Department of Physics, Chemistry and Biology

Master Thesis

Labrador and German shepherd breed differences in dog-human communication

Anna Grozelier

LiTH-IFM- Ex-15/2998-SE

Supervisor: Pr Per Jensen, Dr Lina Roth, Linköping University

Examiner: Dr Karl-Olof Bergman, Linköping University

Department of Physics, Chemistry and Biology

Linköpings universitet

SE-581 83 Linköping, Sweden

Institutionen för fysik, kemi och biologi

Department of Physics, Chemistry and Biology

Datum/Date

2015-06-15

Språk/Language

Engelska/English

Rapporttyp
Report category

Examensarbete
D-uppsats

ISBN
LITH-IFM-A-EX—15/2998—**SE**

ISRN

Serietitel och serienummer **ISSN**
Title of series, numbering

URL för elektronisk version

Handledare/Supervisor Pr Per Jensen, Dr Lina Roth

Ort/Location: Linköping

Titel/Title:

Labrador and German shepherd breed differences in dog-human communication

Författare/Author:

Anna Grozelier

Sammanfattning/Abstract:

As our long-term companions, dogs' communication with us is perhaps the most developed of all human-animal ones. This study was aimed to investigate breed differences of German Shepherds and Labradors in dog-human communication. This was obtained through two tests: a problem-solving task and a pointing test. These two tests target both directions of communication: how much dogs understand and respond to the pointing and how they communicate with humans when facing a problem. Additionally, hair cortisol was measured in the dogs and dog owners filled a behavioural questionnaire (C-BARQ). The main breed difference I found was that Labradors performed better in both tests. I also found that the latency of the dogs' choices in the pointing test correlated with many factors, e.g. they chose quicker when: choosing correctly, when they had many physical contacts with the experimenter in the problem-solving task, when they were more intense, energetic dogs, when they had higher hair cortisol levels and when they had a confident body posture. This indicates that the latency of choice could depend on the confidence of the dog and on the trust in the experimenter as well as on energy level and focus ability. Overall, this study revealed a limited amount of breed differences, compared to a parallel study on Labrador types (hunting and show dogs), showing that intra-breed differences can be more important than inter-breed ones on a behavioural level.

Nyckelord/Keyword:

C-BARQ, communication, cortisol, dog breed, German shepherd, Labrador, pointing test, problem-solving test

Content

1 Abstract

As our long-term companions, dogs' communication with us is perhaps the most developed of all human-animal ones. This study was aimed to investigate breed differences of German Shepherds and Labradors in dog-human communication. This was obtained through two tests: a problem-solving task and a pointing test. These two tests target both directions of communication: how much dogs understand and respond to the pointing and how they communicate with humans when facing a problem. Additionally, hair cortisol was measured in the dogs and dog owners filled a behavioural questionnaire (C-BARQ). The main breed difference I found was that Labradors performed better in both tests. I also found that the latency of the dogs' choices in the pointing test correlated with many factors, e.g. they chose quicker when: choosing correctly, when they had many physical contacts with the experimenter in the problem-solving task, when they were more intense, energetic dogs, when they had higher hair cortisol levels and when they had a confident body posture. This indicates that the latency of choice could depend on the confidence of the dog and on the trust in the experimenter as well as on energy level and focus abilities. Overall, this study revealed a limited amount of breed differences, compared to a parallel study on Labrador types (hunting and show dogs), showing that intra-breed differences can be more important than inter-breed ones on a behavioural level.

Keywords: C-BARQ, communication, cortisol, dog breed, German shepherd, Labrador, pointing test, problem-solving test

2 Introduction

Dogs have been our companions for millennia, more than a hundred thousand years, some claim (Vila et al., 1997), while other research points to 11 to 16 thousand years (Freedman et al., 2014). This close relationship to humans has shaped them to understand us better and they are, for example, able to interpret human cues such as pointing, gazing and head turning (Reid, 2009). A recent study indicates that dogs do not perceive pointing as a command (Scheider et al., 2013), so it is possible that they view it as informative and thus have a true understanding of the gesture.

These outstanding communication abilities in a species morphologically and genetically very different from us started to attract the attention of behavioural scientists a few decades ago. Tests show that dogs understand human cues better than chimpanzees (Miklosi et al., 2004; Kirchhofer et al., 2012), which are a lot closer genetically to us. Indeed, chimpanzees fare better in competing tasks rather than cooperating ones (Hare and Tomasello, 2004) whereas dogs have become very good at cooperating tasks (Pettersson et al., 2011), both with humans and conspecifics (Ostojic and Clayton, 2014). The dogs' abilities have even been compared to those of human infants and toddlers in experiments in which they fared well (Lakatos et al., 2009).

But the human-dog communication goes both ways: dogs interact with us by looks, physical contacts and various human-directed behaviours. One way to study this is to present them with a problem-solving task (Persson et al., 2015). Indeed, presented with a problem to solve, dogs are known to look back to humans in reference, contrary to wolves or even dingoes (Miklosi et al., 2003; Smith and Litchfield, 2013). This looking back behaviour is considered as a form of help request from the dog. Studies revealed that similar behaviours such as gazing and gaze alternation were used as a communicative tool to request toys or food (Gaunet, 2010; Kaminski et al., 2011).

The dog's talent for cooperating with humans is thought to be linked with the selection for tame behaviours. This is supported by results of the silver foxes experiment, is which selection for low-fear of humans generated better skills to understand human cues as pointing (Hare et al., 2005). The emotional reactivity hypothesis states that selection on dogs lowered the stress and fear towards humans, enabling the dogs to feel as comfortable with humans as they do with conspecifics, or even more (Hare and Wrangham, 2002). However, this hypothesis, as well as others, is still debated. Another important hypothesis is the theory of mind: according to this hypothesis, men and dogs evolving together enabled the dogs to develop human-like social skills (Hare and Tomasello, 2005; Reid, 2009; Horowitz, 2011).

