to use an arbitrary marker, used in a communicative manner, to find a hidden reward  
294 (based on Agnetta, Hare, & Tomasello, 2000; Riedel, Buttelmann, Call, & Tomasello,  
295 2006). The experimenter ostensibly (preceded by verbally addressing and making eye

296 contact with the puppy) placed a small yellow block that the puppy had never seen before  
297 next to the baited location. Twelve test trials were conducted.

298 ***Arm pointing (Fig. 1a task 7, Fig. 1b task 6)*** – This task measured the puppy’s ability to  
299 use an arm-pointing gesture to find a hidden reward (based on Hare, Call, & Tomasello,  
300 1998; Miklósi, Polgárdi, Topál, & Csányi, 1998). The experimenter ostensibly (preceded  
301 by verbally addressing and making eye contact with the puppy) pointed with the  
302 contralateral arm, index finger extended, and gazed towards the baited location until the  
303 trial ended. Twelve test trials were conducted.

304 ***Odor control (Fig. 1a task 8, Fig. 1b task 7)*** – This task acted as a control to ensure that  
305 puppies’ performance on the gesture use tasks could not be attributed to olfactory cues or  
306 unintentional cuing by the experimenter (based on Bräuer, Kaminski, Riedel, Call, &  
307 Tomasello, 2006; Hare, Brown, Williamson, & Tomasello, 2002; Miklósi et al., 1998).  
308 After baiting, the experimenter remained still and did not provide any social information.  
309 Eight test trials were conducted.

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

313 ***Working memory (Fig. 1a task 10, Fig. 1b task 10)*** – This task measured the puppy’s ability to  
314 recall the location of a hidden treat after temporal delays of various lengths (based on Doré, Fiset,  
315 Goulet, Dumas, & Gagnon, 1996; Fiset, Beaulieu, & Landry, 2003). It was identical to hiding-  
316 finding warm-ups with the exception that we imposed a delay before the subject was allowed to  
317 search, which increased across blocks of six trials each (5 s, 10 s, 15 s, 20 s). Only individuals who

318 chose correctly on at least four of six trials at 10 s moved on to delays of 15 s, and only those who  
319 chose correctly on at least four of six trials at 15 s moved on to delays of 20 s. The proportion of  
320 trials that the subject first searched in the baited location was used as the dependent measure.

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

324 ***Visual discrimination (Fig. 1a task 11, Fig. 1b task 12)*** – This task measured the puppy's  
325 ability to choose a baited location versus an unbaited location based on visual cues. One  
326 plate contained five pieces of visible kibble and the other was empty. The experimenter  
327 presented the plates directly in front of the puppy before pulling them backward to 50 cm  
328 in front of the puppy, equidistant to the left and right sides. Eight test trials were conducted.  
329 The proportion of trials that the puppy first approached the baited plate (i.e. the puppy's  
330 snout extended over the plate) was used as the dependent measure.

331 ***Auditory discrimination (Fig. 1a task 12, Fig. 1b task 13)*** – This task measured the  
332 puppy's ability to choose a baited location versus an unbaited location based on auditory  
333 cues (based on Bräuer et al., 2006). Two metal bowls, placed approximately 50 cm away  
334 from the puppy, were used as the hiding locations. The experimenter sequentially placed  
335 her hand into each container, audibly dropping the food into only one of the containers.  
336 Eight test trials were conducted. The dependent measure was the proportion of trials that  
337 the subject's first search was to the baited location.

338 ***Odor discrimination (Fig. 1a task 13, Fig. 1b task 14)*** – This task measured the puppy's  
339 ability to choose a baited location versus an unbaited location based on olfactory cues. Two

340 sections of rubber tubing with a 90° bend (“elbows”) were presented, one of which  
341 contained 10 pieces of dry kibble. The ends of the elbows were filled with cotton to prevent  
342 the contents from being visible or audible. The experimenter allowed the subject to sniff  
343 the opening of each elbow individually for 3 s, and then the elbows were presented side by  
344 side for an additional 3 s before being pulled backward 50 cm in front of the puppy,  
345 equidistant to the left and right sides. Puppies were released and allowed to move freely  
346 for 20 s. On each trial, the first and last elbow that the subject approached was recorded, as  
347 well as the cumulative time spent within a marked 10 cm radius around the elbows. Eight  
348 test trials were conducted. The dependent measures were the proportion of trials that the  
349 subject’s first and last responses were directed to the baited location, as well as the  
350 proportion of time that the puppy spent within each of the marked radii around the elbows.

351 This task-by-task description of the DCDB is reprinted from Animal Behaviour, 166, Bray  
352 EE, Gruen ME, Gnanadesikan GE, Horschler DJ, Levy KM, Kennedy BS, Hare BA, MacLean EL,  
353 Cognitive characteristics of 8- to 10-week-old assistance dog puppies, 193-206, Copyright (2020),  
354 with permission from Elsevier.

### 355 **Implementation with young adults**

356 The adult version of the DCDB was identical to the battery implemented with puppies, apart  
357 from the minor changes described below and detailed in the Supplementary Material (Fig. 1b).  
358 Adult dogs were tested at Canine Companions campuses in either Santa Rosa, CA ( $n = 92$ ) or  
359 Oceanside, CA ( $n = 68$ ) within approximately one month of returning for professional training  
360 (minimum = 7 days, maximum = 52 days, average = 23 days). All subjects had previously  
361 completed the DCDB as puppies. On the rare occasions where adults lacked motivation to

362 participate on a given task, the same protocols and abort criteria used for the puppies were applied  
363 (Supplementary Table S3).

364 1) *Removal of vision pretest*

365 a. Because the adult dogs had been selected to enter professional training, their eyes  
366 were thoroughly assessed by a veterinary ophthalmologist and their vision was  
367 deemed adequate. We therefore removed the vision pre-test for adult subjects.

368 2) *Addition of a physical problem-solving task linked to success in guide dogs* (Fig. 1b  
369 task 11)

370 a. We added an independent problem-solving task that has been associated with  
371 training outcomes in a population of guide dogs (Bray, Sammel, Cheney, Serpell,  
372 & Seyfarth, 2017). In this task, the dog was required to complete a series of  
373 familiarization trials to ensure they were sufficiently motivated and able to meet  
374 the physical (motoric) demands of the task. In subsequent test trials, the dog was  
375 required to watch and remember where a treat was hidden within several possible  
376 locations on an apparatus, and then manipulate the apparatus appropriately to  
377 successfully retrieve the food (see Supplementary Material). The dependent  
378 measures for this task are shown in Supplementary Table S4. Although dogs did  
379 not participate in this task as puppies, we include it here as an outcome variable for  
380 Lasso regression models predicting adult performance as a function of multiple  
381 phenotypic measures collected from puppies (see below).

382 3) *Minor age-appropriate modifications*

383 a. We increased the difficulty of the working memory task. Adult subjects were  
384 required to remember where a treat was hidden while accounting for more possible



385           hiding locations (four vs. two) across longer delays (up to 40 s). Given the long trial  
386           times, this was the only task in the battery where if a dog failed to make a choice  
387           within the allotted 30 seconds, the next trial was administered rather than repeating  
388           the trial.

389           b. Where needed, larger stimuli were used (e.g., the ball during the retrieval task, the  
390           platform during the laterality task, the container during the unsolvable task).

391           c. Puppies were not yet leash-trained, and thus were held in place at the start line by  
392           their collar or shoulders prior to the experimenter giving the release command. In  
393           contrast, all adults were leash-trained and thus were held in place at the start line  
394           by a short traffic lead that could subsequently be dropped upon the experimenter  
395           giving the release command. For the laterality task, the handler stood to the side of  
396           the dog (versus straddling the dog) to allow the dog a full range of motion. The side  
397           that the handler stood on was counterbalanced across trials. Therefore, all adults  
398           participated in 16 (versus 15) “up” and “down” trials so that the handler could stand  
399           an equal number of times on the left and right sides.

400           d. After piloting the odor discrimination task with adult dogs, we determined that the  
401           three x 3 s presentations of elbows for dogs to sniff before each test trial (as  
402           implemented in the puppy battery) was frustrating and aversive to many subjects,  
403           and that adults were sufficiently motivated to participate in test trials after a single  
404           initial 3 s presentation of both elbows. Thus, the task was modified such that adults  
405           were given the final 3 s presentation only (in which both elbows were  
406           simultaneously presented) at the start of each of the six test trials.

407 e. With the puppies, the battery consisted of three sessions over three days (Fig. 1a).  
408 Due to the increased attention span and food motivation of adult dogs, the adult  
409 version of the DCDB was implemented in two sessions lasting around 1 to 1.5 hours  
410 each, either on the same day with a break in between or across two different days  
411 (Fig. 1b).

### 412 **Scoring and statistical analysis**

413 All statistical analyses were carried out in R v.3.6.0 (R Development Core Team, 2016).  
414 Most behavioral variables were scored live, but all tasks were videorecorded for reliability  
415 assessment and additional analyses. The following measures were later coded from video: select  
416 variables from cylinder (latency during inhibitory control and reversal learning trials and first side  
417 correct during reversal learning trials), unsolvable (average time manipulating object), and odor  
418 discrimination (time at right and left elbow, from which the variables time in proximity to baited  
419 option and time in proximity to nonbaited option were subsequently calculated).

420 For the live-coded data, independent coders scored from video all trials for 20% of  
421 randomly selected subjects, and interrater reliability was calculated using Pearson correlation for  
422 continuous variables and Cohen's Kappa for categorical variables. For the measures that were not  
423 possible to score live, two coders independently scored data from video. The primary coder scored  
424 all data for analysis, and a reliability coder scored all trials for 20% of randomly selected subjects.

