

This is the final version of the manuscript sent to journal for publication after acceptance. Note the final printed version might be different from this one due to the editing process.

The full citation: Piotti, P., Szabó, D., Bognár, Zs., Egerer, A., Hulsbosch, P., Carson, R. S., Kubinyi, E. 2018. Effect of age on discrimination learning, reversal learning, and cognitive bias in family dogs. *Learning & Behavior*, 1-17. doi: 10.3758/s13420-018-0357-7.

Title: Effect of age on discrimination learning, reversal learning, and cognitive bias in family dogs

Authors: Patrizia Piotti^{1*}, Dóra Szabó¹, Zsófia Bognár¹, Anna Egerer¹, Petrouchka Hulsbosch^{1,2}, Rachel Sophia Carson^{1,3}, Enikő Kubinyi¹

Affiliation:

¹ *Department of Ethology, Eötvös Loránd University, Pázmány P. s 1c, Budapest, 1117, Hungary*

² *Bioengineering Technology, Katholieke Universiteit Leuven, Geel, Belgium*

³ *Kalamazoo College, 1200 Academy St., Kalamazoo, MI 49006, USA*

* Corresponding author: patrizia.piotti@yahoo.it

Abstract

Several studies on age-related cognitive decline in dogs involve laboratory dogs and prolonged training. We developed two spatial tasks that required a single one-hour session. We tested 107 medium-large sized dogs: “young” (N=41, aged 2.5–6.5 years) and “old” (N=66, aged 8–14.5 years). Our results indicated that, in a discrimination learning task and in a reversal learning task, young dogs learned significantly faster than the old dogs, indicating that these two tasks could successfully be used to investigate differences in spatial learning between young and old dogs. We also provide two novel findings. First, in the reversal learning, the dogs trained based on the location of stimuli learned faster than the dogs trained based on stimulus characteristics. Most old dogs did not learn the task within our cut-off of 50 trials. Training based on an object’s location is therefore more appropriate for reversal learning tasks. Second, the contrast between the response to the positive and negative stimuli was narrower in old dogs, compared to young dogs, during the reversal learning task, as well as the cognitive bias test. This measure favours comparability between tasks and between studies. Following the cognitive bias test, we could not find any indication of differences in the positive and negative expectations between young and old dogs. Taken together, these findings do not support the hypothesis that old dogs have more negative expectations than young dogs and the use of the cognitive bias test in older dogs requires further investigation.

Keywords: reversal learning; cognitive bias; dog; learning; memory; ageing

AD, Alzheimer Disease; CB, Cognitive Bias; CBS, Cognitive Bias Score; CCDS, Canine Cognitive Dysfunction Syndrome

Introduction

The family dog as a model for human ageing

An increasing amount of evidence describes the domestic dog (*Canis familiaris*) as a good translational model for the study of ageing (e.g. Cotman & Head, 2008; Cummings et al., 1996; Kaeberlein, Creevy & Promislow, 2016; Head, 2013; Neus Bosch, Pugliese, Gimeno-Bayón, Jose Rodriguez & Mahy, 2012; Schütt et al., 2016). Both dogs and humans spontaneously develop similar medical conditions that increase the risk of death, as they age, such as cancer (Hoffman, Creevy, Franks, O'Neill & Promislow, 2018). Behavioural changes, such as increased anxiety, confusion, alteration of the sleep-wake cycle, may be similarly observed (e.g. dog: Landsberg, Nichol & Araujo, 2012; human: Delphin-Combe et al., 2016). This body of evidence may reflect a naturally occurring and age-related cognitive decline, which follows similar pathways in dogs as it does in humans (Head et al., 1995; Head, 2013). For example, beta-amyloid accumulation has been observed in dogs as early as 9 years of age in the brain region of the prefrontal cortex and from 14 years on in the entorhinal cortex (Head, McCleary, Hahn, Milgram & Cotman, 2000). Atrophy of the cerebral cortex and enlargement of the brain ventricles have also been described in MRI (Magnetic Resonance Imaging) scans of a group 16 years old dogs (Kimotsuki et al., 2005) and in morphometry studies of the cerebral ventricles of both young and old dogs (2-5 years vs 10-12 years, Gonzalez-Soriano et al., 2001). One study looked at MRI scans of a group of 18 dogs aged between 4 and 15 years (with age treated as continuous variable), observing a non-linear relationship between age and brain ventricles enlargement, cortical atrophy. Interestingly, individual differences were also detected, as one 6 year old dog in the sample was as severely affected as the 14 year old dogs in the group; this suggests that some individuals can develop age-related brain degeneration before others and this can be as early as 6 years of age (Su et al., 1998). The same dogs had been tested for their cognitive performance in a different study (Head, Callhan, Muggenburg, Cotman & Milgram, 1998), but no relationships between cognitive findings and brain changes was reported (Su et al., 1998). The dogs had been tested for cognitive performance up to 2 years prior to the MRI scanning (Su et al., 1998; Head et al., 1998) and the results of the cognitive tests were reported according to 3 distinct age groups (young: < 5 years, middle aged: 5-10 years, old: 10+ years; Head et al., 1998), which made any comparison difficult to interpret. Studzinski et al., (2006) also investigated cognitive decline in dogs; they observed spatial deficits starting from the age of 6 and reported that age alone predicted 48.2% of the variability in learning in a task to assess memory. Overall, a large body of evidence indicates that functional decline in cognitive domains, such as learning, memory, executive function, and spatial function, occurs similarly in dogs and humans as they age (for reviews, Cotman & Head, 2008; Head et al., 2013). It should however be noted that, in this body of research, it is difficult to evaluate the effect of age alone as these findings are rarely adjusted to the lifespan variability that is due to factors such as breed, size, or weight (Szabó et al., 2016).

In addition to the physiological reasons described above, the use of family dogs as models for ageing research is also supported by ecological reasons. Through their unique domestication history, dogs have adapted to a specific niche, i.e. the human social environment (Miklósi, Polgárdi, Topál & Csányi, 1998; Hare, Brown, Williamson & Tomasello, 2002). However, a large proportion of the research on dog cognitive decline involves purpose-bred and raised laboratory animals (Head, 2013 for a review). It should be noted that laboratory conditions are unrelated to animals' natural environment (Wood, Desjardins & Fernald, 2011). There is some evidence that family dogs (i.e. dogs living with humans as pets) and purpose-bred research dogs (i.e. kennel reared domestic dogs) diverge in their performance during some cognitive tasks (Lazarowski & Dorman, 2014). Indeed, it has been argued that dogs residing long-term in kennel environments may be affected by cognitive deficits, due to the lack of stimulation provided by their living environment (Miklósi & Topál, 2011; Mongillo et al., 2013). Thus, results from canine cognitive tests, when performed in conditions similar to those of the human

environment, are likely to have strong ecological validity (Waters, 2011). Family dogs, which share the same living environment as humans, are promising subjects for research that is both clinically relevant and provide the necessary vertical integration of findings originated from invertebrate and rodent models (Waters, 2011). Consequently, studies on family dogs' cognitive ageing have begun to emerge (e.g. Chapagain et al., 2017; González-Martínez et al., 2013; Heckler, Tranquilim, Svicerio, Barbosa & Amorim, 2014; Mongillo et al., 2013; Piotti et al., 2017; Wallis et al., 2016; Wallis et al., 2017). One remaining issue is that some of these tests still require prolonged and complex procedures (discussed below); this limits the replicability of such studies, especially outside of the laboratory setting (Heckler et al., 2014 for a discussion). The importance of replicability is being discussed in psychology research, as replications and data reproducibility are necessary to generalise research findings to the general population (Westfall, Judd & Kenny, 2015), and further efforts should be made to increase the replicability of research (Asendorpf et al., 2013).

Cognitive testing of ageing dogs

Decline in the spatial function (i.e. the ability to perceive, remember, and manipulate information within a spatial context, Tapp et al., 2003a), pertaining to executive function (i.e. the selection and monitoring of behaviours for the attainment of goals, Malenka, Nestler & Hyman, 2009), and learning is a part of the normal ageing process (Cotman & Head, 2008). Impairment in the spatial function is particularly of interest because it may be detected before other cognitive deficits emerge (e.g. Dog: Head et al., 1995; Studzinski et al., 2006; Piotti et al 2017; Human: Becker et al., 1988). Reversal learning tasks are often used to detect changes in these two domains as they rely on the dog's ability to differentiate and mentally adjust previously learned behaviours (Cotman & Head, 2008; Milgram et al., 1994). Dogs are initially trained to discriminate between two stimuli (one associated with a reward, usually food, and the other with no reward). Following a delay period, the dogs are presented with the same stimuli, though the discrimination is reversed. Thus the subjects must inhibit their previously learned response to the rewarded object (perseverative behaviour), and learn a new association between the previously unrewarded object and the food reward (Milgram et al., 1994). Older dogs typically require a longer time to learn the associations and commit more errors than younger dogs (Chan et al 2002; Head et al., 1998; Mongillo et al., 2013; Studzinski et al., 2006; Salvin et al., 2011). However, in the literature, grouping based on dogs' age often varied between studies. One study compared old and young dogs, classifying young dogs as 3-5 years and old dogs as 9-12 years (Chan et al., 2002). Another study looked at age both as a continuous variable and after splitting their population into young (< 8 years) and old dogs (≥ 8 years; Mongillo et al., 2013). Others chose more than one cohort; for example, Head et al. (1998) compared young (< 5 years), middle aged (6-9 years) and old dogs (>10 years), while Studzinski et al. (2006) split their sample into 6 cohorts (puppies < 1 years, young 1-2.99 years, adults 3-5.99 years, middle-aged 6-7.99 years, old 8-9.99 years, senior 10-11.99 years). Authors do not always provide the rationale behind their grouping choices or they base it on literature about other species (e.g. primates, Head et al., 1998). However, interestingly, Studzinski et al. (2006) observed deficits in the spatial function from the age of 6 and decline in the performance of reversal learning tasks has been related to age-associated atrophy of the frontal lobe and possibly cortical areas such as the hippocampus and entorhinal cortex (Tapp^[1]_{SEP} et al., 2004; Siwak-Tapp, Head, Muggenburg, Milgram & Cotman, 2008). It should be noted that, in some studies, on simple visual discrimination tasks and procedural learning measures, aged dogs performed as well as younger subjects (Milgram et al., 1994). Authors have suggested that these results may reflect differential susceptibility of various cognitive functions to ageing (Head, 2013) – reversal learning being the more sensitive measure (Milgram et al., 1994). Szabó et al. (2016), however, argued that the disregard of intrinsic and extrinsic factors, such as gender or social environment, may be an underlying cause of

inconsistency in the ageing research findings, which may be addressed through more standardised identification of ageing phenotypes (Szabó et al., 2016). While this may be achieved with testing large and diverse populations, such as those of family dogs, it is also not always feasible due to the training methods of the existing research protocols, which require days or hours of practice. For example, Mongillo et al. (2013) developed a maze-based navigation task designed as learning and reversal learning tasks to test the effect of age on cognitive performance. Comparing dogs younger and older than 8 years, the authors were able to detect an effect of age in the reversal task but not in the learning task or in a retention phase – where dogs were tested 2 weeks after the learning task (Mongillo et al., 2013). However, the test required several weeks to be performed as well as a large apparatus, which makes the procedure difficult to reproduce, e.g. with diseased and demented dogs or in veterinary clinical settings.