A number of studies have aimed to compare hand-raised wolves and dogs to investigate how much influence evolution and genetics have against environment. Experiments on hand-raised wolf puppies and dogs socialised in a similar manner show that the behavioural differences are due to a combination of genetics and social environment (Gacsi et al., 2009a; Udell and Wynne, 2010). Indeed, hand-raised wolves developed similar abilities as dogs for understanding pointing, but at a later age than dog puppies, and the wolves conserved some differences in behaviours like struggling, biting or latency of the search for human eye contact. Viranyi at al. (2008)

suggest that this difference in the search for eye contact is precisely the cause of the differences in understanding pointing signs.

In this project, I want to investigate more subtle differences, between dog breeds themselves. Comparisons between breeds for the use of human cues in pointing test are rare and the size of sampling small (McKinley and Sambrook, 2000; Wobber et al., 2009). The use of problem-solving device in not as common as pointing tests, the devices are varied and the focus is rarely on breed differences or on human-directed behaviours, but instead on the ability itself to solve the problem, or on the different ways to solve it depending on human cues (Shimabukuro et al., 2015). A meta-analysis on breed difference in pointing tasks gathering 14 studies (Dorey et al., 2009) did not identify a significant difference between breeds, but it might be due to several factors including heterogeneity of testing methods. Indeed, Wobber et al. (2009) found differences between working dog breeds and non-working ones, the working breeds being more skilled at interpreting human cues. Many dog breeds were historically shaped by humans with the desire to obtain animals with more specialized skills: hunting, retrieving, herding, guarding (Dorey et al., 2009). The Labrador and German Shepherd are classified in the same genetic cluster (mastiff-type dogs) by genotyping analysis (Parker and Ostrander, 2005), although Labradors are classed as both mastiff-type and hunting dogs. The American Kennel Club catalogues the two breeds in different utility groups, Labradors in the sporting group, German Shepherds in the herding group. The Swedish Kennel club also divides them in two groups: pasture and cattle dogs for the German Shepherd, and retrieving and water dog for the Labrador. However, the German Shepherd dogs (GSDs) have not been used for herding for decades and are nowadays usually pet or working dogs. Hence, we can suppose the selective breeding these past few decades has not been selecting herding specific characteristics. Our study gathered mainly companion dogs, thus we can speculate that we might find low breed differences between dogs bred for the same purpose. Research indicates that recent selective breeding has an important impact on the breed, probably more than the historical background (Svartberg, 2006). Hence, we can nowadays expect the German Shepherds to have little influence left of their herding background.

As a complement to the behavioural tests, this study also included a questionnaire (C-BARQ) completed by the owners. The C-BARQ (Canine Behavioural Assessment & Research Questionnaire) is a recognized behavioural and temperament questionnaire used to screen for behavioural problem, or used in behavioural research (Hsu and Serpell, 2003). Hair cortisol of the dogs was also assessed. Cortisol is a steroid hormone acting on many systems of the body, but it is widely recognized to be a marker of stress and activity (Pollard, 1995). Since the cortisol is incorporated in the hair as it grows, hair cortisol is an indicator of chronic stress (Russell et al., 2012) and a pertinent measure of baseline long-term levels of cortisol (Bennett and Hayssen, 2010; Bryan et al., 2013). Hair cortisol results will together with behavioural and questionnaire results increase our knowledge on possible differences between German Shepherd dogs and Labradors.

The aim of this project was to investigate behavioural breed differences between German Shepherd dogs and Labradors, especially in communication with humans. This study was performed in parallel with another project, investigating type differences within one breed: hunting versus show Labrador Retrievers. The two studies are complementary. Both studies included a pointing test and a problem-solving task. In the problem-solving task, I was particularly interested in whether the dogs would ask for help and if these human directed behaviours correlated with other parameters including e.g. the ability to solve the problem or success in choosing correctly in the pointing test and whether this would differ between breeds. The two breeds have different temperament, as defined by the American Kennel Club and the Swedish Kennel Club: Labradors are active, high energy and friendly dogs, German Shepherds are confident, smart and steady. We could then expect German Shepherds to perform better in the problem-solving test, thanks to their mental abilities.

3 Material & methods

3.1 Animals

The dogs were recruited through announcements on social media and from previous studies (in which three police dogs). The study was performed on 66 dogs, 30 Labradors and 36 German Shepherds. The sex distribution was homogeneous with 18 female and 18 male German Shepherds, 16 female and 14 male Labradors. The age ranged between less than a year to 14 years, was normally distributed for each breed and found by a t-test to have no significant difference between breeds (German Shepherd dogs 3.08 ± 0.67 years, ± SEM, Labrador dogs 4.23 ± 0.42 years).

For each test, a few dogs were excluded of analysis due to non-compliance with the test, did not pass the motivation test (see Methods), or few data points (less than half) in the pointing test (see Methods). For the problem-solving task, two dogs were excluded, one of each breed. For the pointing test, eight dogs were excluded, four of each breed.

3.2 Location and material

The behavioural tests were performed outdoors, at Linköping University, Sweden, in marquee tent (3 x 3 m, Biltema, ref: 14-318) without flooring, with 3 fabric walls and a small fence closing the open side. The tent was used to achieve a uniform environment similar to the other dog project on Labrador types conducted in different locations all over Sweden. The tests were recorded with a full HD camcorder (Canon Legria HF G25). The treats used for the test were dog food pieces of Frolic complete (beef, carrot and cereals kind). Each ring of Frolic was cut in four equal parts, each used as one treat. In case of allergy, either the owner brought their own treats or pieces of sausages were provided.

The problem-solving task was performed with a testing apparatus (Figure 1A) comprised of PVC slabs screwed together, one for the bottom and others defining three pits containing treats. The pits are covered with sliding Plexiglas lids with holes for olfaction. The pits are aligned, and only the side ones are readily accessible, the middle one is unsolvable (closed by screws). A small plate similar to the pits but without sliding lid is used for the motivation test (Fig. 1B, see Methods). For the pointing test only two black buckets were used (21 cm height, 19 cm diameter).

Figure 1 A: The problem-solving test setup (54.8 x 24.8 cm) with treats, side slabs slightly open. B: The motivation test plate (15 x 9.8 cm).

3.3 Methods

The dogs performed two tasks that were video recorded for later analysis: a problem-solving task and a pointing test.