425 All measures were reliable for data collected at both timepoints. For the puppy measures,  
426 there was high inter-rater agreement on both live-coded (Cohen's kappa: mean = 0.94; Pearson's  
427  $r$ : mean = 0.96) and video-coded (Cohen's kappa: mean = 0.93; Pearson's  $r$ : mean = 0.97)  
428 measures. Raw reliability statistics for the puppy data are reported in Bray et al. (2020). Reliability  
429 was also excellent for adult measures with high inter-rater agreement on live-coded (Cohen's

430 kappa: mean = 0.96; Pearson's  $r$ : mean = 0.97) and video-coded (Cohen's kappa: mean = 0.99;  
431 Pearson's  $r$ : mean = 0.93) measures. Raw reliability statistics for testing at this second time point  
432 are presented in Supplementary Tables S5 and S6.

433 To assess changes across ontogeny, we conducted paired sample t-tests on DCDB measures  
434 collected from dogs at ~9 weeks of age, and again in young adulthood (~18-24 months). To  
435 quantify the effect of age at testing on each trait, we calculated Cohen's  $d$  using the R package  
436 "effsize" (Torchiano, 2020), with the 'paired' argument set to true and the 'within' argument set  
437 to false. To assess longitudinal stability of traits measured by the DCDB, we used two analytical  
438 approaches. First, following traditional approaches for assessing the consistency of individual  
439 differences across time, we performed rank-order stability analyses by assessing the Spearman  
440 correlation between phenotypes at the 2 timepoints (Caspi, Roberts, & Shiner, 2005). To test the  
441 directional prediction that phenotypes at timepoint 1 would be positively related to phenotypes at  
442 timepoint 2, we used a directional hypothesis testing framework, following the conventions ( $\delta =$   
443 0.01,  $Y = 0.04$ ) recommended by Rice & Gaines (1994). Second, we fit Bayesian linear mixed-  
444 models (Stan Development Team, 2018) to assess the relationship between phenotypes at  
445 timepoint 1 and timepoint 2, controlling for breed, sex, (adult) testing location, and relatedness  
446 between individuals, using the "rutiltimeflutre" and "rstan" R packages (Flutre, 2020; Stan  
447 Development Team, 2018). For these models, we converted phenotypic measures to z-scores to  
448 facilitate interpretation and comparison of beta coefficients. Models were fit using four  
449 independent MCMC chains with weakly informative Cauchy priors for the beta coefficients  
450 relating phenotypes at timepoint 1 to phenotypes at timepoint 2. Each chain employed a 5,000-  
451 iteration burn-in period followed by 15,000 iterations of sampling, using a 25-sample thinning  
452 interval. The results across chains were merged to obtain the final posterior distributions.

453           In addition to modeling the stability of individual DCDB measures across time, we also  
454 conducted exploratory analyses using multiple phenotypic measures collected from puppies as  
455 predictors of each single adult measure. Thus, rather than focusing on stability in a single given  
456 measure across time, these analyses investigated whether any of the phenotypic measures collected  
457 from puppies predicted variance in adult phenotypes. For tasks with multiple dependent measures,  
458 we first used principal components analysis (PCA) to reduce the number of variables associated  
459 with each task (performed separately for puppies and adults). The collective set of variables  
460 associated with each task was converted to z-scores prior to parallel analysis (Horn, 1965) using  
461 the R package “psych” (Revelle, 2019) to determine the number of components to retain. If  
462 parallel analysis suggested retention of zero components for a task, we retained the original  
463 dependent measures without performing PCA. In all other cases we performed PCA and extracted  
464 the recommended number of components using a varimax rotation to facilitate interpretation of  
465 component loadings. The one exception was the laterality task, for which we also retained the  
466 original dependent measures without performing PCA, due to evidence in the literature that both  
467 bias strength (Barnard, Wells, Hepper, & Milligan, 2017; Branson & Rogers, 2006) and  
468 directionality (Tomkins, Thomson, & McGreevy, 2012; Wells, Hepper, Milligan, & Barnard,  
469 2017) can be important, depending on the associations being tested.

470           We next used Lasso regression, implemented in the R package “glmnet” (Friedman, Hastie,  
471 & Tibshirani, 2009), for variable selection given the high ratio of variables to observations in our  
472 dataset. Lasso regression imposes a penalty ( $\lambda$ ) on the beta coefficients, favoring sparse models by  
473 shrinking many beta coefficients to zero (Friedman, Hastie, & Tibshirani, 2010). To determine the  
474 optimal value for  $\lambda$  in these analyses, we used leave-one-out cross validation to obtain the  $\lambda$  value  
475 that yielded the minimum cross-validated error. Lasso models were fit using 14 DCDB measures

476 from puppies (Table 1) as well as breed, sex, coat color, and adult testing location as predictors for  
 477 each adult outcome measure (Table 2). Finally, we fit unrestricted linear models using the subset  
 478 of variables with non-zero beta coefficients in the Lasso models (Hastie, Tibshirani, Friedman, &  
 479 Franklin, 2005; Hastie, Tibshirani, & Wainwright, 2015).

480 A summary of the primary analyses and their aims is provided in Table 3.

481 *Table 1. Puppy DCDB predictor variables used in Lasso regressions.*

<i>task</i>	<i>measure</i>	<i>type of measure</i>	<i>variables into measure</i>	<i>proportion variance explained</i>
Retrieval	Task engagement	Principal Component	average score (+), tally (+)	95%
Laterality	Laterality index	Z-scored variable	$\frac{[R - L]}{[R + L]} \times 100$	NA
Laterality	Bias strength	Z-scored variable	Absolute value of laterality index	NA
Human interest	Attentive	Principal Component	Average look time (+), average interaction time (+)	58%
Cylinder	Inhibitory control	Principal Component	Inhibitory control score (+)	24%
Cylinder	Reversal learning	Principal Component	Reversal score (+), first side correct (reversal trials) (+)	30%
Cylinder	Quick to solve	Principal Component	Latency (reversal trials) (-), latency (inhibitory control trials) (-)	30%
Unsolvable	Independent	Principal Component	Average time manipulating box (+), average time looking at human (-)	64%
Arm pointing	% trials correct	Z-scored variable	Arm pointing	NA
Communicative marker	% trials correct	Z-scored variable	Communicative marker	NA
Memory	% correct across delays	Principal Component	Short delays (+), long delays (+)	71%
Visual discrimination	% trials correct	Z-scored variable	Visual discrimination	NA
Auditory discrimination	% trials correct	Z-scored variable	Auditory discrimination	NA
Odor discrimination	Time spent near correct location	Principal Component	First choice (+), final choice (+), time in proximity to baited option (+), time in proximity to nonbaited option (-)	49%

483

484 *Table 2. Adult DCDB outcome measures used in Lasso regressions.*

<i>task</i>	<i>measure</i>	<i>type of measure</i>	<i>variables into measure</i>	<i>proportion variance explained</i>
Retrieval	Task engagement	Principal Component	Average score (+), tally (+)	95%
Laterality	Laterality index	Z-scored variable	$\frac{[R - L]}{[R + L]} \times 100$	NA
Laterality	Bias strength	Z-scored variable	Absolute value of laterality index	NA
Human interest	Average look time	Z-scored variable	Average look time	NA
Human interest	Average interaction time	Z-scored variable	Average interaction time	NA
Cylinder	Inhibitory control	Principal Component	Inhibitory control score (+), latency (inhibitory control trials) (-)	28%
Cylinder	Reversal learning	Principal Component	Reversal score (+), latency (reversal trials) (-), first side correct (reversal trials) (+)	38%
Unsolvable	Independent	Principal Component	Average time manipulating box (+), average time looking at human (-)	82%
Arm pointing	% trials correct	Z-scored variable	Arm pointing	NA
Communicative marker	% trials correct	Z-scored variable	Communicative marker	NA
Memory	% correct across delays	Principal Component	Short delays (+), long delays (+)	62%
Visual discrimination	% trials correct	Z-scored variable	Visual discrimination	NA
Auditory discrimination	% trials correct	Z-scored variable	Auditory discrimination	NA
Odor discrimination	Time spent near correct location	Principal Component	First choice (+), final choice (+), time in proximity to baited option (+), time in proximity to nonbaited option (-)	62%
Problem solving A	Success	Principal Component	Correct attempts (+), incorrect attempts (+), latency to solve (-), gaze (-), engage (+)	68%

485

486

487

488

489

490 *Table 3. Summary of primary analyses.*

<i>Statistical Method</i>	<i>Question assessed</i>	<i>Results</i>
Paired-sample t-tests	mean changes across development	Table 4
Spearman correlations	rank-order stability of individual differences	Figure 2
Bayesian linear mixed models	longitudinal trait stability controlling for covariates and genetic relatedness	Figure 2
Lasso regression	associations between puppy phenotypes and adult phenotypes (multiple regression with variable selection)	Table 5

491

492 **Results and Discussion**493 *Development of cognitive traits: changes across ontogeny*

494 The results from paired-sample t-tests are shown in Table 4, along with the puppy and adult  
495 means and effect sizes with 95% confidence intervals. Performance on the majority of measures  
496 improved with age (Table 4). Some of the largest increases were in measures of executive function:  
497 adults substantially outperformed puppies in the cylinder trials involving inhibitory control ( $d =$   
498  $0.78$ ) and reversal learning (reversal score  $d = 0.92$ ; first side correct  $d = 0.80$ ). There was also  
499 large changes in some of the behaviors involving communication and social motivation: the  
500 amount of looking to a human in various contexts dramatically increased from early ontogeny to  
501 young adulthood (human interest: avg look time  $d = 1.14$ ; unsolvable: avg time looking at human  
502  $d = 0.62$ ), and adults were more skilled at using the marker cue ( $d = 0.66$ ).