Measuring emotional changes in aged dogs

There is evidence that, in humans, cognitive state and mood, i.e. long term emotional states, are closely intertwined. For example, depression and emotional dysregulation are related with deficits in working memory, alteration of inhibitory processes, especially in the presence of negative stimuli, and inability to use rewarding stimuli (for a review: Gotlib & Joormann, 2010). Moreover, humans' risk of developing depression increases as they age, e.g. due to sleep-disturbance and disability (Cole & Dendukuri, 2003), which is in turn a risk factor for the development of Alzheimer Disease, a type of dementia (Ownby, Crocco & Acevedo, 2006). Therefore, frailty has an overall negative impact on the emotional state and quality of life of elderly people (Delphin-Combe et al., 2016). Interestingly, geriatric conditions are considered reasons of welfare concern in relation to the quality of life and welfare of dogs as well (Yeates & Main, 2009). According to dog owners, aged pet dogs present behaviour symptoms that are very similar to those reported in humans, such as increased anxiety, alteration of the wake-sleep cycle, and disorientation (Landsberg et al., 2012; Madari et al., 2015; Neilson, Hart, Cliff & Ruehl, 2001). These have a rather high prevalence, ranging from 28% in 11–12 years old dogs to 68% in 15–16 years old dogs (Neilson et al., 2001). However, to our knowledge, no attempts have been made to investigate the relationship between age-related changes in emotional states and cognitive decline. There are recommendations on monitoring the quality of life of old dogs, however the existing methods are prone to subjectivity (Yeates & Main, 2009). For this reason, it was recently suggested that monitoring quality of life in old dogs may benefit from the use of objective methods such as those based on changes in cognitive performance as measured through behaviour tests (Piotti, 2017).

One such test is the “cognitive bias test”, which measures changes in dogs' tendency to approach ambivalent stimuli (Mendl et al., 2010). A cognitive bias (or, more specifically, a judgement bias) is a known phenomenon, where individuals experiencing long term negative emotional states (moods) tend to interpret ambiguous stimuli more negatively than individuals experiencing more positive moods do (Harding, Paul, & Mendl, 2004; Mendl, Burman, Parker & Paul, 2009; Roelofs, Boleij, Nordquist & van der Staay, 2016). One advantage of cognitive biases is that they are responses to threat of punishment or prediction of reward that generalise beyond the trigger stimuli (Paul & Mendl, 2017). Therefore, measurements obtained during the test may be used to infer the long term emotional state of the animal beyond the testing conditions.

The judgment bias phenomenon is described in several species, including dogs, and has led to the development of behaviour tests that can objectively measure the bias in this species (Harding et al., 2004; Mendl et al., 2009). Typically, subjects are trained to discriminate between positive stimuli (P, which predict a reward, e.g. food) and negative stimuli (N, which predict the absence of reward or a mild punishment). Once the discrimination is established, subjects are exposed to one or more ambiguous stimuli (A), which have features that are in

between those of the positive and the negative stimulus. The intensity of response toward each stimulus type is measured. Animals experiencing a negative mood are thought to have negative expectations, i.e. when they are exposed to ambiguous information, they expect a negative outcome rather than the positive one – the opposite is true for animals experiencing a positive mood (Harding et al., 2004). As a consequence, in the cognitive bias test, responses to the ambiguous stimulus (A) that are on average more similar to those directed to the negative stimulus (N) are considered as an indication of a more negative mood; responses that are more similar to those directed to the positive stimulus (P) reflect a more positive mood (Harding et al., 2004; Mendl et al., 2009). In dogs, the test could successfully separate dogs affected by anxiety related problems (separation related anxiety), as opposed to non-anxious counterparts (Mendl et al., 2010). Subsequently, the cognitive bias test was used to monitor the treatment progress of dogs affected by separation related problems (Karagiannis, Burman & Mills, 2015). Additionally, Kis et al. (2015) proposed the canine cognitive bias as a measure of overall well-being. The authors tested dogs with a cognitive bias test after administering them with either a placebo or oxytocin (which has a known association with psychological well-being in humans; William, Kahloon, Fakhry & Ishak, 2011), in combination with the presence or absence of social stimuli. Consistent with the human literature, the test results indicated that oxytocin, especially in the presence of social stimuli, lead to more positive expectations than the placebo (Kis et al., 2015). More recently, the use of the cognitive bias test was extended to the assessment of the well-being of dogs affected by a chronic and putatively painful condition, i.e. syringomyelia (a neurological malformation that involves the formation of sacs filled with fluid in the spinal neural cord; Cockburn, et al., 2017). The test results indicated that a group of dogs affected by syringomyelia had a more negative expectation as opposed to a group of healthy dogs (Cockburn et al., 2017).

It should be noted that there are some limitations in the use of the cognitive bias test; for example it did not prove to be useful in measuring the effect of short-term positive events (Burman et al., 2011). On the contrary, the evidence in the literature indicates that the test is a reliable measure for the long-term effects on wellbeing that are related with chronic health conditions. However, to our knowledge, the cognitive bias test has never been used to assess the decline in the wellbeing of old dogs.

Aims of the study

The aims of the present study were: 1) to design a simple and reproducible version of the reversal learning task, which did not require large apparatuses or several weeks of training and could detect differences in performance between old and young dogs; 2) to assess the two groups of dogs with a cognitive bias test (Kis et al., 2015). For the first goal, the performance of the old dogs was expected to be poorer (in terms of time required to learn and number of trained dogs) during the reversal learning task. For the second goal, the old dogs were expected to have a more negative expectation towards the ambiguous stimulus when compared to the young dogs.

We were also interested in developing tests that could be potentially repeated over time, e.g. to monitor the progress of the condition. However, both the reversal learning tasks and the cognitive bias test intrinsically have a large carryover effect (i.e. during retesting the subjects are influenced by their previous experience). Therefore, we tested the dogs in this study with different stimuli, in order to explore whether these would affect the dogs' performance. The two types of stimuli varied in their physical characteristics and relied on slightly different cognitive tasks and domains, as described by Cotman and Head (2008). For one group of dogs, the stimuli varied based on their location, as described in the original cognitive bias protocol (Mendl et al., 2010); therefore, the tasks relied on egocentric spatial coding (i.e. the animal can rely on the representation of the objects in space relative to its own body axes, such as left-right and front-back), specifically relying on learning (spatial function) and reversal learning

(executive function). For the other group of dogs, the stimuli varied based on their physical characteristics, such as colour and shape; therefore, the tasks relied on visual learning and reversal learning (executive function). Both tasks also relied on visual discrimination learning and reward and object approach learning (learning domain). Previous findings suggested similar effects of age on location and size discrimination (Tapp et al., 2003b); therefore we expected young dogs to perform better than old dogs in both tasks.

In this study we decided to investigate the normal ageing process, i.e. the cognitive changes that occur in all individuals as they age, rather than more severe signs occurring in a smaller part of the population, therefore the subjects were screened based on symptoms of health issues associated with age. In fact, knowing the “normal” ageing process is a prerequisite for the subsequent identification of pathological phenotypes (Szabó et al., 2016).

Materials and methods

Ethics statement

The procedures applied complied with national and EU legislation and institutional guidelines. The study received Ethical Permission from the Eötvös Loránd University, Budapest, Hungary (Permission No.: PE/EA/2019-5/2017). Owners provided written consent to their participation. Our Consent Form was based on the Ethical Codex of Hungarian Psychologists (2004). We took special care to ensure that the consent process was understood completely by the dog owners. In the Consent Form, participants were informed about the identity of the researchers, the aim, procedure, location, expected time commitment of the experiment, the handling of personal and research data, and data reuse. The owners were not informed about the exact aim of the test. The information included the participant’s right to withdraw their consent at any time. Participants could at any point decline to participate and could request for their data not to be used and/or deleted after they were collected during the experiments. The study was performed in strict accordance with the recommendations in the International Society for Applied Ethology guidelines for the use of animals in research.