3.3.1 Problem-solving test

Motivation tests were performed before both the problem-solving and the pointing tests, to ensure that the dog was indeed interested in the treats offered and would eat it in a test context. This was assessed for the problem-solving with a small PVC plate, similar to one of the three plates on the problem-solving apparatus (Figure 1B) but without a sliding lid. The motivation test consisted of offering the dog a treat on the single plate and if the dog ate it this was repeated three times in total. The dogs that passed the motivation test were allowed to participate the problem-solving test. The motivation test was also used to get a feeding score (see Ethogram table 1 and Appendix).

In the problem-solving test, the experimenter and owner stood opposite to each other, each in one corner of the tent (the corners on the fenced side), and the test apparatus was placed in the middle back part of the tent (Figure 2). The dog was released in the tent for three minutes, while the owner and experimenter were to remain neutral and gaze towards the test setup. The experimenter controlled time with a stopwatch. If the dog had not opened any pit and received treat one minute after release, the experimenter opened the two sliding slabs on the sides halfway, enough so that the dog almost had access to the treats.

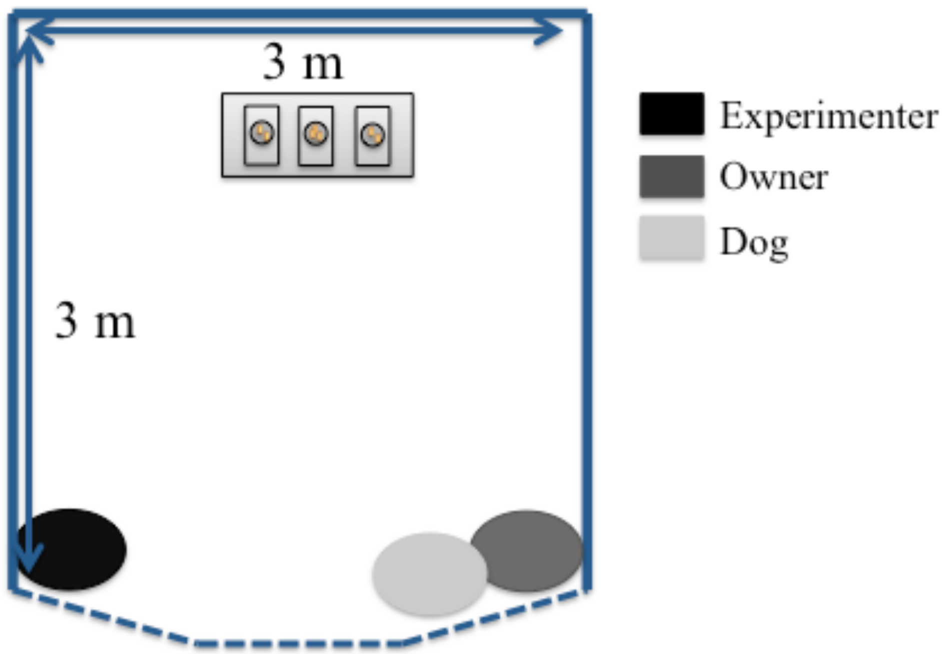

Figure 2: Problem-solving disposition of experimental setting

3.3.2 Pointing test

For the pointing test, the motivation test consisted of offering treats in the buckets, also three times in a row. The experimenter alternated between buckets each time and offered them to the dog while saying “varsågod”. This enabled the dogs to associate the buckets and the “varsågod” with treats. The dogs that passed the motivation test and ate all three treats were allowed to participate the pointing test.

The second test was a pointing test with a procedure similar to Gacsi et al. (2009b). One treat was hidden in one of two identical buckets. The experimenter turned her back towards the dog and owner while doing so to prevent them from seeing which bucket was hiding a treat. Prior to testing, the buckets were both baited with treats to ensure they both smelled similarly and thereby minimising the risk that the dogs could use their sense of smell while choosing. The fence closing the tent was opened in the middle, defining the area the dog and the owner were to stand in the beginning of the pointing task. The owner stayed with the dog in front of the tent (Figure 3). The owner had instructions not to look at the experimenter, but at the dog, in order not to see the pointing. This should prevent the owner from influencing the dog’s choice. If needed, the owner held the dog by the collar, the neck or leash to prevent him/her from approaching the experimental setup before the releasing signal was given. A long lead was also provided to the owner if he/she preferred to keep a hold of the dog at all times and to ensure the experimenter’s security in case of a more jumpy dog. The experimenter established visual contact with the dog, attracting its attention with mouth sounds or calling its name. The experimenter then pointed with the whole arm (Figure 3) towards the bucket containing the treat for about 2 seconds and then went back to neutral arm position and said “varsågod” as a signal for the dog to be released and make its choice. If the dog did not go at this signal, the owner had to repeat “varsågod” him- or herself or release the dog by a sign.

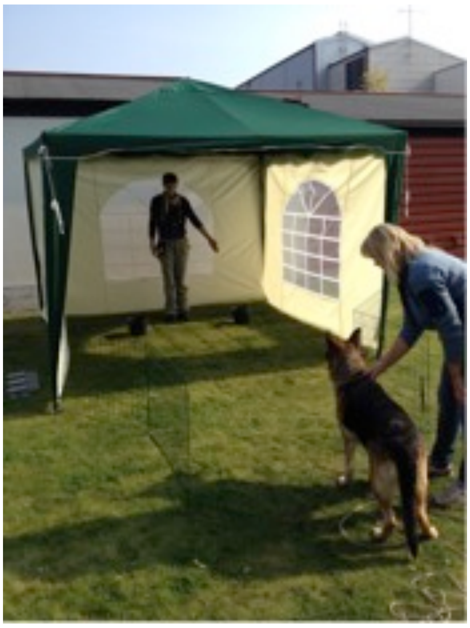

Figure 3. The pointing test with the experimenter pointing towards a bucket.