503 However, there were also a handful of tasks in which dogs performed no differently in  
504 early ontogeny compared to young adulthood (Table 4). In the laterality task, the mean laterality  
505 bias (which incorporates directionality and is reflected by the laterality index) did not differ  
506 between the 2 timepoints, but the strength of this bias significantly increased with age.

507 Additionally, no differences were observed between the two age groups on the visual  
 508 discrimination task or on two measures from the odor discrimination task—final choice and time  
 509 spent in proximity to the baited option—suggesting that the requisite sensory and discriminative  
 510 capabilities reached adult-like states within the first two months of life. Also, in two social  
 511 referencing tasks, while time spent looking to the experimenter’s face significantly increased from  
 512 early ontogeny to young adulthood (human interest mean  $\pm$  SD<sub>puppy</sub> = 6.44  $\pm$  4.00; human interest  
 513 mean  $\pm$  SD<sub>adult</sub> = 15.55  $\pm$  7.31;  $t_{141} = 13.62$ ,  $p < 0.001$ ; unsolvable mean  $\pm$  SD<sub>puppy</sub> = 0.98  $\pm$  1.03;  
 514 unsolvable mean  $\pm$  SD<sub>adult</sub> = 3.33  $\pm$  3.62;  $t_{158} = 7.79$ ,  $p < 0.001$ ), there were no significant  
 515 differences in time spent near the experimenter during the play break period of the human interest  
 516 task or time spent manipulating the container in the unsolvable task. In the cylinder reversal  
 517 learning trials, adults showed significant improvement on two measures (reversal score and first  
 518 side correct), but there was no difference between age groups in the latency to solve the reversal  
 519 trials. Lastly, there was also no difference between age groups in performance on the odor control  
 520 trials, with both groups performing at chance expectation (mean  $\pm$  SD<sub>puppy</sub> = 49.92  $\pm$  15.64; mean  
 521  $\pm$  SD<sub>adult</sub> = 50.70  $\pm$  15.88;  $t_{154} = 0.42$ ,  $p = 0.68$ ).

522 *Table 4. Within-subject age differences by task. Medium to large effect sizes are indicated in*  
 523 *bold.*

<i>variable</i>	<i>units</i>	<i>puppy mean</i>	<i>adult mean</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>effect size (Cohen's d)</i>	<i>lower 95% CI</i>	<i>upper 95% CI</i>
retrieval: average score	rating system (see text)	3.3	3.7	2.96	159	< 0.01	0.23	0.08	0.39
retrieval: tally	No. tallies	3.01	6.17	5.46	159	< 0.01	0.43	0.27	0.59
laterality: laterality index	$\frac{[R - L]}{[R + L]} \times 100$	-7.71	-10.7	-0.48	159	0.63	-0.04	-0.19	0.12



laterality: bias strength	absolute value of laterality index	40.88	59.06	5.40	159	< 0.01	0.43	0.26	0.59
human interest: avg look time	No. of seconds	6.44	15.55	13.62	141	< 0.01	<b>1.14</b>	0.93	1.35
human interest: avg interaction time	No. of seconds	18.57	19.32	1.20	141	0.23	0.10	-0.06	0.27
cylinder: inhibitory control score	% trials correct	51.19	75.94	9.75	157	< 0.01	<b>0.78</b>	0.60	0.95
cylinder: latency (inhibitory control trials)	No. of seconds	3.99	3.35	-2.54	157	0.01	-0.20	-0.36	-0.04
cylinder: reversal score	% trials correct	29.7	59.59	11.53	155	< 0.01	<b>0.92</b>	0.73	1.11
cylinder: first side correct (reversal trials)	% trials correct	23.01	57	9.99	155	< 0.01	<b>0.80</b>	0.62	0.98
cylinder: latency (reversal trials)	No. of seconds	6.65	6.25	-0.96	155	0.34	-0.08	-0.23	0.08
unsolvable task: avg time looking at human	No. of seconds	0.98	3.3	7.79	158	< 0.01	<b>0.62</b>	0.45	0.79
unsolvable task: avg time manipulating box	No. of seconds	12.78	13.51	1.40	158	0.16	0.11	-0.05	0.27
arm pointing	% trials correct	69.5	77.14	3.75	155	< 0.01	0.30	0.14	0.46
communicative marker	% trials correct	76.11	89.32	8.25	157	< 0.01	<b>0.66</b>	0.48	0.83
memory (short delays)	% trials correct	63.22	73.09	2.41	57	0.02	0.32	0.05	0.58
visual discrimination	% trials correct	91.33	90.08	-0.90	159	0.37	-0.07	-0.23	0.08
auditory discrimination	% trials correct	59.2	65.47	2.87	158	< 0.01	0.23	0.07	0.39
odor discrimination: first choice	% trials correct	53.31	60.94	3.35	155	< 0.01	0.27	0.11	0.43
odor discrimination: final choice	% trials correct	72.22	71.77	-0.25	155	0.81	-0.02	-0.18	0.14
odor discrimination: time in proximity to baited option	No. of seconds	61.32	64.33	1.14	155	0.26	0.09	-0.07	0.25

odor discrimination: time in proximity to non-baited option	No. of seconds	18.35	20.7	1.92	155	0.06	0.15	-0.00	0.31
---	-------------------	-------	------	------	-----	------	------	-------	------

---

524

525 *Stability of cognitive traits: early life predictors of adult phenotypes*526 *Longitudinal stability.* We first assessed longitudinal stability of cognitive traits by

527 analyzing the one-to-one correspondence between measures at the two developmental timepoints.

528 The main results from these analyses are shown in Fig. 2.

529 Across traits, rank-order stability analyses yielded Spearman correlations ranging from -

530 0.07 to 0.19 (Fig. 2). Sixteen of the correlation coefficients were positive, and only six were

531 negative. A one-sample t-test on the rank-order correlation coefficients indicated that the mean

532 correlation coefficient was significantly greater than zero (mean  $\pm$  SE = 0.06  $\pm$  0.02,  $t_{21} = 3.42$ ,  $p$ 533  $< 0.01$ ), suggesting overall positive relationships between the same traits measured at the 2

534 timepoints. Five individual traits had significant rank-order correlations, all of which were positive

535 (Fig. 3). These traits included a measure of attention to a human face during communication

536 (human interest: average looking time,  $r_s = 0.19$ ,  $p = 0.02$ ), independent persistence during an537 unsolvable task (unsolvable task: average time manipulating box,  $r_s = 0.17$ ,  $p = 0.02$ ), performance538 in the reversal phase of the cylinder task (cylinder: reversal score,  $r_s = 0.16$ ,  $p = 0.03$ ), accuracy in539 detecting a baited location via odor (odor discrimination: final choice,  $r_s = 0.15$ ,  $p = 0.04$ ), and540 sensitivity to human communication using an arbitrary cue (communicative marker,  $r_s = 0.15$ ,  $p =$ 

541 0.04).

542 The results from Bayesian linear mixed models controlling for breed, sex, testing location,

543 and relatedness between individuals supported similar conclusions. The mean beta coefficients

544 from the posterior distributions for puppy phenotype as a predictor of adult phenotype ranged

545 between -0.09 and 0.16 (Fig. 2). Fifteen of these beta coefficients were positive, and seven were  
546 negative. A one-sample t-test indicated that the mean of these beta coefficients was significantly  
547 greater than 0 (mean  $\pm$  SE =  $0.05 \pm 0.02$ ,  $t_{21} = 3.19$ ,  $p < 0.01$ ). For four measures with positive  
548 associations between the puppy and adult phenotypes (cylinder: reversal score, odor  
549 discrimination: final choice, communicative marker, and cylinder: latency (reversal trials)), the  
550 90% credible interval (Kruschke, 2014) for the beta coefficient did not contain zero, indicating a  
551 credible positive relationship between these puppy and adult phenotypes. Therefore, while  
552 individual phenotypes changed substantially across development, for a subset of traits involving  
553 interest in and communication with humans, as well as persistence, reversal learning, and odor  
554 discrimination, individual differences in puppies were modestly predictive of adult phenotypes  
555 (Fig. 2).

556 *Lasso regression models.* We next used a multiple regression approach to identify a set of  
557 phenotypic measures collected from puppies that were associated with adult performance on  
558 DCDB tasks. For this analysis we excluded all variables from Problem Solving B, because a large  
559 percentage (17%) of subjects were unable to pass familiarization trials and thus did not have data  
560 for test trials. For the remaining tasks, all of which had many fewer missing observations (mean  $\pm$   
561  $SE_{\text{puppy}} = 1.88 \pm 0.01\%$ ; mean  $\pm$   $SE_{\text{adult}} = 1.45 \pm 0.01\%$ ), missing data were imputed using a k-  
562 nearest neighbors approach.