Subjects

Power analysis was performed using G*power in order to estimate the required sample size for the study, aiming for a medium effect size ($d = .65$) and a power of .08. The analysis yielded a recommended sample of $N = 35$ per group (old and young). A group of old dogs ($N_{old} = 76$) had already been recruited and tested, following the same procedure as the current work, for a different study (unpublished). Therefore, a matching sample of young dogs had been recruited following the same inclusion and exclusion criteria, although the sample size had been kept closer to that calculated *a priori* ($N_{young} = 43$) in order to avoid inflating the risk of type I errors. For both groups, the inclusion criteria were for the dogs to be free from overt signs of cognitive decline (as reported by the owner) and/or medical problems, which might have affected the performance in the tasks (e.g. untreated neurological conditions, motor-sensory deficits). All dogs were assessed for their motor-sensory skills by a qualified veterinary surgeon or a physiotherapist, following on a previously defined procedure (Bognár et al., in prep). For the young dogs group, dogs between 2 and 6.5 years were recruited; for the old dogs group, dogs had to be above 8 years of age. The age limits had been set based on previous findings about the age when the first signs of cognitive decline were detected (Studzinski et al., 2006; Szabó et al., 2016 for a review). The sample included 54 mix breeds and 69 pure breeds from 17 different types of breed. In order to reduce the confounding effect of dogs’ size on ageing speed (see Szabó et al., 2016), only dogs of medium to large size were included in the sample. Twelve dogs were tested but excluded from the analysis (Supplemental material) due to procedural mistakes during the data collection ($N = 11$) or due to technical issues with the

electronic spread sheet (N = 1). The demographic characteristics of the final sample (N = 107) are described in table 1.

Table 1. Demographic information of the dogs included in the analysis.

Groups		N	Mdn Age [years] (min – max)	M : F (Neutered N)	Mdn Weight [Kg] (min – max)	Mdn Height [cm] (min – max)
Young	41	4.5 (2.5-6.5)	1 : 1 (27)	20.2 (9.5 – 44.2)	50.00 (27.00 - 84.00)	
Old	66	10.5 (8-14.5)	1 : 1 (53)	18.65 (5.9 – 78)	50.00 (35.00 - 70.00)	

Study design

Stimuli and randomisation

Different sets of stimuli were used to train the dogs to discriminate the stimuli based on their location (“location” group) or based on their physical characteristics, i.e. size, colour, and shape (“physical characteristics” group).

For the test performed based on location, dogs had to associate the presence (P) or absence (N) of food with the specific location where the stimulus was placed. The stimulus was a blue plate (plastic, round shape, 20 cm of diameter) and it was placed on the floor either on the left or the right hand side of the experimenter. For about half of the dogs in the location group, the positive location was on the left and the negative was on the right, while for the other half the opposite was true; the same balancing method was used for the colour group (table 2).

Table 2. Groups composition. The number of individuals is shown based on the counter-balancing for type of discrimination (based on location or physical characteristics of the stimuli)

Groups	N total	Location (P = left)	Colour (P = white)
Young	41	21 (10)	20 (9)
Old	66	28 (13)	38 (18)

Notes: P = positive

For the test performed based on the physical characteristics of the stimuli, the stimuli were always placed on the floor right in front of the experimenter and the dog had to learn the association between presence (P) or absence of food (N) based on the shape and colour of the stimuli. These consisted of a set of 3 plates; two plates were used as P and N stimuli and the third plate, with intermediate characteristics between the P and N, was used as an ambiguous stimulus. The two plates used as P and N (Fig 1) were a large, black, rectangular plate (plastic, 23 x 15.5 cm) and a small, white, round plate (plastic, 12 cm). The ambiguous stimulus (A) was a plate of a size that was in between the other two and of a colour (plastic, round, green colour, 15 cm of diameter), which had contrast and chromatic characteristics that are supposed to be perceived as in between black and white, according to dogs’ sensory abilities (Pongrácz et al., 2017).

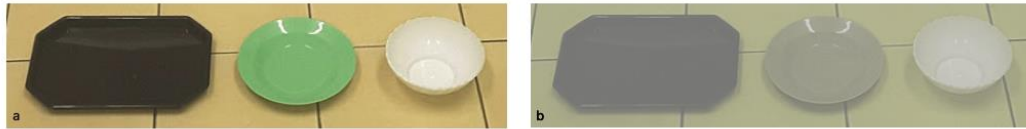


Figure 1. Stimuli used in the physical characteristics group. The two images show the stimuli in the way they are perceived by humans (a) and dogs (b, altered to dog-vision setting through a dedicated [image processing](http://dog-vision.com) tool, <http://dog-vision.com>). The three plates varied in shape, size and colour. The middle plate (used as ambiguous cue) had intermediate characteristics between the other two plates.

All dogs had pre-tested to ensure that they could reliably find a small object placed on the floor – this was done to ensure that their visual abilities were sufficient for them to perform successfully in the current test. Additionally, in all cases, from their starting position the dog could not see the food inside of the plate as we were interested for the dogs to decide whether to approach a stimulus or not based on the outcome associated with it. Finally, previous findings indicate that dogs do not rely on smell in such settings (Szetei et al., 2003); nevertheless, all plates were smeared with food prior to testing in order to control for odour cues. A thin slice of sausage or a piece of cheese of equal size was used as food reward in all conditions (one dog was on a medical diet, so was tested using pieces of food from their own diet); the dog owners were asked about the dog’s preference and dietary requirements.

Testing area

The study was performed at the Department of Ethology, Eötvös Loránd University (Budapest, Hungary), in a small room (2.8 x 5 m). Only the dog’s owner, the dog, and the experimenter were in the room. The room was empty beside a chair placed at one end of the room (for the owner). The owner held the dog’s collar in order to keep the dog in front of the chair (on a spot marked with tape on the floor). The experimenter was standing approximately 3 meters in front of the dog-owner dyad. Two cameras were mounted on the ceiling, one behind the experimenter and the other behind the dog-owner dyad (Fig 2).

For the “location” group, the food plates were placed on the floor right in front of the experimenter (L2) or towards her left (L3) or right-hand side (L1) and at arm’s length distance from her (P and N). For the “colour” group, the plates were always placed in front of the experimenter (L2).

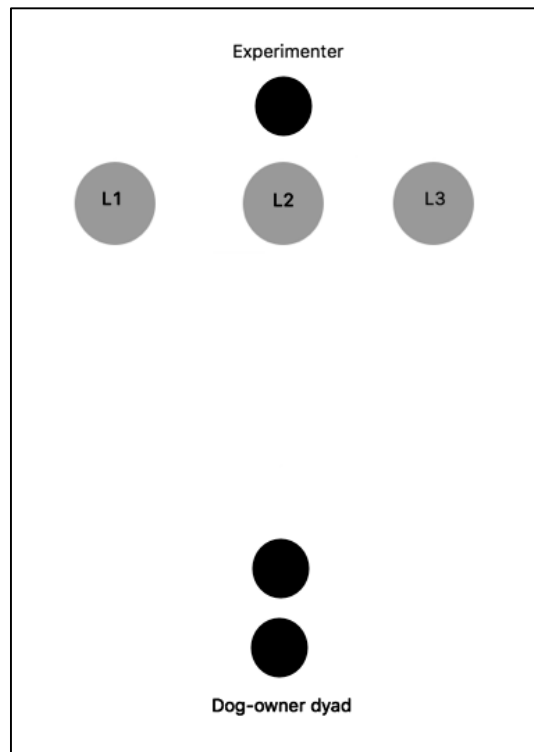


Figure 2. Room set up. The owner and the dog were at one side of the room (black circles at the bottom of the figure) and the experimenter was standing opposite to them (black circle at the top of the figure). The grey circles in front of the experimenter mark the positions where the plates could be placed based on the type of the stimuli (L1 and L3 = positions for the P and N stimuli in the “location” group; L2 = position for the ambiguous stimulus as well as the P and N stimuli in the “physical characteristics” group).

Protocol

The study consisted of four different phases. In the “discrimination learning” (phase 1), dogs were trained to discriminate between the P and N stimuli (“discrimination learning”), then they were given a break of a few minutes. Kis et al. (2015) observed that some dogs appeared to not discriminate well the stimuli after a short delay; therefore, as in their study, the dogs in our study underwent additional training-like trials (phase 2, “consolidation”). In the “cognitive bias test” (phase 3), the dogs were then tested with a similar procedure as in Kis et al. (2015), described below. Finally, the dogs were exposed to a “reversal learning” (phase 4), where the P and N valence of the stimuli was swapped around. The order of presentation of the P, N, and A stimuli was identical for all dogs.

The discrimination and the reversal learning lasted between 15 and 35 minutes each, depending on the dog, and the consolidation and cognitive bias test each lasted less than 5 minutes. The whole procedure lasted about one hour.

1) Discrimination learning

The aim of this phase was to teach the dogs that one stimulus was associated with food (P = positive) and the opposite stimulus was associated with no food (N = negative). This was achieved by presenting the dogs with the two stimuli in consecutive trials; the stimuli were presented in a pre-determined pseudo-random order, with no more than two trials of the same type being presented consecutively (for the order of presentations see the supplemental material).

At the beginning of each trial, the owner was asked to put the dog on the leash while the experimenter baited (or did not bait, depending on the trial valence, N or P) one of the plates.

The experimenter then put the plate on the floor (based on the group type and the trial valence, and according to the pseudo-random order), and in front of her: for the location group this was placed either towards the left-hand side or the right-hand side of the dog; for the location the plate corresponding to that trial valence was placed right in front of the experimenter.