I considered that the dog had made a choice when his/her nose came at a distance of a few centimetres from one bucket. When a choice was made, the experimenter removed the bucket that was not chosen. This was to prevent the dog from having access to both buckets. Hence the dog could only access the treat if he/she chose the correct bucket first. If the dog spent more than 3 seconds in front of the experimenter or moved around the tent without approaching any buckets for more than 5 seconds, it was considered a “no choice”. The pointing test was repeated 20 times, with a short break after the 10 first. Two random orders of pointing were generated to alternate between buckets (see Appendix). The order was generated so to have no more than two points in the same direction. For each dog, one order was used for the first session, the other for the second one. The order used for first and second session were alternated between each dog being tested so that two contiguous dogs would have different orders of points.

3.3.3 Additional data outside the tests

The dogs’ long-term cortisol levels were assessed by hair cortisol analysis and hair samples were collected directly after the behavioural tests. The hair samples were cut from the middle of the neck of the dogs, cutting the hair as close to the skin as possible without hurting the dogs. I took part in some steps of the cortisol analysis, which was performed with a methanol extraction and Radioimmunoassay (RIA) measure, protocol (see Appendix) developed at the Department of Medical and Health Sciences, Linköping University (Morelius et al., 2004; Karlén et al., 2011).

All the owners also filled a Canine Behavioral Assessment and Research Questionnaire, C-BARQ (Hsu and Serpell, 2003) with 31 additional questions on the dog's background. The C-BARQ is a questionnaire with guidelines to calculate from clusters of questions scores for temperament parameters. The C-BARQ used in this study is a slightly modified Swedish version adapted by Dr Kenth Svartberg, Stockholm University, Sweden. The questionnaire results were analysed according to Serpell's and Svartberg's instructions and a publication (Duffy et al., 2008), to calculate overall scores for different behavioural traits. Only the scores relevant for my hypotheses were further used i.e. the scores for trainability, excitability, stranger-directed aggression and attachment/attention seeking behaviour.

3.4 Data collection from recordings

Like mentioned, all tests were recorded by videotape. The behavioural results from the problem-solving test were extracted with the Noldus software Observer XT 10.5, recording the behaviours detailed in the Ethogram (Table 1). This Ethogram is an adapted version of the one used in a previous study using the same test apparatus (Persson et al., 2015). The initial list of behaviours included barking, but it was excluded since only five individuals presented the behaviour. The duration and frequency of each behaviour from the 3 minutes of test were recorded with a continuous sampling method. For the problem-solving itself, I extracted the latency for solving each problem and if the dog solved each of them. To ensure good reliability and low inter-observer variation, the two students in charge of the dog projects analysed together two dogs with the Observer software, one dog from each project. The two students then scored one additional dog independently and results were visually compared and were found to match.

Table 1. Ethogram of the recorded behaviour during the problem-solving test

Behaviour group	***Behaviour***	***Definition***
Position	*Test setup*	*The dog's head is within its own body length of the test setup*
	Experimenter	*The dog's head is within its own body length of the experimenter*
	Owner	*The dog's head is within its own body length of the owner*
	Elsewhere inside	*The dog's head is not within its own body length of either the test setup, the experimenter or the owner*
	Outside	*The dog has at least its head and shoulders outside the tent (in case of an escape)*
Activity	*Standing*	*The dog is standing with its four paws touching the ground*
	Sitting	*The dog is sitting down, its buttocks touching the ground*
	Lying down	*The dog is lying down, its belly touching the ground*
	Walking	*The dog is lifting its paws and is moving forward on the horizontal plan*
	Movement	*The dog is lifting its paws without moving forward*
Test	*Test setup interactions*	*Physical interactions with the test setup*
	First problem solved	*The dog accessed and ate the treats in one of the pits of the test setup*
	Second problem solved	*The dog accessed and ate the treat in the other pit of the test setup*
Human interactions	*Eye contact experimenter*	*The dog is either positioned at the test setup, between the test setup and the experimenter or at the experimenter while gazing towards the face of the experimenter.*
	Eye contact owner	*The dog is either positioned at the test setup, between the test setup and the owner or at the owner while gazing towards the face of the owner.*
	Physical contact experimenter	*The dog is positioned close to the experimenter and in physical contact.*
	Physical contact owner	*The dog is positioned close to the owner and in physical contact.*
Other	*Feeding Score*	*A score from 1 to 3 (late to early feeding) explaining the time it took for the dogs to eat the treats in the initial motivation test. See Appendix for details.*
	Body Posture Score	*A score from 1 to 5 (high to low) of the overall body posture. See Appendix for details.*
	Escape attempt	*The dog's nose is touching the bottom of the tent's wall*

The problem-solving test was also used to record transition indexes and the dog's intensity, which were both scored from the movies with the movie player software VLC (Version 2.1.5). Firstly, three transition indexes were scored: when the dog walked directly between the owner and the test setup, between the experimenter and the test setup, or between the owner and experimenter. Scored transitions had to be direct, take less than 5 seconds to ensure the dog's intention (a dog stopping midway to sniff for example was not recorded as a transition). We considered direct transitions between a human and the test to be a help seeking behaviour, similar to gaze alternation.

I defined the dog's intensity score by studying the movements, pace, activity of the dog as well as the level of interaction with the test setup (Ethogram of this measurement in Appendix). Since this measure was subjective, scorings of 20 dogs were made by three observers to ensure high inter-evaluator reliability. The three observers were the two students in charge of the related dog projects mentioned before, plus another ethology student. The criteria of scoring were explained in detail to the other student. Plots of all scores showed more variability between the supplementary observer and the other observers than between the two students in charge of the dog projects. A Spearman correlation test was performed between the results of the two dog project leaders (as the data was found not to be normally distributed by a Shapiro test). The Spearman test found a significant correlation ($p < 0.05$) with a coefficient of $r = 0.70$. Considering the subjectivity of the measure, we deemed the correlation level high enough to go on with the scoring of all dogs and to include the parameter in our analysis.

The pointing test's data was extracted with the Mac video editing software iMovie (Version 10.0.6) to obtain the choices and their latencies. The latency of choice is defined as the time between the last "varsågod" (or the last sign that released the dog) and the choice itself (when the dog's nose was a few centimetres away from a bucket). In the analysis, I chose to eliminate the "no choice" occurrences, according to Hare et al. (2010), as they would artificially increase the number of wrong choices while they reflect something different (such as lack of interest for the test, or a lack of understanding).