563 The results of Lasso regressions using puppy phenotypic measures to predict adult  
564 phenotypes are shown in Table 5. As described above, the predictor variables for these models  
565 were obtained by performing PCA on each puppy task with multiple measures as well as  
566 converting all remaining measures to z-scores (Table 1), and the outcome variables for these  
567 models were obtained by following this same procedure for the adult tasks (Table 2). For 9 of 15

568 models, all beta coefficients were shrunk to zero, leaving an intercept only model (data not shown).  
 569 However, models for the remaining six adult measures all retained some puppy phenotypic  
 570 measures as predictor variables. Unconstrained linear models using these predictor variables  
 571 revealed several plausible associations. First, adult performance in the human interest task was  
 572 positively predicted by puppy performance on the arm pointing, retrieval, and short-term memory  
 573 tasks. Given that the outcome and two of the three predictor variables all involve communication  
 574 and dyadic interaction with humans, this result may capture a developmentally stable suite of traits  
 575 involving cooperative interaction with humans. Second, for the cylinder task, adult performance  
 576 on the reversal learning trials was positively predicted by puppy performance on the inhibitory  
 577 control trials and negatively predicted by a slow latency to solve both the inhibitory control and  
 578 reversal learning trials as a puppy. Third, adult performance on the inhibitory control trials of the  
 579 cylinder task was positively predicted by puppy performance on the short-term memory task and  
 580 by a right-paw preference on the laterality task. Given the positive relationships between variables  
 581 involving impulse control and working memory, the latter two models may reflect developmental  
 582 stability in traits related to executive function. Furthermore, the association between early paw  
 583 preference and later impulsivity is intriguing and parallels findings in the human literature, in  
 584 which left-handed people are more likely to show impairments in impulsivity and hyperactivity  
 585 (e.g., Reid & Norvilitis, 2000; Schmidt, Carvaho, & Simoes, 2017; Shaw & Brown, 1991; Simoes,  
 586 Carvalho, & Schmidt, 2017).

587 *Table 5. Linear models predicting adult phenotypes from puppy phenotypes. Significant predictors*  
 588 *are indicated in bold.*

<i>adult outcome</i>	<i>puppy predictor</i>	$\beta$	<i>t</i>	<i>p</i>
----------------------	------------------------	---------	----------	----------

	communicative marker	0.0906	1.1517	0.2512
human interest: avg look time $r^2 = 0.17$	<b>arm pointing</b>	<b>0.2216</b>	<b>2.8286</b>	<b>0.0053</b>
	<b>retrieval (high engagement)</b>	<b>0.2027</b>	<b>2.6959</b>	<b>0.0078</b>
	<b>memory: all delays</b>	<b>0.1511</b>	<b>2.0418</b>	<b>0.0429</b>
auditory discrimination $r^2 = 0.11$	arm pointing	0.1268	1.6908	0.0929
	visual discrimination	0.1401	1.5459	0.1242
	human interest: attentive	0.1241	1.7182	0.0878
	<b>cylinder: reversal learning</b>	<b>0.2446</b>	<b>3.4026</b>	<b>0.0000</b>
	memory: all delays	-0.0496	-0.686	0.4937
problem solving A: success $r^2 = 0.30$	laterality: laterality index	0.1208	1.7173	0.088
	visual discrimination	-0.1064	-1.2165	0.2257
	<b>auditory discrimination</b>	<b>0.1431</b>	<b>1.9663</b>	<b>0.0511</b>
	<b>retrieval (high engagement)</b>	<b>0.2485</b>	<b>3.5372</b>	<b>0.0000</b>
	cylinder: reversal learning	-0.0656	-0.9478	0.3448
cylinder: reversal learning $r^2 = 0.11$	<b>cylinder: quick to solve</b>	<b>0.1557</b>	<b>2.0561</b>	<b>0.0414</b>
	<b>cylinder: inhibitory control</b>	<b>0.1759</b>	<b>2.3023</b>	<b>0.0226</b>
	unsolvable: independent	0.12	1.5727	0.1178
cylinder: inhibitory control $r^2 = 0.07$	<b>laterality: laterality index</b>	<b>0.1842</b>	<b>2.3505</b>	<b>0.0200</b>
	<b>memory: all delays</b>	<b>0.1809</b>	<b>2.346</b>	<b>0.0202</b>

---

memory: all delays $r^2 = 0.04$	laterality: strength	bias	0.1435	1.7640	0.0797
	communicative marker		0.1239	1.4890	0.1385

---

589

590 **General Discussion**

591 We tested a sample of candidate assistance dogs ( $n = 160$ ) at 2 timepoints on a series of  
592 tasks that measured diverse aspects of cognition to explore the early development and stability of  
593 individual differences in cognitive traits. Over the developmental period that we investigated  
594 (approximately 9 weeks to 21 months), performance on most cognitive tests exhibited age-related  
595 improvement. For example, performance on tasks involving executive function (e.g., memory,  
596 impulse control, reversal learning) and social motivation (e.g., retrieval, looking toward humans,  
597 using communicative cues) all improved with age, a finding that is largely consistent with the few  
598 previously published studies exploring the early development of dog cognition (Dorey, Udell, &  
599 Wynne, 2010; Lazarowski et al., 2020; Passalacqua et al., 2011; Watowich et al., 2020; Wynne,  
600 Udell, & Lord, 2008; but see also Hare et al., 2002; Riedel et al., 2008 & Gàcsi et al., 2009).  
601 Particularly large effects were observed on inhibitory control and reversal learning trials of the  
602 cylinder task and looking time during the human interest task. On the other hand, there were a  
603 handful of cognitive measures on which puppy performance was indistinguishable from adult  
604 performance, including persistence at an unsolvable task, time interacting with the human during  
605 play breaks in the human interest task, direction of paw preference, performance on the visual  
606 discrimination task, time spent near the baited option during the odor discrimination task, and  
607 performance on the odor control task (at chance for both age groups).

608           The findings from this study contribute to the debate in the literature about the evolution  
609 of social skills in dogs. In line with several prior studies (Agnetta et al., 2000; Gácsi, Györi, et al.,  
610 2009; Gácsi, Kara, et al., 2009; Hare et al., 2002; Kaminski, Schulz, & Tomasello, 2012; Riedel  
611 et al., 2008; Rossano et al., 2014; Virányi et al., 2008), our data suggest that dogs are attuned to  
612 human communicative gestures from early in development, prior to extensive exposure to humans,  
613 as they reliably follow both conventional and novel gestures to find a food reward at above chance  
614 levels (while failing to do so in the absence of any social cues). We also find that these abilities  
615 improve over time, with adult dogs exhibiting small (arm pointing: *Cohen's d* = 0.30) to medium  
616 (communicative marker: *Cohen's d* = 0.66) increases in gesture following ability. The current  
617 study design precludes us from determining the extent to which this improvement results from  
618 simple maturational processes versus specific environmental experiences. Finally, we also find  
619 evidence that individual differences on these measures exhibit some stability across development.  
620 Therefore, while absolute ability tends to increase across ontogeny, relative ability between  
621 individual dogs is correlated in early development and young adulthood.

622           The longitudinal nature of our study allowed us to investigate the stability of individual  
623 differences across development. We found that some cognitive measures – including propensity  
624 to retrieve, auditory discrimination, and interaction time during human interest – showed marked  
625 interindividual change over development. Conversely, several other cognitive measures –  
626 including social gaze toward humans, use of human communicative signals, independent  
627 persistence at a problem, odor discrimination, and inhibitory control – exhibit significant rank-  
628 order stability across development, suggesting an early emerging and relatively stable pattern of  
629 individual differences (Fig. 3). Further, we found evidence that for some traits – including human  
630 interest, auditory discrimination, independent problem solving, inhibitory control, and reversal

631 learning – adult phenotypes can be predicted by leveraging multiple predictor variables collected  
632 from puppies.

633         Performance on the reversal learning trials of the cylinder task, which requires inhibition  
634 of a previously rewarded behavior and is therefore a measure of impulsivity (Izquierdo & Jentsch,  
635 2012), had one of the highest rank-order correlations between early development and young  
636 adulthood. Furthermore, in our Lasso regression models, adult reversal learning scores were  
637 predicted by multiple measures related to inhibitory control in early development. These results  
638 are consistent with reports in the human literature. For example, the Dunedin Multidisciplinary  
639 Health and Development Study, a longitudinal study that followed a cohort of 1,000 children in  
640 New Zealand, also found that self-control, measured via questions pertaining to impulsivity,  
641 hyperactivity, and inattention, was moderately stable from childhood to young adulthood ( $r = 0.30$ ,  
642  $p \leq 0.001$ ) (Moffitt et al., 2011).

643         Although this study was conducted in a population of prospective working dogs, if these  
644 findings hold across other populations, they have the potential to inform human-animal  
645 interactions by facilitating the prediction of adult dog characteristics. Conversely, this research  
646 also indicates that there are certain traits for which such prediction would likely be futile. Past  
647 studies have documented how features of the dog, including behavior, can affect the human-animal  
648 relationship (Curb, Abramson, Grice, & Kennison, 2013; Duffy, Kruger, & Serpell, 2014; Hsu &  
649 Serpell, 2003). Thus, on a practical level, having an objective tool through which to screen  
650 behavior at the age around which adoption usually occurs, coupled with the emerging knowledge  
651 of which behaviors are stable over time, could be extremely useful in enabling responsible and  
652 successful pet adoptions.