As soon as the plate was on the floor, the owner unleashed the dog. Owners had been instructed to unleash the dogs exactly as the plate touched the floor, so to avoid biasing the dog's behaviour. Dogs had up to 15 seconds to reach the plate and eat the food, while the experimenter looked straight ahead without making eye contact with the dog; once the dog ate the food, or after the 15 seconds had elapsed, the experimenter took the plate and the trial was over. If the dog did not immediately approach the food plates when released, the owner was allowed to encourage it with short sentences, such as "Go", or gently touching the dog. For each trial, the latency to reach the plate was recorded (defined as the amount of time from when the food plate touched the ground to when the dog crossed the line on the floor that marked the plates' position); if the dog did not approach the plate, the experimenter recorded the no-choice and the trial was given a latency of 15 seconds. Dogs were deemed to have learned the association between the stimulus and the food when the longest latency to reach the P plate in the previous 5 trials was shorter than the latency to reach any of the previous N trials ("learning threshold"). Dogs underwent a maximum of 50 training trials; if they did not reach the learning criterion by the 50th trial, they were excluded from further testing. If the dog refused to participate in the test (i.e. did not leave the location of the chair), the training was interrupted. For each dog, the experimenter recorded the latency to reach the food in each trial. These were imported in a spreadsheet at the time of testing; after the first 10 trials, the spreadsheet compared the last 5 latencies to P and the last 5 latencies to N for each trial in order to assess whether the dog had reached the learning threshold. Overall, the following variables were recorded and calculated:

- *number of learning trials*: count variable indicating the number of training trials (this could be either the number of trials required to learn the task, the number of trials before the task was interrupted because the dog refused to participate, or the maximum number of trials (50) if the dog did not learn);
- *learning success*: binary variable indicating whether the dog reached the learning criterion (0 = did not achieve learning, 1 = achieved learning).
- *latency*: time to reach the plate in each trial (continuous variable), if the dog did not make a choice, the maximum latency (15 seconds) was assigned to that trial;
- *learning mean latency P*; *learning mean latency N*: continuous variable indicating the mean latency to the last 5 P and 5 N trials;

2) Consolidation

The dogs that passed the learning threshold in the discrimination learning phase, underwent the consolidation phase ($N = 88$). The aim of this phase was to have a measure of how stable the learning remained after about 5 minute break. This consisted of 10 trials (5 P and 5 N) identical to the training trials (reported in the supplemental material). The following variables were recorded:

- *consolidation success*: binary variable, indicating whether the dog reached the same threshold as in the learning (0 = fail, 1 = success)
- *latency* to reach the plate in each trial (continuous variable), if the dog did not make a choice, the maximum latency (15 seconds) was assigned to that trial;
- consolidation mean latency P*; *consolidation mean latency N*: continuous variable indicating the mean latency to the last 5 P and 5 N trials;

3) Cognitive bias test

Dogs ($N = 88$) were tested following the procedure described in Kis et al. (2015). The test consisted of 3 trials: one negative (N), one positive (P), and one ambiguous (A). The trials were presented in a fixed order (N, P, A), which was identical for all dogs (reported in the supplemental material). The ambiguous trial was not baited; otherwise, the 3 trials were identical to those in the training. The following variable was recorded:

- *latency* to reach the plate in each trial i.e. P, N, A (continuous variable), if the dog did not make a choice, the maximum latency (15 seconds) was assigned to that trial.

4) Reversal learning

This phase was identical to the discrimination learning phase, except only the dogs that had passed the learning threshold were included ($N = 88$) and the P and N were reversed. For example, for a dog in the location group, if the positive stimulus (P) was on the left, now it was on the right; for a dog in the colour group, if the P was the white plate, now it was the black plate. As in the learning phase the dogs underwent a maximum of 50 trials; the learning threshold was the same as in the discrimination learning phase (reported in the supplemental material). The following variables were recorded:

- *number of reversal trials*: count variable indicating the number of trials (this could be either the number of trials required to learn the reversal discrimination, the number of trials before the task was interrupted, or the maximum number of trials (50) if the dog did not learn);
- *reversal learning success*: binary variable indicating whether the dog reached the learning criterion (0 = did not achieve learning, 1 = achieved learning)
- *latency* to reach the plate in each trial (continuous variable), if the dog did not make a choice, the maximum latency (15 seconds) was assigned to that trial;
- *reversal learning mean latency P*; *reversal learning mean latency N*: continuous variable indicating the mean latency to the last 5 P and 5 N trials;

Statistical analysis

Data were analysed using R statistical language (R Core Team, 2017). The recorded variables were subjected to a Shapiro-Wilk normality test (Supplemental material), which indicated that some of the data were not normally distributed; therefore, non-parametric and semi-parametric tests were used. All tests were two-tailed.

1) Discrimination learning

Our variable regarding the number of trials was right censored (i.e. some dogs did not reach the learning outcome within the 50 trials or became uncooperative, leading to interruption of the phase). Therefore a survival analysis was performed on the discrimination learning using a Cox proportional hazard regression model (Cox regression) implemented in the R package “survival.” A Cox regression model for survival-time (time-to-event) was fit to calculate the effect of the age group upon the time (i.e. number of trials) on the occurrence of learning; types of stimuli were included in the model as covariate to evaluate their effect on the learning. The assumption of proportional hazards was tested using the function “cox.zph” and by plotting the estimates over the time to event (supplemental material). The analysis provided the Hazard ratios (HRs) for each parameter, i.e. the ratio of the risk (probability) for the event to occur (i.e. for the dog to learn) for different levels of the explanatory variables. For each survival analysis completed, significant results were presented with hazard ratios (HRs), 95% confidence intervals (CIs), p -values (p), and median survival times (i.e. the time taken to learn). The R squared of the model and test significance (log-rank test) were also reported: the R squared indicates the proportion of the variance in the dependent variable that is attributable to the variables in the model; the log-rank test is a non-parametric test which tests the null-hypothesis that the distribution in the groups under consideration is not different.

A delta score was also calculated to measure the difference between the average latency to the last 5 P and 5 N trials of that given phase ($\Delta_{\text{discr}} = \text{mean lat}_N - \text{mean lat}_P$). Only the dogs which had reached the learning criterion were considered when calculating the delta score. A small delta indicates that the dogs' speeds to the N and to the P are very similar. We then used a Wilcoxon rank-sum test to look at how the delta varied between the two age groups.

2) Consolidation phase

In order to assess any age differences in the dogs' ability to remember the discrimination after a short break, the "consolidation success" variable was analysed using a Chi-squared test of independence, comparing the success and failure between old and young dogs.

A delta score was calculated for the consolidation phase, using the same formula adopted for the discrimination learning ($\Delta_{\text{cons}} = \text{mean lat}_N - \text{mean lat}_P$) and a Wilcoxon rank-sum test was used to assess differences between old and young dogs.

3) Cognitive bias test

The cognitive bias test was analysed as suggested by Mendl et al. (2010). Positive and negative expectations were measured using the formula provided by Mendl et al. (2010), which controls for the possibility that differences in the latency to the ambiguous cue location (A) was due to differences between the dogs in the running speeds. This score ("cognitive bias score", CBS) was calculated by adjusting the A latency by taking into account the baseline latency to P and to N in the test phase:

$$\text{CBS} = ((\text{Lat}_A - \text{Lat}_P) * 100) / (\text{Lat}_N - \text{Lat}_P)$$

This formula varies slightly from the one described in Mendl et al. (2010), in that in our sample there was one trial per condition, therefore our CBS score is calculated from absolute values rather than averages. A Wilcoxon sum-rank test was used to analyse the differences in the CBS between the two age groups and between types of stimuli (alpha = .025 to correct for multiple comparisons).

Startling et al. (2014) recommended comparing average individual latencies to P and to N in the cognitive bias test, because the average latency to reach P needs to be significantly less than the average length to reach N for the test to be meaningful. However, in our case only one score for each stimulus valence was available (as our cognitive bias test had only 1 P and 1 N). Therefore, a delta score was used instead, although in this case it was calculated as the difference between the latency to N and the latency to P ($\Delta_{\text{CBS}} = \text{lat}_N - \text{lat}_P$). The main interest was to ensure that the delta was positive for all dogs. However, the difference between young and old dogs was assessed as well using a Wilcoxon rank-sum test.

4) Reversal learning

The effect of the age group on the time (i.e. number of trials) necessary for the occurrence of learning was analysed using Cox regression models, with the same modalities followed for the discrimination learning. Two additional Cox regression models were calculated for the location and the physical characteristics groups separately.

Latencies were analysed with a Wilcoxon rank-sum test, after calculating a delta score as in the discrimination learning ($\Delta_{\text{rev}} = \text{mean Lat}_N - \text{mean Lat}_P$).

Results

1) Discrimination learning

A Cox proportional hazard model (Table 3) indicated that, during the discrimination learning phase, the number of trials that the dogs required to reach the learning criterion was influenced by the age group, with the young dogs learning faster when compared to the old dogs (Table 4, Fig 3). The type of stimuli did not significantly affect the number of trials required to learn (Table 3).

Table 3. Cox Proportional Hazards estimates of the determinants of learning in the discrimination learning phase, by age group and type of stimuli.

Variable	Coefficient ^a	HR ^b	Test statistics	<i>p</i>
Age group^c				
(Old)				
Young	.6513	1.91	2.99	.003
Type of stimuli^c				
(Physic.)				
Location	-.3952	.67	-1.83	.066
Total N		107		
Wald test		12.25 (2 df), <i>p</i> = .002		
Test for proportional hazards		$\chi^2 = 4.083$, <i>p</i> = .129		
R ²		.10		

^a Coefficient = log (relative hazard).

^b HR = Hazard Ratio.

^c Reference category in parenthesis.

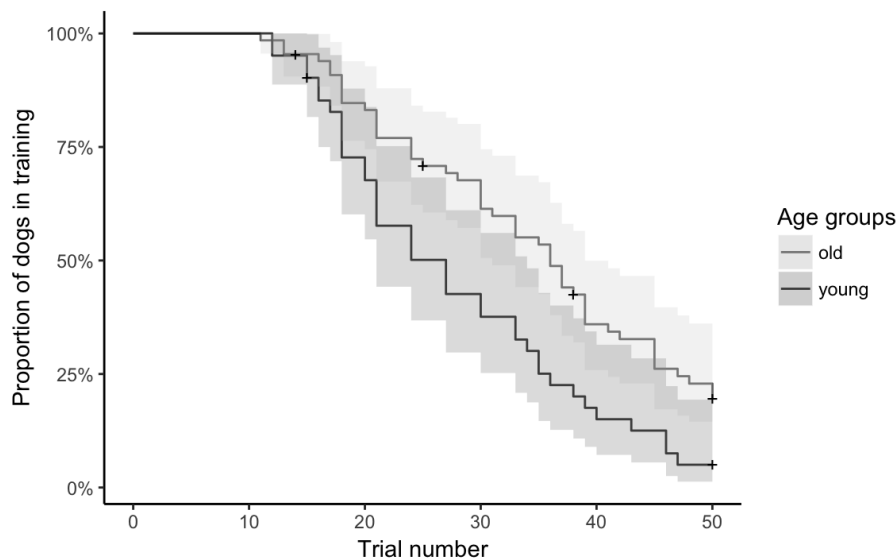


Figure 3. Survival curve for the number of trials to reach the learning criterion in the discrimination learning according to the age group. The darker line represents the young dogs, the lighter line represents the old dogs. The shadowed areas reflect the confidence intervals and the crosses (+) indicate censored data (i.e. dogs for which the task was interrupted but they had not learned).