3.5 Data analysis

The statistics were performed using the statistical packages R (version 3.1.0 built for Mac OS Mavericks) and SPSS (version 22.0.0.0), bar plot graphs were generated with Graphpad Prism (version 6.0).

The data were checked for normal distribution by plotting each behaviour's distribution against a normal curve. In the cases where the data was not normally distributed, I tried transforming the data by logarithm to base 10 or by square root. If neither logarithm nor square root generated normal distribution, the according parameters were analysed with non-parametric tests. Descriptive statistics (mean, standard error of the mean, maximum and minimum) were obtained for each parameter.

General Linear Model (GLM) tests were performed on the normally distributed data, using breed and sex as fixed factors to check if the parameters varied between the two breeds and the two genders.

I used Mann-Whitney U tests on the non-normally distributed data to compare each parameter between breed and sex.

I also tested if the dogs' performance in the pointing test was above chance level. This was tested with a one-way t-test against a mean accuracy of 0.5.

In addition, Spearman correlation tests (since not all the behaviours were found to be normally distributed) were performed on a global scale throughout all scaled parameters. All correlations were studied and I selected the ones of interest. Correlations between behavioural items that were strongly depending on each other e.g. time spend in the test setup zone positively correlated with the time spend interacting with the test were excluded of the analysis since it does not reveal anything new.

4 Results

Table 2. Table of significant results and tendencies. PST: Problem-Solving test, PT: Pointing test.

From	Behaviour differing	P-value	U/F value	Difference between	Mean
PST	Latency to solve first problem	0.095	U= 387	GSD	118±11s
				Labrador	91.8±12s
PST	Transitions in GSD	0.020		Experimenter-test	2.52±0.35
				Owner-test	1.57±0.21
PST	Physical contact with owner	0.087	U= 402	Males	0.44±0.26s
				Females	1.62±0.47s
PST	Time interacting with test setup	0.013	$F_{(1,63)}$= 6.52	GSD	67.0±9.5s
				Labrador	37.6±6.1s
PST	Time interacting with test setup	0.042	$F_{(1,63)}$= 4.32	Males	61.4±8.7s
				Females	40.4±7.1s
PST	Intensity scores	0.012	U= 326	GSD	3.9±0.2
				Labrador	3.1±0.2
PST	Time spent moving	0.010	$F_{(1,63)}$= 12	GSD	34.3±3.9s
				Labrador	57.7±6.0s
PST	Time spent lying down	0.060	U= 387	GSD	11.07±3.7s
				Labrador	4.77±2.2s
PST	Escape attempt	0.0002	U= 230	GSD	4.0±0.4
				Labrador	1.7±0.4
PST	Time spent moving	0.010	$F_{(1,63)}$= 12	Males	51.3±6.1s
				Females	38.6±4.1s
PST	Time spent sitting	0.012	U= 339	Males	5.64±2.15s
				Females	16.3±3.9s
PST	Time spent in owner zone	0.010	$F_{(1,63)}$= 7.1	Males	22.5±3.5s
				Females	46.6±6.8s
PST	Time spent in test setup zone	0.021	$F_{(1,63)}$= 5.6	Males	78.0±8.6s
				Females	53.2±7.7s
PT	Accuracy second session	0.038	$F_{(1,57)}$= 4.5	GSD	0.65±0.03
				Labrador	0.74±0.03
PT	Accuracy both sessions combined	0.040	$F_{(1,57)}$= 4.4	GSD	0.62±0.03
				Labrador	0.71±0.03
PT	Latency of choice	0.060		Correct choice	2.1±0.04s
				Wrong choice	2.0±0.03s
Other	Hair cortisol	0.012	$F_{(1,65)}$= 6.6	GSD	14.9±2.0 pg mg^{-1} hair
				Labrador	23.5±4.5 pg mg^{-1} hair
Other	C-BARQ stranger aggression	0.010	U= 733	GSD	0.48±0.1
				Labrador	0.15±0.05

4.1 Problem-solving test

4.1.1 Performance in the test

The latency to solve the first problem in the problem-solving test showed the tendency of a difference between breed (U= 387, p= 0.095): German Shepherds (118 ± 11 s) tended to be slower to solve the problem than Labradors (91.8 ± 12 s). The feeding score showed no significant differences between breeds or sex.

4.1.2 Human attention: physical and visual contact

Although neither of the transition scores differed between breeds, the German Shepherds had significantly more owner-test transitions than experimenter-test transitions (Figure 4, 2.52 ± 0.35 and 1.57 ± 0.21 respectively; p= 0.02), a result that was not found in Labradors (p= 0.76).

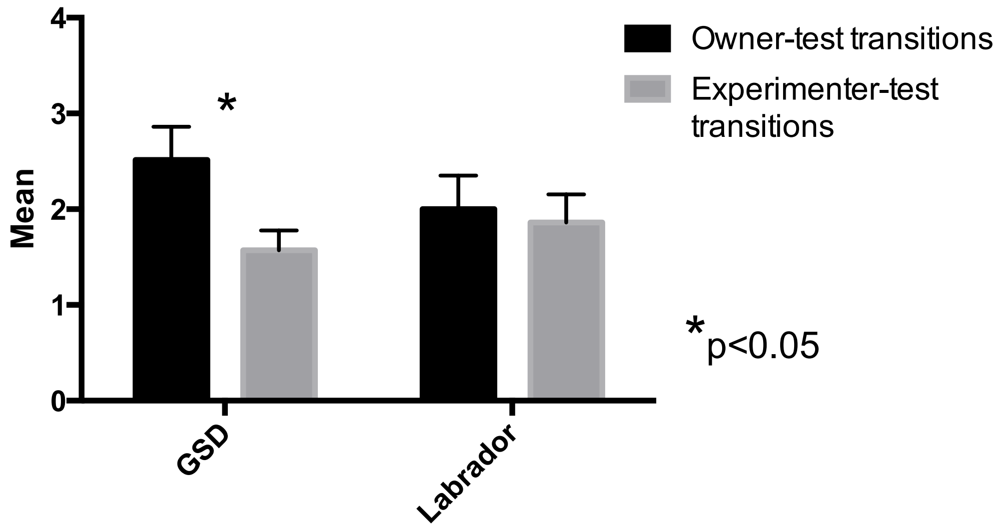

Figure 4. Mean number of transitions between human (owner or experimenter) and test setup for both breeds in the problem-solving experiment.