653           From an applied perspective, understanding the developmental course of cognition and  
654 temperament will be crucial to more efficient selection of assistance dogs. Studies are beginning  
655 to document not only the functional impacts (e.g. increasing independence; Hall, MacMichael,  
656 Turner, & Mills, 2017) but also the psychosocial benefits (O'Haire & Rodriguez, 2018; Rodriguez,  
657 Bryce, Granger, & O'Haire, 2018; Rodriguez, LaFollette, Hediger, Ogata, & O'Haire, 2020) that  
658 these highly-trained dogs provide to their handlers. However, most candidate assistance dogs, even  
659 among populations specifically bred for these roles, are ultimately released from training programs  
660 due to behaviors incompatible with their working role (Bray et al., 2019). Thus, our findings speak  
661 to the possibility of screening for relevant characteristics early in a dog's life, and we identify a  
662 subset of traits for which this approach may be most profitably employed. It is particularly  
663 promising that adult skills in the realms of dyadic communication and executive function can be  
664 predicted by puppy performance on these tasks since they are likely to directly impact the human-  
665 animal bond, a key component of any assistance dog team (Burrows, Adams, & Spiers, 2008;  
666 LaFollette, Rodriguez, Ogata, & O'Haire, 2019). In fact, previous research has linked individual  
667 differences on similar measures to working dog success (Bray, Sammel, Cheney, et al., 2017;  
668 Lazarowski et al., 2020; MacLean & Hare, 2018). Furthermore, individual differences on these  
669 measures presumably arise in part due to genetic mechanisms, and future work will benefit from  
670 characterizing the heritability and molecular bases of these traits. Thus, in addition to contributing  
671 to our knowledge of the ontogeny of canine cognition, our findings may also help guide the  
672 processes of screening, selecting, and breeding working dogs in the future.

673

674

675 **Acknowledgments**

676 We thank Ben Allen, Erika Albrecht, Kacie Bauer, Ashtyn Bernard, Amelia Byrony, James  
677 Brooks, Molly Byrne, Meg Callahan, Alexzia Clark, Victoria Coon, Elizabeth Carranza, Averill  
678 Cantwell, Mary Chiang, Amanda Chira, Allison Doty, Laura Douglas, Alex Evans, Erin Hardin,  
679 Victoria Holden, Emily Humphrey, Julia Kemper, Jennifer Geary, Kyla Guinon, Lindsey Lang,  
680 Jessica Nelson, Camden Olson, Facundo Ortega, Gianna Ossello, Alessandra Ostheimer, Amber  
681 Robello, Kerri Rodriguez, Camila Risueno-Pena, Ashley Ryan, Holland Smith, Paige Smith, Lily  
682 Tees, and Mia Wesselkamper for help with data collection and video coding. We thank the staff  
683 of Canine Companions for Independence and their dedicated volunteer breeder caretakers for  
684 accommodating six months of research with their assistance dog puppies at the Canine Early  
685 Development Center, and over a year of research with their assistance dogs in professional training  
686 at two of their regional campuses. This research was supported in part by grants from the Office  
687 of Naval Research (ONR N00014-17-1-2380 and N00014-20-1-2545 to EM and ONR N00014-  
688 16-1-2682 to BH), the American Kennel Club Canine Health Foundation (Grant No. 02518 to EB  
689 and EM), and the Eunice Kennedy Shriver National Institute of Child Health & Human  
690 Development of the National Institutes of Health (Award No. R01HD097732 to BH). The contents  
691 of this publication are solely the responsibility of the authors and do not necessarily represent the  
692 views of the Foundation or the National Institutes of Health.

693 **Author contributions:** The study was designed by all authors. The analysis was primarily  
694 designed and conducted by EEB and ELM. The paper was written primarily by EEB and ELM with  
695 significant contributions and revisions from MG, GEG, DH, BH, KL, and BK. All authors gave  
696 their final approval for publication and agree to be held accountable for the work performed  
697 therein.

698 **Figure captions**

699 **Fig 1** Tasks comprising the Dog Cognitive Development Battery (DCDB). A) Order of DCDB  
700 tasks implemented in early development (~9 weeks), consisting of three ~45-minute sessions  
701 spread out over three days. B) Order of DCDB tasks implemented in early adulthood, consisting  
702 of two ~1-1.5-hour sessions administered either on the same day or over two consecutive days. In  
703 both panels, the constructs that each task was designed to measure are indicated in bold. A version  
704 of Fig. 1a was published in *Animal Behaviour*, 166, Bray EE, Gruen ME, Gnanadesikan GE,  
705 Horschler DJ, Levy KM, Kennedy BS, Hare BA, MacLean EL, Cognitive characteristics of 8- to  
706 10-week-old assistance dog puppies, 193-206, Copyright (2020), reprinted with permission from  
707 Elsevier.

708 **Fig 2** Longitudinal stability of DCDB traits. Circles reflect the rank-order correlation coefficient  
709 between phenotypic measures collected from puppies and adults. Filled circles reflect significant  
710 correlations and open circles reflect correlations with  $p$  values  $> 0.05$ . For Bayesian mixed model  
711 analyses, the turquoise bars span the interquartile range of the posterior probability distribution for  
712 the beta coefficient relating puppy phenotypes and adult phenotypes; black lines span the 90%  
713 credible interval of the posterior distribution.

714 **Fig 3** Traits with significant longitudinal stability. Points and error bars reflect the mean and  
715 standard error of the adult phenotype

716

717

718

**719 Declarations**

720 Funding: This research was supported in part by grants from the Office of Naval Research (ONR  
721 N00014-17-1-2380 to EM and ONR N00014-16-1-2682 to BH), the American Kennel Club  
722 Canine Health Foundation (Grant No. 02518 to EB and EM), and the Eunice Kennedy Shriver  
723 National Institute of Child Health & Human Development of the National Institutes of Health  
724 (Award No. R01HD097732 to BH).

725 Conflict of interest: The authors declare no conflicts of interest.

726 Ethics approval: All testing procedures were reviewed and adhered to regulations set forth by the  
727 University of Arizona Institutional Animal Care and Use Committee (IACUC # 16-175) and were  
728 collected in accordance with relevant guidelines and regulations.

729 Contributions: The study was designed by all authors. The analysis was primarily designed and  
730 conducted by EEB and ELM. The paper was written primarily by EEB and ELM with significant  
731 contributions and revisions from MG, GEG, DH, BH, KL, and BK. All authors gave their final  
732 approval for publication and agree to be held accountable for the work performed therein.

733 Data availability: The datasets generated during and analyzed during the current study are available  
734 from the corresponding author on reasonable request.

735

736

737

738

739

740 **References**

- 741 Agnetta, B., Hare, B., & Tomasello, M. (2000). Cues to food location that domestic dogs (*Canis*  
742 *familiaris*) of different ages do and do not use. *Anim Cogn*, 3(2), 107-112.  
743 doi:10.1007/s100710000070
- 744 Annett, M. (1994). Handedness as a continuous variable with dextral shift: sex, generation, and  
745 family handedness in subgroups of left-and right-handers. *Behavior genetics*, 24(1), 51-  
746 63.
- 747 Arden, R., Bensky, M. K., & Adams, M. J. (2016). A Review of Cognitive Abilities in Dogs,  
748 1911 Through 2016 More Individual Differences, Please! *Curr Dir Psychol Sci*, 25(5),  
749 307-312. doi:10.1177/0963721416667718
- 750 Barnard, S., Wells, D. L., Hepper, P. G., & Milligan, A. D. (2017). Association between lateral  
751 bias and personality traits in the domestic dog (*Canis familiaris*). *J Comp Psychol*,  
752 131(3), 246-256. doi:10.1037/com0000074
- 753 Batt, L. S., Batt, M., Baguley, J., & McGreevy, P. (2008). Stability of motor lateralisation in  
754 maturing dogs. *Laterality*, 13(5), 468.
- 755 Batt, L. S., Batt, M. S., Baguley, J. A., & McGreevy, P. D. (2009). The relationships between  
756 motor lateralization, salivary cortisol concentrations and behavior in dogs. *J Vet Behav*,  
757 4(6), 216-222. doi:10.1016/j.jveb.2009.02.001
- 758 Ben-Aderet, T., Gallego-Abenza, M., Reby, D., & Mathevon, N. (2017). Dog-directed speech:  
759 why do we use it and do dogs pay attention to it? *Proc R Soc B*, 284(1846), 20162429.  
760 doi:10.1098/rspb.2016.2429
- 761 Bensky, M. K., Gosling, S. D., & Sinn, D. L. (2013). The world from a dog's point of view: a  
762 review and synthesis of dog cognition research *Advances in the Study of Behavior* (Vol.  
763 45, pp. 209-406): Elsevier.