Table 4. Median number of trials necessary to reach the learning criterion in the discrimination learning phase, according to age groups.

Variable	N	Events (learning)	Median	95% CI
Age group				
Old	66	50	36	31-39
Young	41	38	27	21-34

The Wilcoxon rank-sum test (Table 12) revealed that there was no significant difference in the delta score between young and old dogs (Fig 6).

As table 4 indicates, 82% of the dogs ($N_{\text{events}} = 50 + 38 = 88$) reached the learning criterion during the discrimination learning phase. The dogs which did not learn the discrimination were excluded from further testing.

2) Consolidation phase

The Chi-square test of independence indicated that there was no significant difference between old and young dogs in their success during the consolidation phase (Table 5), $\chi^2_{(1)} = 1.36, p = .243$.

Table 5. Contingency table for the consolidation success variable.

	Success		Total
	0 (%)	1 (%)	
Old	26 (53%)	24 (47%)	50 (100%)
Young	15 (40%)	23 (60%)	38 (100%)
Total	41 (47%)	47 (53%)	

Note: N = 88

The Wilcoxon rank-sum test (Table 12) revealed that there was no significant difference in the delta score between young and old dogs (Fig 6). Additionally, for some of the dogs, the delta score was equal or below 0 ($N_{\text{consolidation}} = 3$).

3) Cognitive bias test

Two Wilcoxon rank-sum tests (Table 6) revealed that there was no significant difference in the CBS between the young and the old dogs or between the two types of stimuli.

Table 6. Results of the Wilcoxon rank-sum test for the cognitive bias score (CBS), as compared between old and young dogs and between type of stimuli (location and physical characteristics). The effect size is reported ($r = z / \sqrt{N}$), significance level was corrected for multiple comparisons ($\alpha = .025$)

	<i>W</i>	<i>p</i>	<i>r</i>
Old-Young	1043	.433	-.08
Location-Physic.	986	.867	-.02

Note: N = 88

A Wilcoxon rank-sum test (Table 12) revealed that the delta score was significantly smaller in the old dogs compared to the young dogs (Fig 6). Again, for some of the dogs the delta was negative or equal to 0 ($N_{\text{CB}} = 2$).

4) Reversal learning

A second Cox proportional hazard model (Table 7) indicated that, during the reversal learning phase, the number of trials that the dogs required to reach the learning criterion was influenced by both the age group and type of stimuli. The young dogs required fewer trials to reach the learning criterion, compared to the old dogs; similarly, the dogs in the location group required less trials to learn, compared to the dogs in the physical characteristics group.

Table 7. Cox Proportional Hazards estimates of the determinants of learning in the reversal learning phase, by age group and type of stimuli.

Variable	Coefficient ^a	HR ^b	Test statistics	<i>p</i>
Age group^c				
(Old)				
Young	1.1091	3.03	4.15	<.001
Type of stimuli^c				
(Location)				
Physic.	-1.46	.23	-5.34	<.001
Total N		88		
Wald test		36.10 (2 df), <i>p</i> < .001		
Test for proportional hazards		$\chi^2 = .815, p = .665$		
R ²		.37		

^a Coefficient = log (relative hazard).

^b HR = Hazard Ratio.

^c Reference category in parenthesis.

As the Hazard Ratios for the two types of stimuli differ significantly, with the dogs in the physical characteristics having 74% less of probability to learn the task, compared to the dogs in the location group, the two groups were further assessed separately, in order to obtain predictions regarding the effect of age alone for each of them.

Another Cox proportional hazard model (Table 8) indicated that, during the reversal learning phase of the location group, the number of trials that the dogs required to reach the learning criterion was influenced by the age group, with young dogs requiring fewer trials to reach the learning criterion, compared to the old dogs (Table 9, Fig 4).

Table 8. Cox Proportional Hazards estimates of the determinants of learning in the reversal learning phase in the location group, by age group.

Variable	Coefficient ^a	HR ^b	Test statistics	<i>p</i>
Age group^c				
(Old)				
Young	.91	2.49	2.658	.008
Total N		42		
Wald test		7.07 (1 df), <i>p</i> = .008		
Test for proportional hazards		$\chi^2 = 1.14, p = .285$		
R ²		.15		

^a Coefficient = log (relative hazard).

^b HR = Hazard Ratio.

^c Reference category in parenthesis.

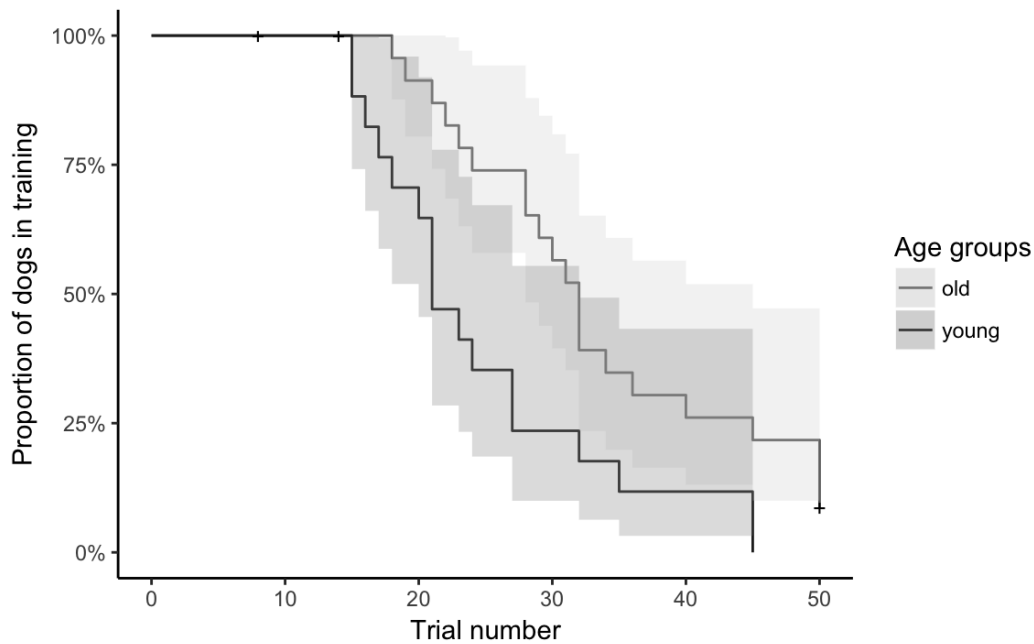


Figure 4. Survival curve for the number of trials to reach the learning criterion in the reversal learning phase of the location group according to the age group. The darker line represents the young dogs, the lighter line represents the old dogs. The shadowed areas reflect the confidence intervals and the crosses (+) indicate censored data (i.e. dogs for which the task was interrupted but they had not learned).

Table 9. Median number of trials necessary to reach the learning criterion in the reversal learning phase of the location group, according to age groups.

Variable	N	Events (learning)	Median	95% CI
Age group				
Old	24	20	32	28-45
Young	18	17	21	20-32

The same analysis was repeated for the physical characteristics group and the Cox proportional hazard model (Table 10) indicated that, during the reversal learning, the number of trials that the dogs required to reach the learning criterion was influenced by the age group, with young dogs requiring fewer trials to reach the learning criterion, compared to the old dogs (Table 11, Fig 5).

Table 10. Cox Proportional Hazards estimates of the determinants of learning in the reversal learning phase in the physical characteristics group, by age group.

Variable	Coefficient ^a	HR ^b	Test statistics	<i>p</i>
Age group^c				
(Old)				
Young	1.351	3.862	3.092	.002
Total N		46		
Wald test		9.56 (1 df), <i>p</i> = .002		
Test for proportional hazards		$\chi^2 = .01, p = .99$		
R ²		.20		

^a Coefficient = log (relative hazard).

^b HR = Hazard Ratio.

^c Reference category in parenthesis.

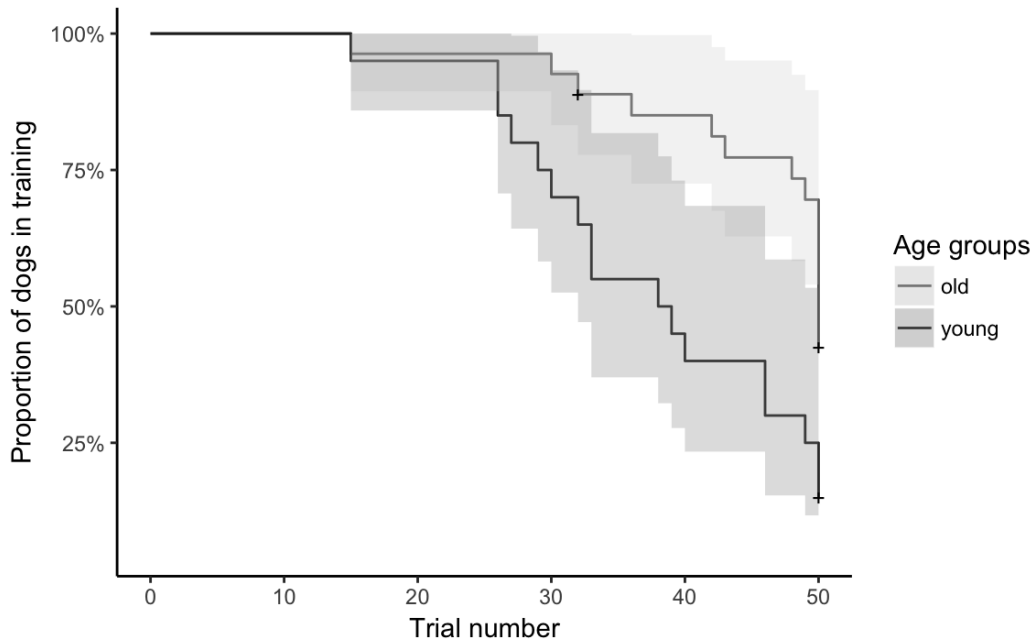


Figure 5. Survival curve for the number of trials to reach the learning criterion in the reversal learning phase of the physical characteristics group according to the age group. The darker line represents the young dogs, the lighter line represents the old dogs. The shadowed areas reflect the confidence intervals and the crosses (+) indicate censored data (i.e. dogs for which the task was interrupted but they had not learned).