I found no breed difference on the duration or frequency of physical contact, neither with the owner nor the experimenter. However, females in general had a tendency towards more physical contact with their owner than males (1.62 ± 0.47 s and 0.44 ± 0.26 s respectively; U= 402, p= 0.087).

No breed or sex difference was found for the duration or number of occurrences of eye contact, either with the owner or with the experimenter.

4.1.3 Test interaction

In the problem-solving test, Labradors interacted significantly longer with the test setup than German Shepherds (67.0 ± 9.5 s and 37.6 ± 6.1 s respectively; $F_{(1,63)}$= 6.52, p= 0.013).

The duration of physical interaction with the test setup during the problem-solving test significantly differed between sexes ($F_{(1,63)}$= 4.32, p= 0.042): males interacted longer with the test than females (61.4 ± 8.7 s and 40.4 ± 7.1 s respectively).

4.1.4 Intensity, movements, position

The intensity scores from the problem-solving test significantly differed between breeds (U= 326, p= 0.012): the Labradors were more intense than German Shepherds (3.1 ± 0.2 and 3.9 ± 0.2 respectively). Accordingly, Labradors spent more time moving in the problem-solving test than German Shepherds (57.7 ± 6.0 s and 34.3 ± 3.9 s respectively; $F_{(1,63)}$= 12, p= 0.01). Furthermore, the German Shepherds tended to spend more time lying down than Labradors (11.0 ± 3.7 s and 4.7 ± 2.2 s respectively; U= 387, p= 0.06). Nevertheless, the German Shepherds tried to escape significantly more often than Labradors (4.0 ± 0.4 times and 1.7 ± 0.4 times respectively; U= 230, p = 0.0002).

In the problem-solving test the duration of movement significantly differed between sexes ($F_{(1,63)}$= 12, p= 0.01): males spent more time moving than females (51.3 ± 6.1 s and 38.6 ± 4.1 s respectively). Accordingly, a significant difference between sexes was found for the time spent sitting down (U= 339, p= 0.012): females sat longer periods of time in total than males did (16.3 ± 3.9 s and 5.64 ± 2.15 s respectively). In addition, females were shown to spend more time in the owner zone than males (46.6 ± 6.8 s and 22.5 ± 3.5 s respectively; $F_{(1,63)}$= 7.1, p= 0.01). Males were instead found to spend significantly longer time in the test setup zone than females (78.0 ± 8.6 s and 53.2 ± 7.7 s respectively; $F_{(1,63)}$= 5.6, p= 0.021).

4.2 Pointing test

The results of the pointing test were analysed in a one-way t-test to determine if the dogs' accuracy of choice was above chance. The tests showed that the dogs did perform above chance on a group level, for each pointing session (first session: 0.63 ± 0.03, $t_{(57)}$= 5.0, p= 0.000; second session: 0.69 ± 0.02, $t_{(57)}$= 8.5, p=0.000) and overall (0.66 ± 0.02, $t_{(57)}$= 7.8, p=0.00). However, Labradors showed significantly better pointing accuracy in the pointing test for the second session than the GSDs (Figure 5, 0.74 ± 0.03 and 0.65 ± 0.03 respectively, $F_{(1,57)}$= 4.5, p= 0.038) and for both sessions combined (Labradors 0.71 ± 0.03, GSD 0.62 ± 0.03, $F_{(1,57)}$= 4.4, p= 0.04).

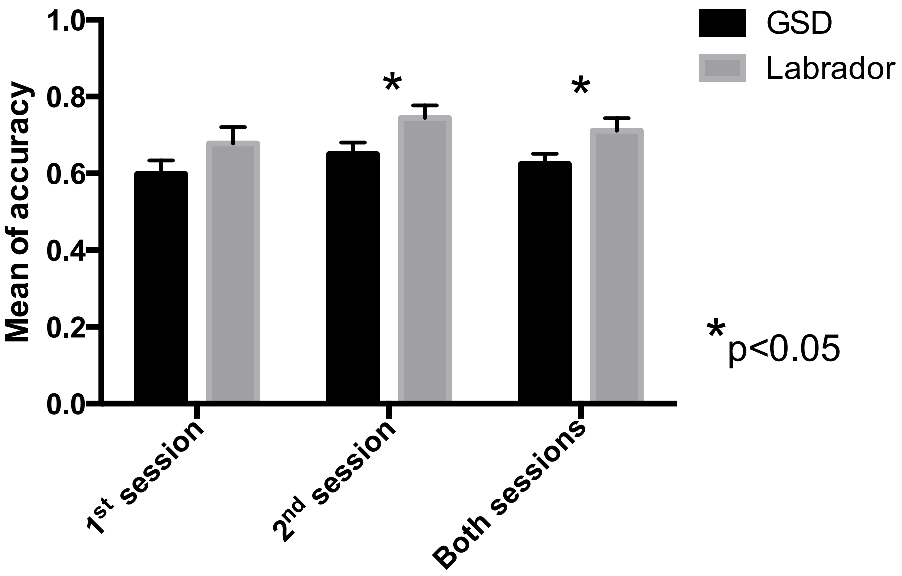

Figure 5. Mean of accuracy in pointing test for both breeds.

A t-test was performed for the latency of choice in the pointing test between correct and wrong choice: the dogs tend to choose quicker (p= 0.06) when choosing correctly than when choosing wrong (2.0 ± 0.03 s and 2.1 ± 0.04 s respectively) but no breed difference was found for the latency of choice in the pointing test, for either session or both.