- 764 Bhattacharjee, D., Dev, N., Gupta, S., Sau, S., Sarkar, R., Biswas, A., . . . Bhadra, A. (2017).  
765 Free-ranging dogs show age related plasticity in their ability to follow human pointing.  
766 *PloS one*, *12*(7). doi:10.1371/journal.pone.0180643
- 767 Bibost, A.-L., & Brown, C. (2014). Laterality influences cognitive performance in rainbowfish  
768 *Melanotaenia duboulayi*. *Anim Cogn*, *17*(5), 1045-1051. doi:10.1007/s10071-014-0734-3
- 769 Branson, N., & Rogers, L. (2006). Relationship between paw preference strength and noise  
770 phobia in *Canis familiaris*. *J Comp Psychol*, *120*(3), 176-183. doi:10.1037/0735-  
771 7036.120.3.176
- 772 Bräuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making inferences about  
773 the location of hidden food: social dog, causal ape. *J Comp Psychol*, *120*(1), 38-47.  
774 doi:10.1037/0735-7036.120.1.38
- 775 Bray, E. E., Gruen, M., Gnanadesikan, G. E., Horschler, D. J., Levy, K., Kennedy, B. S., . . .  
776 MacLean, E. (2020). Cognitive characteristics of 8- to 10-week-old assistance dog  
777 puppies. *Anim Behav*, *166C*, 193-206. doi:10.1016/j.anbehav.2020.05.019
- 778 Bray, E. E., Levy, K. M., Kennedy, B. S., Duffy, D. L., Serpell, J. A., & MacLean, E. L. (2019).  
779 Predictive models of assistance dog training outcomes using the Canine Behavioral  
780 Assessment and Research Questionnaire and a standardized temperament evaluation.  
781 *Front Vet Sci*, *6*, 49. doi:10.3389/fvets.2019.00049
- 782 Bray, E. E., MacLean, E. L., & Hare, B. (2014). Context specificity of inhibitory control in dogs.  
783 *Anim Cogn*, *17*(1), 15-31. doi:10.1007/s10071-013-0633-z
- 784 Bray, E. E., Sammel, M. D., Cheney, D. L., Serpell, J. A., & Seyfarth, R. M. (2017). The effects  
785 of maternal investment, temperament, and cognition on guide dog success. *P Natl Acad*  
786 *Sci USA*, *114*(34), 9128–9133. doi:10.1073/pnas.1704303114

- 787 Bray, E. E., Sammel, M. D., Seyfarth, R. M., Serpell, J. A., & Cheney, D. L. (2017).  
788       Temperament and problem solving in a population of adolescent guide dogs. *Anim Cogn*,  
789       20(5), 923-939. doi:10.1007/s10071-017-1112-8
- 790 Brubaker, L., Dasgupta, S., Bhattacharjee, D., Bhadra, A., & Udell, M. A. (2017). Differences in  
791       problem-solving between canid populations: Do domestication and lifetime experience  
792       affect persistence? *Anim Cogn*, 20(4), 717-723. doi:10.1007/s10071-017-1093-7
- 793 Brucks, D., Marshall-Pescini, S., Wallis, L. J., Huber, L., & Range, F. (2017). Measures of Dogs'  
794       Inhibitory Control Abilities Do Not Correlate across Tasks. *Front Psychol*, 8, 849.  
795       doi:10.3389/fpsyg.2017.00849
- 796 Burrows, K. E., Adams, C. L., & Spiers, J. (2008). Sentinels of safety: Service dogs ensure  
797       safety and enhance freedom and well-being for families with autistic children. *Qual*  
798       *Health Res*, 18(12), 1642-1649. doi:10.1177/1049732308327088
- 799 Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and  
800       change. *Annu. Rev. Psychol.*, 56, 453-484. doi:10.1146/annurev.psych.55.090902.141913
- 801 Christie, L.-A., Studzinski, C. M., Araujo, J. A., Leung, C. S., Ikeda-Douglas, C. J., Head, E., . . .  
802       Milgram, N. W. (2005). A comparison of egocentric and allocentric age-dependent  
803       spatial learning in the beagle dog. *Prog Neuro-Psychopharmacol Biol Psychiatry*, 29(3),  
804       361-369.
- 805 Curb, L. A., Abramson, C. I., Grice, J. W., & Kennison, S. M. (2013). The relationship between  
806       personality match and pet satisfaction among dog owners. *Anthrozoös*, 26(3), 395-404.
- 807 Doré, F. Y., Fiset, S., Goulet, S., Dumas, M.-C., & Gagnon, S. (1996). Search behavior in cats  
808       and dogs: interspecific differences in working memory and spatial cognition. *Anim Learn*  
809       *Behav*, 24(2), 142-149.

- 810 Dorey, N. R., Udell, M. A., & Wynne, C. D. (2010). When do domestic dogs, *Canis familiaris*,  
811 start to understand human pointing? The role of ontogeny in the development of  
812 interspecies communication. *Anim Behav*, *79*(1), 37-41.  
813 doi:10.1016/j.anbehav.2009.09.032
- 814 Duffy, D., Kruger, K., & Serpell, J. (2014). Evaluation of a behavioral assessment tool for dogs  
815 relinquished to shelters. *Preventive veterinary medicine*, *117*(3-4), 601.
- 816 Duranton, C., & Gaunet, F. (2016). Effects of shelter housing on dogs' sensitivity to human  
817 social cues. *J Vet Behav*, *14*, 20-27. doi:10.1016/j.jveb.2016.06.011
- 818 Fagnani, J., Barrera, G., Carballo, F., & Bentosela, M. (2016). Is previous experience important  
819 for inhibitory control? A comparison between shelter and pet dogs in A-not-B and  
820 cylinder tasks. *Anim Cogn*, *19*(6), 1165–1172. doi:10.1007/s10071-016-1024-z
- 821 Fiset, S., Beaulieu, C., & Landry, F. (2003). Duration of dogs'(Canis familiaris) working  
822 memory in search for disappearing objects. *Anim Cogn*, *6*(1), 1-10. doi:10.1007/s10071-  
823 002-0157-4
- 824 Flutre, T. (2020). Timothee flutre's personal r code in rutilstimflutre.  
825 <https://github.com/timflutre/rutilstimflutre>.
- 826 Fratkin, J. L., Sinn, D. L., Patall, E. A., & Gosling, S. D. (2013). Personality consistency in dogs:  
827 a meta-analysis. *PloS one*, *8*(1), e54907. doi:10.1371/journal.pone.0054907
- 828 Friedman, J., Hastie, T., & Tibshirani, R. (2009). Glmnet: Lasso and elastic-net regularized  
829 generalized linear models. R package version 4.0. [https://cloud.r-project.org/  
830 package=glmnet](https://cloud.r-project.org/package=glmnet).
- 831 Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear  
832 models via coordinate descent. *J Stat Softw*, *33*(1), 1-22.



- 833 Gácsi, M., Györi, B., Virányi, Z., Kubinyi, E., Range, F., Belényi, B., & Miklósi, Á. (2009).  
834 Explaining dog wolf differences in utilizing human pointing gestures: selection for  
835 synergistic shifts in the development of some social skills. *PloS one*, 4(8), e6584.  
836 doi:10.1371/journal.pone.0006584
- 837 Gácsi, M., Kara, E., Belényi, B., Topál, J., & Miklósi, Á. (2009). The effect of development and  
838 individual differences in pointing comprehension of dogs. *Anim Cogn*, 12(3), 471-479.  
839 doi:10.1007/s10071-008-0208-6
- 840 Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech  
841 situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic  
842 analysis. *Sci Rep*, 7(1), 1-10. doi:10.1038/s41598-017-13883-2
- 843 Goddard, M., & Beilharz, R. (1984a). A factor analysis of fearfulness in potential guide dogs.  
844 *Appl Anim Behav Sci*, 12(3), 253-265. doi:10.1016/0168-1591(84)90118-7
- 845 Goddard, M., & Beilharz, R. (1984b). The relationship of fearfulness to, and the effects of, sex,  
846 age and experience on exploration and activity in dogs. *Appl Anim Behav Sci*, 12(3), 267-  
847 278.
- 848 Goddard, M., & Beilharz, R. (1986). Early prediction of adult behaviour in potential guide dogs.  
849 *Appl Anim Behav Sci*, 15(3), 247-260. doi:0.1016/0168-1591(86)90095-X
- 850 Güntürkün, O., Ströckens, F., & Ocklenburg, S. (2020). Brain Lateralization: A Comparative  
851 Perspective. *Physiol Rev*, 100(3), 1019-1063. doi:10.1152/physrev.00006.2019
- 852 Hall, S. S., MacMichael, J., Turner, A., & Mills, D. S. (2017). A survey of the impact of owning  
853 a service dog on quality of life for individuals with physical and hearing disability: a pilot  
854 study. *Health qual life out*, 15, 59. doi:10.1186/s12955-017-0640-x

- 855 Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social  
856 cognition in dogs. *Science*, 298(5598), 1634-1636. doi:10.1126/science.1072702
- 857 Hare, B., Call, J., & Tomasello, M. (1998). Communication of food location between human and  
858 dog (*Canis familiaris*). *Evol Commun*, 2(1), 137-159. doi:10.1075/eoc.2.1.06har
- 859 Harvey, N. D., Craigon, P. J., Blythe, S. A., England, G. C., & Asher, L. (2017). An evidence-  
860 based decision assistance model for predicting training outcome in juvenile guide dogs.  
861 *PloS one*, 12(6), e0174261. doi:10.1371/journal.pone.0174261
- 862 Harvey, N. D., Craigon, P. J., Sommerville, R., McMillan, C., Green, M., England, G. C., &  
863 Asher, L. (2016). Test-retest reliability and predictive validity of a juvenile guide dog  
864 behavior test. *J Vet Behav*, 11, 65-76. doi:10.1016/j.jveb.2015.09.005
- 865 Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical  
866 learning: Data mining, inference and prediction. *Math Intell*, 27(2), 83-85.
- 867 Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: the lasso  
868 and generalizations*: CRC press.
- 869 Head, E. (2013). A canine model of human aging and Alzheimer's disease. *Biochim Biophys  
870 Acta*, 1832(9), 1384-1389.
- 871 Head, E., Callahan, H., Cummings, B., Cotman, C., Ruehl, W., Muggenberg, B., & Milgram, N.  
872 (1997). Open field activity and human interaction as a function of age and breed in dogs.  
873 *Physiol Behav*, 62(5), 963-971.
- 874 Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis.  
875 *Psychometrika*, 30(2), 179-185.
- 876 Hsu, Y., & Serpell, J. A. (2003). Development and validation of a questionnaire for measuring  
877 behavior and temperament traits in pet dogs. *J Am Vet Med Assoc*, 223(9), 1293-1300.