Table 11. Median number of trials necessary to reach the learning criterion in the reversal learning phase of the physical characteristics group, according to age groups.

Variable	N	Events (learning)	Median	95% CI
Age group				
Old	26	8	39	-
Young	20	16	38.5	32 - 50

A Wilcoxon rank-sum test (Table 12) revealed that the delta score was significantly smaller in the old dogs compared to the young dogs (Fig 6). Furthermore, for several dogs, the delta score was negative or equal to 0 ($N_{\text{reversal}} = 7$).

Table 12. Results of the Wilcoxon rank-sum test for the delta score, as compared between old and young dogs. The effect size is reported ($r = z / \sqrt{N}$), significance level was corrected for multiple comparisons ($\alpha = .025$)

	<i>Mdn (min, max)</i>		<i>W</i>	<i>p</i>	<i>r</i>
	Young	Old			
Discrimination learning	3.01 (.11, 10.26)	1.03 (.19, 10.17)	739	.075	-.19
Consolidation	4.57 (-.58, 12.06)	2.02 (-.23, 10.40)	728	.060	-.19
Cognitive bias test	9.38 (.08, 13.81)	.86 (-.62, 12.99)	567	.001	-.34
Reversal learning	1.27 (.00, 8.65)	.58 (-.18, 9.19)	649	.011	-.27

Note: $N = 88$

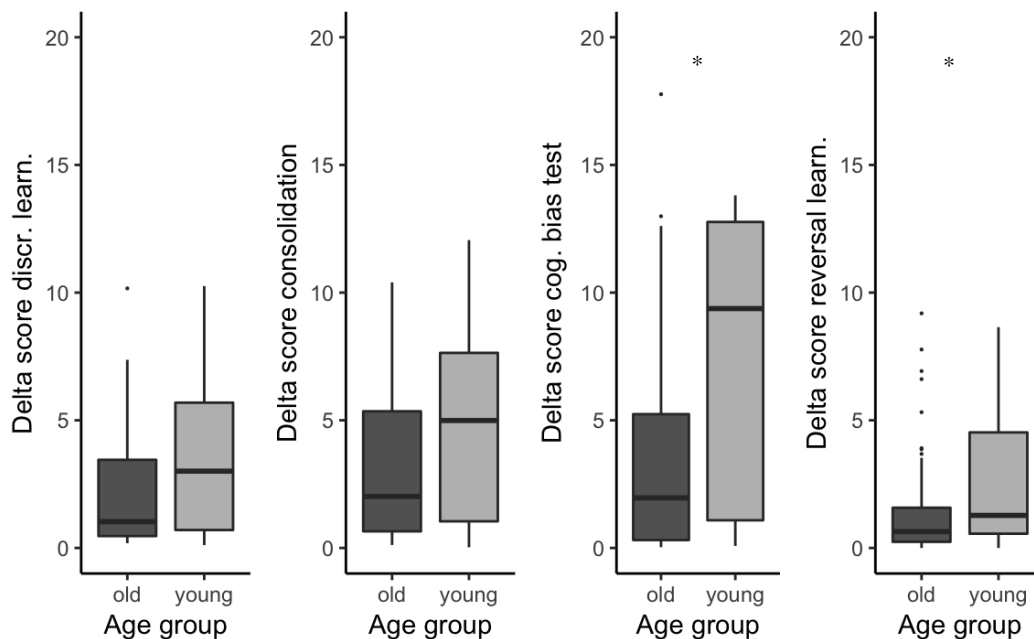


Figure 6. Box-plot graph for the delta scores of the discrimination learning, consolidation, cognitive bias and reversal learning phases. The tick bar represents the median score for the

group, the boxes are the 25% and 75% quartiles and the error bars represent the minimum and maximum values; * indicates $p < .025$.

Discussion

The first aim of this study was to develop a reversal learning task which could detect age-related changes in the cognitive function of family dogs. We found that both the reversal learning and the preceding discrimination learning tasks could detect age differences in a group of family dogs without overt medical problems. Younger dogs (2-6.5 years), as a group, required fewer repetitions than older dogs (7-14.5 years) to discriminate between two stimuli; the same was true for the reversal learning task. In contrast to previous discrimination and reversal learning tests designed for the same purpose, our protocol could be performed in a short time frame (about 1 hour).

Our second aim was to assess mood differences between the two groups, old and young dogs, through a cognitive bias test. The two groups did not vary in their performance in the test, as measured by the cognitive bias score (Mendl et al., 2010). This suggests that, contrary to our hypothesis, successfully ageing dogs might not have more negative expectations in the presence of ambiguous stimuli. However, a delta score (measuring a difference in the speed approaching stimuli predicting a reward, P, and stimuli predicting the absence of reward, N) showed that, during the cognitive bias test and at the end of the reversal learning, older dogs approached the two stimuli (P and N) at a very similar speed. This delta score may be interpreted as an indirect measure of the contrast or magnitude of the discrimination between stimuli predicting a reward and stimuli predicting the absence of reward, and is suggested here as a novel measure in the study of dog ageing. The small contrast observed in the speed to approach P and to approach N could be due to lower motivation in old dogs, or cognitive reasons, e.g. decline in memory function or spatial function. Either way, because the old dogs responded very similarly to both P and N or were sometimes even faster in reaching the N than the P, the results regarding the cognitive bias test performed on this sample are limited in their interpretability in relation to mood. Nevertheless, the relationship between age and the delta score highlights the fact that cognitive tasks requiring training might not always be feasible in old dog populations.

Another novelty of our study is the indication that the types of stimuli affect the level of difficulty of the task. Especially during the reversal learning, it was clear that the learning was more difficult (required more trials) for the dogs trained based on the physical characteristics of the stimuli.

The literature shows inconsistent evidence regarding age effects on discrimination learning (detected: Milgram et al., 2002; Studzinski et al., 2006; Tapp et al., 2003a,b; Tapp et al., 2004; Wallis et al., 2016; not detected: Christie et al., 2005; Milgram et al., 1994; Mongillo et al., 2014). On the contrary, the age effect is more stable when reversal learning tasks are employed (Christie et al., 2005; Milgram et al., 1994; Mongillo et al., 2014; Tapp et al., 2003a,b; Tapp et al., 2004). On the other hand, dog populations, age ranges and types of tasks (as well as task design) vary largely in the literature (Szabó et al., 2016); therefore, any direct comparison between tasks should be cautious. Mongillo et al. (2013) developed a spatial learning task for family dogs. The authors reported no difference in the discrimination learning between young dogs (3-7.3 years) and old dogs (8-14 years) and concluded that the task is likely too simple (Mongillo et al., 2013). It should be noted that 25% of the dogs in their sample (young: 23%; old: 28%) did not learn the discrimination. In our sample the rate in the learning failure was relatively similar for the old dogs (27%), but not for the young ones (7%), suggesting that the discrimination learning alone is more difficult for the old compared to the young dogs. Mongillo et al. (2013) suggested that the differences in performance that they observed between their discrimination learning and reversal learning tasks were likely due to variability

in the level of difficulty between the two tasks. However, easier tasks should not be deemed as not useful; in fact, they allow for measuring a wider range of variability within a population. For example we observed that the dogs in our study found size, shape and colour discrimination more difficult than egocentric spatial learning in the reversal learning task (i.e. dogs in the “physical characteristics” group required more trials than dogs in the “location” group), although we cannot know which of these three components were used by the dogs and which were more difficult to them. Possibly as a consequence of this we observed a ceiling effect in the “physical characteristics” group, as the median number of trials of the old dogs corresponded to the maximum number of trials for the task. One might also argue that older dogs had impaired visual abilities and could not see the plates. Although, all dogs included in the current study had been preselected to ensure that they were able to reliably find small objects placed on floor. We cannot completely exclude that differences in the visual function affected old dogs’ ability to discriminate the stimuli in the “physical characteristics” group. Regardless of the reason, our results show that excessively difficult tasks might reduce the variance within the tested population, making it impossible to investigate within group differences.

The focus on the discrimination learning task brings an additional level of detail in the assessment of dogs’ cognitive function, especially overcoming the issue of the ceiling effect observed in the reversal learning task, which might be relevant in the assessment of dogs with a higher deficit. Thus, tasks which are difficult for successfully ageing dogs and lead to a ceiling effect, cannot be used for testing dogs with cognitive dysfunction. On the contrary, as previously mentioned, reversal learning is a very consistent measure of age related decline (e.g. Christie et al., 2005; Milgram et al., 1994; Mongillo et al., 2014; Tapp et al., 2004). Our findings therefore suggest that the discrimination learning and reversal learning, based on location or combined with learning based on physical characteristics, may be an efficient task for the objective investigation of age-related cognitive decline in family dogs.

One possible reason for the discrepancies reported in the literature, in relation to learning and reversal learning, may be the fact that not all brain regions are affected by ageing at the same pace and not all subjects are equally affected at a given age; moreover, different cognitive functions rely on different brain areas. For example, Tapp et al. (2004) observed increasing atrophy in the frontal lobe of dogs as their age increased (starting from 8 – 11 years) as well as an inverse correlation between frontal lobe volume and a discrimination learning and reversal learning task. Moreover, Studzinski et al. (2006) observed deficits in the acquisition of a discrimination task in dogs aged 10+ years, compared to dogs below 8 years of age. However, they also reported that not all dogs in their sample showed spatial impairment, with 25% of the dogs above the age of 10 showing no deficit in their performance. Similarly, Head et al. (2000) reported signs of beta-amyloid accumulation in some brain areas (enthorinal cortex, prefrontal cortex) from the age of 9 years, but only in subsets of dogs (42% vs 75% respectively), which was not consistently affected until the age of 14 years. Finally, Siwak-Tapp et al. (2007) found that hippocampal neurogenesis is significantly lower in dogs aged 3-5 years, compared to dogs aged 13-15 years. Additionally, hippocampal neurogenesis was negatively correlated with dogs’ performance in a number of cognitive tasks, including object discrimination, size discrimination, and black/white discrimination (Siwak-Tapp et al., 2007). The dogs in our sample were screened for neurological conditions, sensory-motor decline, and general signs of pathological ageing. Therefore our results could mirror cognitive changes of successfully aging dogs.