4.3 Stress and activity: hair cortisol

The breeds significantly differed in hair cortisol levels ($F_{(1,65)}$= 6.6, p= 0.012): Labradors have significantly higher cortisol levels than German Shepherds (23.5 ± 4.5 pg mg^{-1} hair and 14.9 ± 2.0 pg mg^{-1} hair respectively). No sex difference was found for the hair cortisol levels.

4.4 C-BARQ questionnaire

Only one of the C-BARQ scores tested presented a significant variation between breed, none between sex. The stranger directed aggression differed between breeds: German Shepherds had significantly higher scores than Labradors (0.48 ± 0.1 and 0.15 ± 0.05 respectively; U= 733, p= 0.01).

4.5 Correlations

Spearman's correlations were performed across all scaled parameters from both tests and the parameters outside the test, such as cortisol, age and questionnaire results.

Table 3. Significant correlations between parameters from the questionnaire, the problem-solving experiment and the pointing test.

	Correlated parameters	r	P-value
Age of dogs	Number of physical contacts with owner	0.28	0.03
	Duration of physical contacts with owner	0.29	0.02
Number of physical contact with owner	Accuracy of choice first session	-0.29	0.03
	Accuracy of choice second session	-0.30	0.04
	Accuracy of choice both sessions combined	-0.34	0.01
Duration of physical contact with owner	Accuracy of choice second session	-0.30	0.02
	Accuracy of choice both sessions combined	-0.33	0.01
Cortisol	Duration spent outside the tent	-0.27	0.03
	Latency of choice in pointing first session	-0.34	0.01
	Latency of choice in pointing second session	-0.38	0.003
	Latency of choice in pointing both sessions combined	-0.40	0.002
Intensity	Latency of first problem solved	0.31	0.01
	Latency of choice in pointing first session	0.53	0.00
	Latency of choice in pointing second session	0.59	0.00
	Latency of choice in pointing both sessions combined	0.58	0.00
Duration spent outside the tent	Owner-test transitions	0.34	0.006
	Number of looks to the owner	0.27	0.03
Escape attempts	Number of looks to the owner	0.28	0.03
	Number of times in the owner zone	0.54	0.00
	Number of times in the test setup zone	0.50	0.00
	Number of times in the experimenter zone	0.27	0.03
Body posture	Latency of choice in first pointing session	0.28	0.04
	Latency of choice in second pointing session	0.29	0.03
	Latency of choice for both pointing sessions	0.29	0.03
Number of physical contacts with experimenter	Latency of choice in first pointing session	-0.30	0.03
	Latency of choice in second pointing session	-0.27	0.04
	Latency of choice for both pointing sessions	-0.32	0.02
Experimenter-test transitions	Latency of choice in first pointing session	-0.36	0.005
	Latency of choice in second pointing session	-0.28	0.03
	Latency of choice for both pointing sessions	-0.38	0.003

A significant positive correlation was found between age and physical contact with the owner: the younger the dogs, the more they had physical contact with their owner (Table 3). In addition, the frequency and duration of physical contact with the owner correlated negatively with the accuracy of choice in the pointing task (Figure 6): the more physical contact with their owner the dogs had, the less accurate they were in the pointing task.

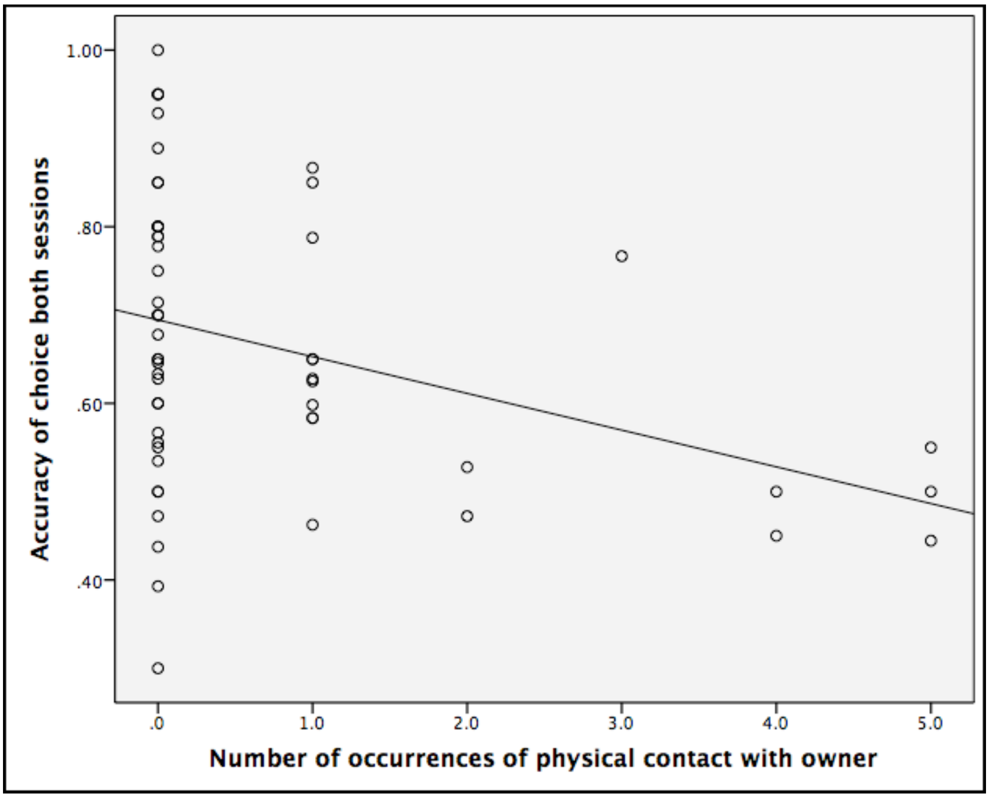

Figure 6. Accuracy of choice for both sessions combined in the pointing test in function of the number of occurrences of physical contact with the owner in the problem-solving experiment.

Several correlations were found between hair cortisol levels and other parameters (Table 3). Cortisol level was negatively correlated with the duration spent outside the tent during the problem solving: the lower the cortisol levels, the more time they spent outside. Cortisol also correlated negatively with the latency of problem-solving: the higher the cortisol levels, the quicker the dogs were to solve the first of the problems.