- 878 Izquierdo, A., & Jentsch, J. D. (2012). Reversal learning as a measure of impulsive and  
879 compulsive behavior in addictions. *Psychopharmacology*, *219*(2), 607-620.  
880 doi:10.1007/s00213-011-2579-7
- 881 Jones, A. C., & Gosling, S. D. (2005). Temperament and personality in dogs (*Canis familiaris*):  
882 A review and evaluation of past research. *Appl Anim Behav Sci*, *95*(1-2), 1-53.  
883 doi:10.1016/j.applanim.2005.04.008
- 884 Kaminski, J., Schulz, L., & Tomasello, M. (2012). How dogs know when communication is  
885 intended for them. *Dev Sci*, *15*(2), 222-232. doi:10.1111/j.1467-7687.2011.01120.x
- 886 Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*:  
887 Academic Press.
- 888 LaFollette, M., Rodriguez, K., Ogata, N., & O'Haire, M. (2019). Military Veterans and Their  
889 PTSD Service Dogs: Associations Between Training Methods, PTSD Severity, Dog  
890 Behavior, and the Human-Animal Bond. *Front Vet Sci*, *6*(23), 1-11.  
891 doi:10.3389/fvets.2019.00023
- 892 Lampe, M., Bräuer, J., Kaminski, J., & Virányi, Z. (2017). The effects of domestication and  
893 ontogeny on cognition in dogs and wolves. *Sci Rep*, *7*, 11690. doi:10.1038/s41598-017-  
894 12055-6
- 895 Lazarowski, L., & Dorman, D. C. (2015). A comparison of pet and purpose-bred research dog  
896 (*Canis familiaris*) performance on human-guided object-choice tasks. *Behav Process*,  
897 *110*, 60-67. doi:10.1016/j.beproc.2014.09.021
- 898 Lazarowski, L., Krichbaum, S., Waggoner, L. P., & Katz, J. S. (2020). The development of  
899 problem-solving abilities in a population of candidate detection dogs (*Canis familiaris*).  
900 *Anim Cogn*. doi:10.1007/s10071-020-01387-y

- 901 Lazarowski, L., Rogers, B., Waggoner, L. P., & Katz, J. S. (2019). When the nose knows:  
902 ontogenetic changes in detection dogs'(Canis familiaris) responsiveness to social and  
903 olfactory cues. *Anim Behav*, *153*, 61-68. doi:10.1016/j.anbehav.2019.05.002
- 904 Lazarowski, L., Strassberg, L. R., Waggoner, L. P., & Katz, J. S. (2019). Persistence and human-  
905 directed behavior in detection dogs: Ontogenetic development and relationships to  
906 working dog success. *Appl Anim Behav Sci*, *220*, 104860.  
907 doi:10.1016/j.applanim.2019.104860
- 908 Lazzaroni, M., Marshall-Pescini, S., Manzenreiter, H., Gosch, S., Přibilová, L., Darc, L., . . .  
909 Range, F. (2020). Why do dogs look back at the human in an impossible task? Looking  
910 back behaviour may be over-interpreted. *Anim Cogn*, 1-15. doi:10.1007/s10071-020-  
911 01345-8
- 912 MacLean, E. L., & Hare, B. (2018). Enhanced selection of assistance and explosive detection  
913 dogs using cognitive measures. *Front Vet Sci*, *5*(236), 1-14.  
914 doi:10.3389/fvets.2018.00236
- 915 MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., . . . Zhao, Y.  
916 (2014). The evolution of self-control. *Proc Natl Acad Sci USA*, *111*(20), E2140-E2148.  
917 doi:10.1073/pnas.1323533111
- 918 MacLean, E. L., Herrmann, E., Suchindran, S., & Hare, B. (2017). Individual differences in  
919 cooperative communicative skills are more similar between dogs and humans than  
920 chimpanzees. *Anim Behav*, *126*, 41-51. doi:10.1016/j.anbehav.2017.01.005
- 921 Magat, M., & Brown, C. (2009). Laterality enhances cognition in Australian parrots. *Proc R Soc*  
922 *B*, *276*(1676), 4155-4162. doi:10.1098/rspb.2009.1397

- 923 Marshall-Pescini, S., Virányi, Z., Kubinyi, E., & Range, F. (2017). Motivational Factors  
924 Underlying Problem Solving: Comparing Wolf and Dog Puppies' Explorative and  
925 Neophobic Behaviors at 5, 6, and 8 Weeks of Age. *Front Psychol.*, 8(180).  
926 doi:10.3389/fpsyg.2017.00180
- 927 Marshall-Pescini, S., Virányi, Z., & Range, F. (2015). The Effect of Domestication on Inhibitory  
928 Control: Wolves and Dogs Compared. *PloS one*, 10(2), e0118469.  
929 doi:10.1371/journal.pone.0118469
- 930 Miklósi, Á., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z., & Csányi, V. (2003). A simple reason  
931 for a big difference: wolves do not look back at humans, but dogs do. *Curr Biol*, 13(9),  
932 763-766. doi:10.1016/S0960-9822(03)00263-X
- 933 Miklósi, Á., Polgárdi, R., Topál, J., & Csányi, V. (1998). Use of experimenter-given cues in  
934 dogs. *Anim Cogn*, 1(2), 113-121. doi:10.1007/s100710050016
- 935 Milgram, N. W., Head, E., Muggenburg, B., Holowachuk, D., Murphey, H., Estrada, J., . . .  
936 Cotman, C. (2002). Landmark discrimination learning in the dog: effects of age, an  
937 antioxidant fortified food, and cognitive strategy. *Neurosci Biobehav Rev*, 26(6), 679-  
938 695.
- 939 Milgram, N. W., Head, E., Weiner, E., & Thomas, E. (1994). Cognitive functions and aging in  
940 the dog: acquisition of nonspatial visual tasks. *Behav Neurosci*, 108(1), 57.
- 941 Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H. L., . . .  
942 Ross, S. (2011). A gradient of childhood self-control predicts health, wealth, and public  
943 safety. *Proc Natl Acad Sci USA*, 108(7), 1-6.

- 944 O'Haire, M. E., & Rodriguez, K. E. (2018). Preliminary efficacy of service dogs as a  
945 complementary treatment for posttraumatic stress disorder in military members and  
946 veterans. *J Consult Clin Psychol*, *86*(2), 179. doi:10.1037/ccp0000267
- 947 Ollivier, F., Plummer, C., & Barrie, K. (2007). Ophthalmic examination and diagnostics. Part 1:  
948 the eye examination and diagnostic procedures. *Veterinary ophthalmology*. 4th ed. Ames,  
949 Iowa: Blackwell Publishing, 438-483.
- 950 Passalacqua, C., Marshall-Pescini, S., Barnard, S., Lakatos, G., Valsecchi, P., & Previde, E. P.  
951 (2011). Human-directed gazing behaviour in puppies and adult dogs, *Canis lupus*  
952 *familiaris*. *Anim Behav*, *82*(5), 1043-1050. doi:10.1016/j.anbehav.2011.07.039
- 953 Piotti, P., Szabó, D., Bognár, Z., Egerer, A., Hulsbosch, P., Carson, R. S., & Kubinyi, E. (2018).  
954 Effect of age on discrimination learning, reversal learning, and cognitive bias in family  
955 dogs. *Learn Behav*, *46*(4), 537-553. doi:10.3758/s13420-018-0357-7
- 956 R Development Core Team. (2016). R: a language and environment for statistical computing.  
957 Vienna, Austria: R Foundation for Statistical Computing. Retrieved from [http://www.R-](http://www.R-project.org)  
958 [project.org](http://www.R-project.org)
- 959 Reid, H. M., & Norvilitis, J. M. (2000). Evidence for anomalous lateralization across domain in  
960 ADHD children as well as adults identified with the Wender Utah rating scale. *J*  
961 *Psychiatr Res*, *34*(4-5), 311-316. doi:10.1016/S0022-3956(00)00027-3
- 962 Revelle, W. (2019). psych: Procedures for Psychological, Psychometric, and Personality  
963 Research. R package version 1.9.12. Retrieved from [https://CRAN.R-](https://CRAN.R-project.org/package=psych)  
964 [project.org/package=psych](https://CRAN.R-project.org/package=psych)

- 965 Rice, W. R., & Gaines, S. D. (1994). 'Heads I win, tails you lose': testing directional alternative  
966 hypotheses in ecological and evolutionary research. *Trends Ecol Evol*, 9(6), 235-237.  
967 doi:10.1016/0169-5347(94)90258-5
- 968 Riedel, J., Buttelmann, D., Call, J., & Tomasello, M. (2006). Domestic dogs (*Canis familiaris*)  
969 use a physical marker to locate hidden food. *Anim Cogn*, 9(1), 27-35.  
970 doi:10.1007/s10071-005-0256-0
- 971 Riedel, J., Schumann, K., Kaminski, J., Call, J., & Tomasello, M. (2008). The early ontogeny of  
972 human–dog communication. *Anim Behav*, 75(3), 1003-1014.  
973 doi:10.1016/j.anbehav.2007.08.010
- 974 Riemer, S., Mills, D. S., & Wright, H. (2014). Impulsive for life? The nature of long-term  
975 impulsivity in domestic dogs. *Anim Cogn*, 17(3), 815-819. doi:10.1007/s10071-013-  
976 0701-4
- 977 Riemer, S., Müller, C., Virányi, Z., Huber, L., & Range, F. (2016). Individual and group level  
978 trajectories of behavioural development in Border collies. *Appl Anim Behav Sci*, 180, 78-  
979 86. doi:10.1016/j.applanim.2016.04.021
- 980 Rodriguez, K. E., Bryce, C. I., Granger, D. A., & O'Haire, M. E. (2018). The effect of a service  
981 dog on salivary cortisol awakening response in a military population with posttraumatic  
982 stress disorder (PTSD). *Psychoneuroendocrinology*, 98, 202-210.  
983 doi:10.1016/j.psyneuen.2018.04.026
- 984 Rodriguez, K. E., LaFollette, M. R., Hediger, K., Ogata, N., & O'Haire, M. E. (2020). Defining  
985 the PTSD service dog intervention: perceived importance, usage, and symptom  
986 specificity of psychiatric service dogs for military veterans. *Frontiers in psychology*, 11,  
987 1638.