One advantage of the use of family dogs as models for the study of ageing is the large and diverse population available, as long as non-invasive methods are employed. The discrimination learning and reversal learning tests described in our study could be used in combination with imaging techniques (e.g. MRI, fMRI, EEG) to further investigate the role of structural and functional changes in the brain on cognitive function. Future research should also

investigate ways to overcome the issue with the non-learning dogs. This is a limitation of the current study (as well as any study requiring training), as the dogs that do not learn the discrimination cannot be assessed in such cognitive tasks. One way to attempt overcoming this issue could be to increase the number of training trials during the discrimination learning; however, this strategy has a limitation in that it prolongs the overall duration of the task.

One interesting finding is the variation in the delta score (i.e. difference between the latency to N and the latency to P, which could be defined as positive-negative contrast) between old and young dogs. Originally derived from the cognitive bias test research (Starling et al., 2011), the measure should provide an indication of dogs' ability to discriminate between the two stimuli. It is therefore interesting that we could observe two phenomena: 1) based on the delta score, there was a marked difference between the young and the old dogs, both during the cognitive bias test and the reversal learning; the young dogs tended to have a higher delta in terms of central tendency (median value) as well as spread (min-max range). It is interesting that the difference between the groups appeared in the two phases where the discrimination may be more difficult for the dogs. Dogs might have found the discrimination difficult during the cognitive bias because of the delay from the training (despite of the use of a block of consolidation trials), and in the reversal learning due to the cognitive conflict with the previous learning. We did not analyse further delta difference between tests as this was beyond the scope of the current paper; however, another advantage of measures such as a delta score is that it allows for comparability between tasks and between papers.

Our findings do not support the hypothesis that old dogs have more negative expectations than young dogs. However, as previously mentioned, the fact that the dogs in our sample did not consistently discriminate between the P and the N, deemed the results of the cognitive bias test to be difficult to interpret (Roelofs et al., 2016; Starling et al., 2014). Kis et al. (2015) reported that dogs could not remember the discrimination between P and N after a delay and indeed we observed that a large number of dogs did not repeat their discrimination learning performance during the consolidation phase (5 P and 5 N trials). This is something that should be explored more in the cognitive bias research, especially when the training protocols or the types of stimuli vary or there is a time gap between training and testing. Another reason for the inconsistent findings that were observed in the cognitive bias test could be the limited number of test trials (i.e., 3 trials). The protocol had been previously used and yielded successful results in detecting differences between groups of dogs with the same number of trials (Kis et al., 2015). One solid justification for the use of few trials is that, with repetitions, subjects of the cognitive bias test readily learn that the ambiguous cues are unrewarded (a phenomenon called 'loss of ambiguity'), which may lead to false conclusions (Doyle et al., 2010). In contrast, repeated tests may be particularly relevant in the case of ageing research (e.g. to monitor changes over time), therefore further investigations aiming to find a balance between loss of ambiguity and susceptibility to random errors is necessary. It has also been pointed out that personality traits (e.g. trait anxiety) might affect the results of the cognitive bias test (Roelofs et al., 2016), which is particularly relevant for a between subject design. For example, Starling et al. (2014) reported individual differences in baseline judgement bias in dogs. This should be taken into account, for example by matching the cognitive bias results with other behavioural observations and personality measures of the animal (Roelofs et al., 2016). Research should also investigate other measures potentially related to welfare, such as anticipatory behaviour (Clegg & Deflour, 2018).

Overall, our findings indicate that a discrimination task based on the location of two objects, accompanied by a reversal learning task, may be an efficient way to measure age related decline in learning performance in family dogs. Further studies are necessary to better understand to which degree variation in the performance of this test might relate to functional changes in the brain or whether the test is predictive to cognitive decline. The task may be associated to colour and shape discrimination, although care should be taken to avoid ceiling

effects, e.g. by increasing the maximum number of trials in the test. Future research should investigate the stability over time of the learned associations. This is particularly important when other tests (for example the cognitive bias test) rely on this association. However, the effectiveness of a cognitive bias test on a population of aged dogs is yet to be demonstrated.

Acknowledgements: We would like to thank Stiegman B., Deés A., Hemző V., Böröczki R., Marosi S., for their help with data collection and coding, and the dog owners who took part in the study with their dogs. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 680040) and was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences for EK and by the Budapest Semester in Cognitive Science program (BSCS-US LLC, Kalamazoo, Michigan) and the Hungarian Academy of Cognitive Science for RSC.

References

Asendorpf, J.B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J.J., Fiedler, K., Fiedler, S., Funder, D.C., Kliegl, R., Nosek, B.A. and Perugini, M., 2013. Recommendations for increasing replicability in psychology. *European Journal of Personality*, 27(2), pp.108-119.

Becker, J. T., Huff, F. J., Nebes, R. D., Holland, A., & Boller, F. (1988). Neuropsychological function in Alzheimer's disease. *Arch Neurol*, 45, 263-268.

Bognár Zs.*, Piotti P.*, Szabó D., Le Nézet L., Kubinyi E. Behavioural measures of dogs' responsiveness to visual and auditory stimuli. * these authors contributed equally to the work (in prep)

Burman, O., McGowan, R., Mendl, M., Norling, Y., Paul, E., Rehn, T., & Keeling, L. (2011). Using judgement bias to measure positive affective state in dogs. *Applied Animal Behaviour Science*, 132(3), 160-168.

Chapagain, D., Virányi, Z., Wallis, L. J., Huber, L., Serra, J., & Range, F. (2017). Aging of attentiveness in border collies and other pet dog breeds: the protective benefits of lifelong training. *Frontiers in aging neuroscience*, 9, 100.

Christie, L., Studzinski, C.M., Araujo, J.A., Leung, C.S., Ikeda-douglas, C.J., Head, E., Cotman, C.W., Milgram, N.W., 2005. A comparison of egocentric and allocentric age-dependent spatial learning in the beagle dog. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 29, 361e369.

Clegg, I. L., & Delfour, F. (2018). Cognitive judgement bias is associated with frequency of anticipatory behavior in bottlenose dolphins. *Zoo biology*.

Cockburn, A., Smith, M., Rusbridge, C., Fowler, C., Paul, E.S., Murrell, J.C., Blackwell, E.J., Casey, R.A., Whay, H.R. and Mendl, M., 2017. Evidence of negative affective state in Cavalier King Charles Spaniels with syringomyelia. *Applied Animal Behaviour Science*.

- Cotman, C. W., & Head, E. (2008). The canine (dog) model of human aging and disease: dietary, environmental and immunotherapy approaches. *Journal of Alzheimer's Disease*, *15*(4), 685-707.
- Cummings, B. J., Head, E., Ruehl, W., Milgram, N. W., & Cotman, C. W. (1996). The canine as an animal model of human aging and dementia. *Neurobiology of Aging*, *17*(2), 259–268.
- Delphin-Combe, F., Bathsavanis, A., Rouch, I., Liles, T., Vannier-Nitenberg, C., Fantino, B., Dauphinot, V. and Krolak-Salmon, P., 2016. Relationship between anxiety and cognitive performance in an elderly population with a cognitive complaint. *European journal of neurology*, *23*(7), pp.1210-1217.
- Doyle, R. E., Vidal, S., Hinch, G. N., Fisher, A. D., Boissy, A., and Lee, C. (2010). The effect of repeated testing on judgement biases in sheep. *Behav. Processes* *83*, 349–352. doi: 10.1016/j.beproc.2010.01.019
- Gonzalez-Soriano, J., Garcia, P. M., Contreras-Rodriguez, J., Martinez-Sainz, P., & Rodriguez-Veiga, E. (2001). Age-related changes in the ventricular system of the dog brain. *Annals of Anatomy-Anatomischer Anzeiger*, *183*(3), 283-291.
- González-Martínez, Á., Rosado, B., Pesini, P., García-Belenguer, S., Palacio, J., Villegas, A., Suárez, M.L., Santamarina, G. and Sarasa, M., 2013. Effect of age and severity of cognitive dysfunction on two simple tasks in pet dogs. *The Veterinary Journal*, *198*(1), pp.176-181.
- Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual review of clinical psychology*, *6*, 285-312.'
- Harding, E. J., Paul, E. S., & Mendl, M. T. (2004). Cognitive bias and affective state. *Nature*, *427*(January), 312. <http://doi.org/10.1038/427312a>
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, *298*(5598), 1634-1636.
- Head, E. (2013). A canine model of human aging and Alzheimer's disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, *1832*(9), 1384-1389.
- Head, E., Callahan, H., Muggenburg, B. A., Cotman, C. W., & Milgram, N. W. (1998). Visual-discrimination learning ability and β -amyloid accumulation in the dog. *Neurobiology of aging*, *19*(5), 415-425.
- Head, E., McCleary, R., Hahn, F. F., Milgram, N. W., & Cotman, C. W. (2000). Region-specific age at onset of β -amyloid in dogs★. *Neurobiology of aging*, *21*(1), 89-96.
- Head, E., Mehta, R., Hartley, J., Kameka, M., Cummings, B. J., Cotman, C. W., ... & Milgram, N. W. (1995). Spatial learning and memory as a function of age in the dog. *Behavioral neuroscience*, *109*(5), 851.
- Hoffman, J. M., Creevy, K. E., Franks, A., O'Neill, D. G., & Promislow, D. E. (2018). The companion dog as a model for human aging and mortality. *Aging Cell*.

- Heckler, M. C., Tranquilim, M. V., Svicero, D. J., Barbosa, L., & Amorim, R. M. (2014). Clinical feasibility of cognitive testing in dogs (*Canis lupus familiaris*). *Journal of Veterinary Behavior: Clinical Applications and Research*, 9(1), 6-12.
- Hua, J., Hoummady, S., Muller, C., Pouchelon, J. L., Blondot, M., Gilbert, C., & Desquilbet, L. (2016). Assessment of frailty in aged dogs. *American journal of veterinary research*, 77(12), 1357-1365.
- Kaeberlein, M., Creevy, K. E., & Promislow, D. E. (2016). The dog aging project: translational geroscience in companion animals. *Mammalian genome*, 27(7-8), 279-288.
- Karagiannis, C. I., Burman, O. H., & Mills, D. S. (2015). Dogs with separation-related problems show a “less pessimistic” cognitive bias during treatment with fluoxetine (ReconcileTM) and a behaviour modification plan. *BMC Veterinary Research*, 11(1), 80.
- Kimotsuki, T., Nagaoka, T., Yasuda, M., Tamahara, S., Matsuki, N., & Ono, K. (2005). Changes of magnetic resonance imaging on the brain in beagle dogs with aging. *Journal of veterinary medical science*, 67(10), 961-967.
- Kis, A., Hernádi, A., Kanizsár, O., Gácsi, M., & Topál, J. (2015). Oxytocin induces positive expectations about ambivalent stimuli (cognitive bias) in dogs. *Hormones and behavior*, 69, 1-7.
- Landsberg, G. M., Nichol, J., & Araujo, J. A. (2012). Cognitive dysfunction syndrome: a disease of canine and feline brain aging. *Veterinary Clinics: Small Animal Practice*, 42(4), 749-768.
- Madari, A., Farbakova, J., Katina, S., Smolek, T., Novak, P., Weissova, T., Novak, M. and Zilka, N., 2015. Assessment of severity and progression of canine cognitive dysfunction syndrome using the CANine DEmentia Scale (CADES). *Applied Animal Behaviour Science*, 171, pp.138-145.
- Malenka, R. C., Nestler, E. J., & Hyman, S. E. (2009). Chapter 6: widely projecting systems: monoamines, acetylcholine, and orexin. *Snyder A, Brown RY. Molecular Neuropharmacology: A Foundation for Clinical Neuroscience*, 147-148.
- Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, 118(3-4), 161-181.
- Mendl, M., Brooks, J., Basse, C., Burman, O., Paul, E., Blackwell, E., & Casey, R. (2010). Dogs showing separation-related behaviour exhibit a “pessimistic” cognitive bias. *Current Biology*, 20(19).
- Miklósi, Á., Polgárdi, R., Topál, J., & Csányi, V. (1998). Use of experimenter-given cues in dogs. *Animal cognition*, 1(2), 113-121.
- Miklósi, Á., & Topál, J. (2011). On the hunt for the gene of perspective taking: pitfalls in methodology. *Learning & Behavior*, 39(4), 310-313.

- Milgram, N.W., E. Head, E. Weiner, E. Thomas, Cognitive functions and aging in the dog: acquisition of nonspatial visual tasks, *Behav. Neurosci.* 108 (1994) 57–68.
- Mongillo, P., Araujo, J. A., Pitteri, E., Carnier, P., Adamelli, S., Regolin, L., & Marinelli, L. (2013). Spatial reversal learning is impaired by age in pet dogs. *Age*, 35(6), 2273–2282.
- Neilson, J. C., Hart, B. L., Cliff, K. D., & Ruehl, W. W. (2001). Prevalence of behavioral changes associated with age-related cognitive impairment in dogs. *Journal of the American Veterinary Medical Association*, 218(11), 1787-1791.
- Neus Bosch, M., Pugliese, M., Gimeno-Bayón, J., Jose Rodriguez, M., & Mahy, N. (2012). Dogs with cognitive dysfunction syndrome: a natural model of Alzheimer's disease. *Current Alzheimer Research*, 9(3), 298-314.
- Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: The utility of a cognitive approach. *Neuroscience and Biobehavioral Reviews*, 29(3), 469–491. <http://doi.org/10.1016/j.neubiorev.2005.01.002>
- Paul, E. S., & Mendl, M. T. (2018). Animal emotion: Descriptive and prescriptive definitions and their implications for a comparative perspective. *Applied Animal Behaviour Science*.
- Piotti, P. (2017). Positive emotions and quality of life in dogs. *Animal Sentience: An Interdisciplinary Journal on Animal Feeling*, 2(14), 17.
- Piotti, P., Szabó, D., Wallis, L., Bognár, Z., Stiegmann, B.S., Egerer, A., Marty, P. and Kubinyi, E., 2017. The effect of age on visuo-spatial short-term memory in family dogs. *Pet Behaviour Science*, (4), pp.17-19.
- Pongrácz, P., Miklósi, Á., Dóka, A., & Csányi, V. (2003). Successful application of video-projected human images for signalling to dogs. *Ethology*, 109(10), 809-821.
- Pongrácz, P., Ujvári, V., Faragó, T., Miklósi, Á., & Péter, A. (2017). Do you see what I see? The difference between dog and human visual perception may affect the outcome of experiments. *Behavioural processes*, 140, 53-60.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Roelofs, S., Boleij, H., Nordquist, R. E., & van der Staay, F. J. (2016). Making decisions under ambiguity: judgment bias tasks for assessing emotional state in animals. *Frontiers in behavioral neuroscience*, 10, 119.
- Rofina, J.E., Van Ederen, A.M., Toussaint, M.J.M., Secreve, M., Van Der Spek, A., Van Der Meer, I., Van Eerdenburg, F.J.C.M. and Gruys, E., 2006. Cognitive disturbances in old dogs suffering from the canine counterpart of Alzheimer's disease. *Brain research*, 1069(1), pp.216-226.
- Schütt, T., Helboe, L., Pedersen, L. Ø., Waldemar, G., Berendt, M., & Pedersen, J. T. (2016). Dogs with cognitive dysfunction as a spontaneous model for early Alzheimer's disease: a

translational study of neuropathological and inflammatory markers. *Journal of Alzheimer's Disease*, 52(2), 433-449.

Szetei, V., Miklósi, Á., Topál, J., & Csányi, V. (2003). When dogs seem to lose their nose: an investigation on the use of visual and olfactory cues in communicative context between dog and owner. *Applied Animal Behaviour Science*, 83(2), 141-152.

Siwak-Tapp, C. T., Head, E., Muggenburg, B. A., Milgram, N. W., & Cotman, C. W. (2007). Neurogenesis decreases with age in the canine hippocampus and correlates with cognitive function. *Neurobiology of learning and memory*, 88(2), 249-259.

Siwak-Tapp, C. T., Head, E., Muggenburg, B. A., Milgram, N. W., & Cotman, C. W. (2008). Region specific neuron loss in the aged canine hippocampus is reduced by enrichment. *Neurobiology of aging*, 29(1), 39-50.

Starling, M. J., Branson, N., Cody, D., Starling, T. R., & McGreevy, P. D. (2014). Canine sense and sensibility: tipping points and response latency variability as an optimism index in a canine judgement bias assessment. *PLoS One*, 9(9), e107794.

Studzinski, C. M., Christie, L. A., Araujo, J. A., Burnham, W. M., Head, E., Cotman, C. W., & Milgram, N. W. (2006). Visuospatial function in the beagle dog: an early marker of cognitive decline in a model of human aging and dementia. *Neurobiology of learning and memory*, 86(2), 197-204.

Su, M.Y., Head, E., Brooks, W.M., Wang, Z., Muggenburg, B.A., Adam, G.E., Sutherland, R., Cotman, C.W. and Nalcioğlu, O., 1998. Magnetic resonance imaging of anatomic and vascular characteristics in a canine model of human aging. *Neurobiology of aging*, 19(5), pp.479-485.

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

Tapp, P. D., Siwak, C. T., Estrada, J., Head, E., Muggenburg, B. A., Cotman, C. W., & Milgram, N. W. (2003). Size and reversal learning in the beagle dog as a measure of executive function and inhibitory control in aging. *Learning & Memory*, 10(1), 64-73.

Tapp, P.D., Siwak, C.T., Gao, F.Q., Chiou, J.Y., Black, S.E., Head, E., Muggenburg, B.A., Cotman, C.W., Milgram, N.W. and Su, M.Y., 2004. Frontal lobe volume, function, and β -amyloid pathology in a canine model of aging. *Journal of Neuroscience*, 24(38), pp.8205-8213.

Vergheze, J., Annweiler, C., Ayers, E., Barzilai, N., Beauchet, O., Bennett, D.A., Bridenbaugh, S.A., Buchman, A.S., Callisaya, M.L., Camicioli, R. and Capistrant, B., 2014. Motoric cognitive risk syndrome Multicountry prevalence and dementia risk. *Neurology*, 83(8), pp.718-726.

Wallis, L. J., Virányi, Z., Müller, C. A., Serisier, S., Huber, L., & Range, F. (2016). Aging effects on discrimination learning, logical reasoning and memory in pet dogs. *Age*, 38(1), 6.

Wallis, L. J., Range, F., Kubinyi, E., Chapagain, D., Serra, J., & Huber, L. (2017, November). Utilising dog-computer interactions to provide mental stimulation in dogs especially during

ageing. In *Proceedings of the Fourth International Conference on Animal-Computer Interaction* (p. 1). ACM.

Waters, D. J. (2011). Aging research 2011: exploring the pet dog paradigm. *ILAR journal*, 52(1), 97-105.

Westfall, J., Judd, C. M., & Kenny, D. A. (2015). Replicating studies in which samples of participants respond to samples of stimuli. *Perspectives on Psychological Science*, 10(3), 390-399.

William, W., Kahloon, M., Fakhry, H., Ishak, W.W., 2011. Oxytocin role in enhancing well-being: a literature review. *J. Affect. Disord.* 130, 1–9.

Wood, L. S., Desjardins, J. K., & Fernald, R. D. (2011). Effects of stress and motivation on performing a spatial task. *Neurobiology of learning and memory*, 95(3), 277-285.