- 988 Rosati, A. G., Wobber, V., Hughes, K., & Santos, L. R. (2014). Comparative developmental  
989 psychology: How is human cognitive development unique? *Evol Psychol*, *12*(2), 448-  
990 473. doi:10.1177/147470491401200211
- 991 Rossano, F., Nitzschner, M., & Tomasello, M. (2014). Domestic dogs and puppies can use  
992 human voice direction referentially. *Proc R Soc B*, *281*, 20133201.  
993 doi:10.1098/rspb.2013.3201
- 994 Schmidt, S. L., Carvahó, A. L. N., & Simoes, E. N. (2017). Effect of handedness on auditory  
995 attentional performance in ADHD students. *Neuropsychiatr Dis Treat*, *13*, 2921.
- 996 Scott, J. P., & Fuller, J. L. (1965). *Genetics and the Social Behavior of the Dog*. Chicago:  
997 University of Chicago Press.
- 998 Sforzini, E., Michelazzi, M., Spada, E., Ricci, C., Carenzi, C., Milani, S., . . . Verga, M. (2009).  
999 Evaluation of young and adult dogs' reactivity. *J Vet Behav*, *4*(1), 3-10.  
1000 doi:10.1016/j.jveb.2008.09.035
- 1001 Shaw, G. A., & Brown, G. (1991). Laterality, implicit memory and attention disorder. *Educ Stud*,  
1002 *17*(1), 15-23. doi:10.1080/0305569910170102
- 1003 Simoes, E. N., Carvalho, A. L. N., & Schmidt, S. L. (2017). What does handedness reveal about  
1004 ADHD? An analysis based on CPT performance. *Res Dev Disabil*, *65*, 46-56.  
1005 doi:10.1016/j.ridd.2017.04.009
- 1006 Slabbert, J., & Odendaal, J. (1999). Early prediction of adult police dog efficiency—a  
1007 longitudinal study. *Appl Anim Behav Sci*, *64*(4), 269-288. doi:10.1016/S0168-  
1008 1591(99)00038-6
- 1009 Stan Development Team. (2018). RStan: the R interface to Stan. R package version 2.17.3.  
1010 <http://mc-stan.org>.



- 1011 Starling, M. J., Branson, N., Thomson, P. C., & McGreevy, P. D. (2013). Age, sex and  
1012 reproductive status affect boldness in dogs. *The Veterinary Journal*, *197*(3), 868-872.  
1013 doi:10.1016/j.tvjl.2013.05.019
- 1014 Svartberg, K., Tapper, I., Temrin, H., Radesäter, T., & Thorman, S. (2005). Consistency of  
1015 personality traits in dogs. *Anim Behav*, *69*(2), 283-291.  
1016 doi:10.1016/j.anbehav.2004.04.011
- 1017 Tapp, P. D., Siwak, C. T., Estrada, J., Head, E., Muggenburg, B. A., Cotman, C. W., & Milgram,  
1018 N. W. (2003). Size and reversal learning in the beagle dog as a measure of executive  
1019 function and inhibitory control in aging. *Learn Memory*, *10*(1), 64-73.
- 1020 Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*(4),  
1021 410-433.
- 1022 Tomkins, L. M., Thomson, P. C., & McGreevy, P. D. (2010). First-stepping Test as a measure of  
1023 motor laterality in dogs (*Canis familiaris*). *J Vet Behav*, *5*(5), 247-255.  
1024 doi:10.1016/j.jveb.2010.03.001
- 1025 Tomkins, L. M., Thomson, P. C., & McGreevy, P. D. (2012). Associations between motor,  
1026 sensory and structural lateralisation and guide dog success. *The Veterinary Journal*,  
1027 *192*(3), 359-367. doi:10.1016/j.tvjl.2011.09.010
- 1028 Tomkins, L. M., Williams, K., Thomson, P., & McGreevy, P. (2012). Lateralization in the  
1029 domestic dog (*Canis familiaris*): Relationships between structural, motor, and sensory  
1030 laterality. *Journal of Veterinary Behavior*, *7*(2), 70-79.
- 1031 Torchiano, M. (2020). effsize: Efficient Effect Size Computation. R package version 0.8.0.  
1032 Retrieved from <https://CRAN.R-project.org/package=effsize>

- 1033 Turcsán, B., Tátrai, K., Petró, E., Topál, J., Balogh, L., Egyed, B., & Kubinyi, E. (2020).  
1034 Comparison of Behavior and Genetic Structure in Populations of Family and Kenneled  
1035 Beagles. *Front Vet Sci*, 7(183). doi:10.3389/fvets.2020.00183
- 1036 van Horik, J. O., Beardsworth, C. E., Laker, P. R., Whiteside, M. A., & Madden, J. R. (2020).  
1037 Response learning confounds assays of inhibitory control on detour tasks. *Anim Cogn*,  
1038 23(1), 215-225. doi:10.1007/s10071-019-01330-w
- 1039 van Horik, J. O., Langley, E. J., Whiteside, M. A., Laker, P. R., Beardsworth, C. E., & Madden,  
1040 J. R. (2018). Do detour tasks provide accurate assays of inhibitory control? *Proc R Soc B*,  
1041 285(1875), 20180150. doi:10.1098/rspb.2018.0150
- 1042 Virányi, Z., Gácsi, M., Kubinyi, E., Topál, J., Belényi, B., Ujfalussy, D., & Miklósi, Á. (2008).  
1043 Comprehension of human pointing gestures in young human-reared wolves (*Canis lupus*)  
1044 and dogs (*Canis familiaris*). *Anim Cogn*, 11(3), 373-387. doi:10.1007/s10071-007-0127-y
- 1045 Wallis, L. J., Range, F., Müller, C. A., Serisier, S., Huber, L., & Virányi, Z. (2014). Lifespan  
1046 development of attentiveness in domestic dogs: drawing parallels with humans. *Front*  
1047 *Psychol*, 5, 71. doi:10.3389/fpsyg.2014.00071
- 1048 Wallis, L. J., Szabó, D., & Kubinyi, E. (2020). Cross-Sectional Age Differences in Canine  
1049 Personality Traits; Influence of Breed, Sex, Previous Trauma, and Dog Obedience Tasks.  
1050 *Frontiers in Veterinary Science*, 6, 493.
- 1051 Watowich, M. M., MacLean, E. L., Hare, B., Call, J., Kaminski, J., & Miklósi, Á. (2020). Age  
1052 influences domestic dog cognitive performance independent of average breed lifespan.  
1053 *Anim Cogn*. doi:10.1007/s10071-020-01385-0
- 1054 Wells, D. L. (2003). Lateralised behaviour in the domestic dog, *Canis familiaris*. *Behavioural*  
1055 *processes*, 61(1-2), 27-35.

- 1056 Wells, D. L., Hepper, P. G., Milligan, A. D., & Barnard, S. (2017). Cognitive bias and paw  
1057 preference in the domestic dog (*Canis familiaris*). *J Comp Psychol*, *131*(4), 317-325.  
1058 doi:10.1037/com0000080
- 1059 Whiteside, M. A., Bess, M. M., Frasnelli, E., Beardsworth, C. E., Langley, E. J., van Horik, J. O.,  
1060 & Madden, J. R. (2020). No evidence that footedness in pheasants influences cognitive  
1061 performance in tasks assessing colour discrimination and spatial ability. *Learn Behav*, *48*,  
1062 84-95. doi:10.3758/s13420-019-00402-8
- 1063 Wilsson, E., & Sundgren, P.-E. (1997). The use of a behaviour test for the selection of dogs for  
1064 service and breeding, I: Method of testing and evaluating test results in the adult dog,  
1065 demands on different kinds of service dogs, sex and breed differences. *Appl Anim Behav*  
1066 *Sci*, *53*(4), 279-295. doi:10.1016/S0168-1591(96)01174-4
- 1067 Wilsson, E., & Sundgren, P.-E. (1998). Behaviour test for eight-week old puppies—heritabilities  
1068 of tested behaviour traits and its correspondence to later behaviour. *Appl Anim Behav Sci*,  
1069 *58*(1), 151-162. doi:10.1016/S0168-1591(97)00093-2
- 1070 Wynne, C. D., Udell, M. A., & Lord, K. A. (2008). Ontogeny's impacts on human-dog  
1071 communication. *Anim Behav*, *76*(4), e1-e4. doi:10.1016/j.anbehav.2008.03.010
- 1072 Zaine, I., Domeniconi, C., & Wynne, C. D. (2015). The ontogeny of human point following in  
1073 dogs: When younger dogs outperform older. *Behav Process*, *119*, 76-85.  
1074 doi:10.1016/j.beproc.2015.07.004
- 1075