The intensity score positively correlated with the latency of problem-solving (considering a high score was actually a low intensity dog): the less intense the dog was, the more time it needed to solve the problem. Intensity score was also positively correlated with the latency of choice of first, second and both sessions combined in the pointing test: the more intense the dog was, the shorter latency of choice (Table 3).

The duration spent outside the tent during the problem-solving test correlated with the number of looks at the owner and the owner-test transition: the more the dog looked at the owner and moved between owner and test setup, the more time it spent outside.

The body posture score correlated positively to the latency of choice in each pointing sessions and in both sessions combined in the pointing test: the higher body posture score (high body posture score means a more insecure and frightful dog accordingly to the scale; see in Methods and Ethogram in Appendix), the longer time the dogs needed to choose.

The duration of physical contact with the experimenter in the problem-solving test correlated negatively with the latency of choice during the pointing test (Figure 7): the more physical contact the dogs had with the experimenter during the problem-solving test, the quicker they were to make their choice in the pointing test. Similarly, the experimenter-test transitions in the problem-solving test correlated negatively with the latency of choice in the pointing test: the more experimenter-test transitions the dogs performed in the problem-solving test, the quicker they were to choose in the pointing test.

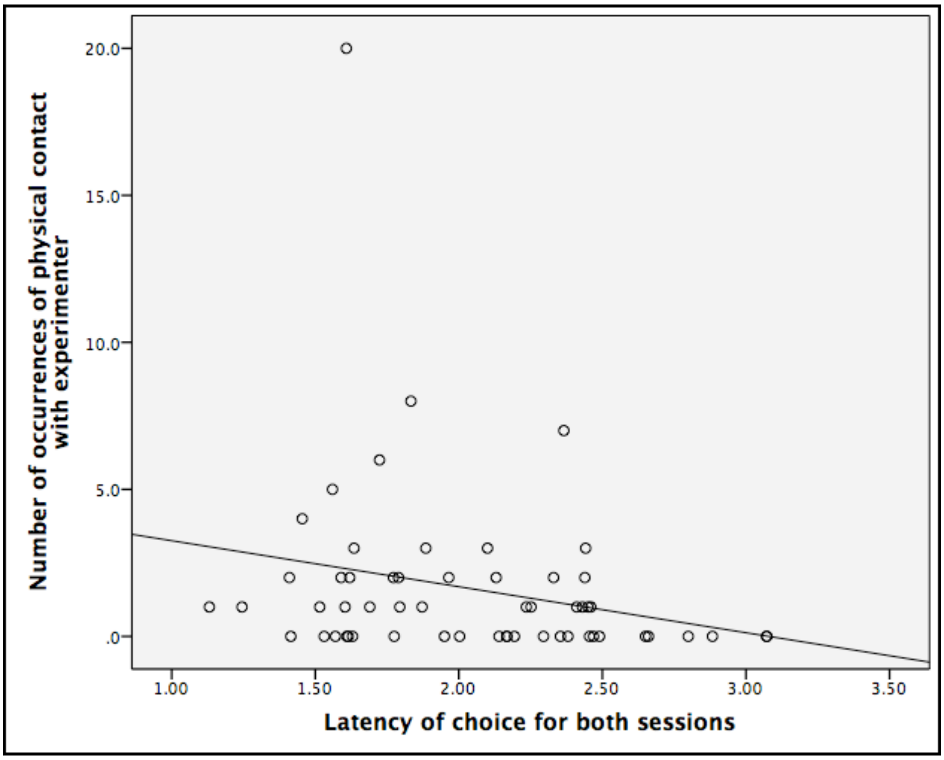

Figure 7. Number of occurrences of physical contact with the experimenter in the problem-solving experiment in function of the latency of choice for both sessions combined in the pointing test.

5 Discussion

The aim of this study was to investigate breed differences associated with dog-human communication, through two tests aimed to target both directions of communication. Interestingly, Labradors performed better in the pointing test and the problem-solving test, while German Shepherds escaped the tent more and were less active.

Overall, the Labradors performed better in both tests: they solved the problem-solving quicker and were more accurate in the pointing test. Some dog owners argued against using treats as motivation for German Shepherd dogs (as German Shepherds are considered to be less food motivated than Labroadors), but the feeding score did not differ significantly between breeds. Hence, I do not consider food motivation as being a factor in the breed difference in the general test performance. However, the low accuracy in the pointing test could be a result of the German Shepherds' suspiciousness towards strangers (Wilsson and Sundgren, 1997), since the experimenter was a stranger to the dogs. This speculation could also explain the difference shown between experimenter-test transitions and owner-test transitions and the difference in stranger-directed aggression revealed by the C-BARQ questionnaire. I then hypothesize that the lower performances of German Shepherds could be due to their lack of trust in strangers, a distrust that might be reinforced by the times they make a wrong choice and are prevented by the experimenter to access the treat in the other bucket in the pointing test. German Shepherds are often used as guarding or defence dogs, a purpose for which the distrust of strangers is an advantage. Hence this behavioural characteristic is likely to be the produce of purposeful selection.

In the problem-solving, the time spent outside correlated with the owner-test transitions and the number of looks to the owner. One possible explanation could be that the dogs chose to escape when they did not manage to get access to the treats and their owners did not respond to their help-seeking behaviour.

Dogs performed above chance on a group level in the pointing test, which coincides with other studies (Soproni et al., 2002; Viranyi et al., 2008). Studies have shown various breed groups perform over chance level (Dorey et al., 2009; Gacsi et al., 2009b).

Predictably, the latency of choice in the pointing test correlated positively with the intensity of the dog and negatively with the hair cortisol levels, more intense and higher cortisol dogs chose quicker. Results pointed to the fact that the dogs had a lower latency of choice in the pointing test when they were confident dogs, as shown by the body posture score correlation, and also a lower latency when they were doing the correct choice. Interestingly, the latency of choice also correlated with the duration of physical contact with the experimenter, indicating that the latency of choice could also depend on the trustfulness of the dog in strangers, like the experimenter, and on the confidence in one's choice. Moreover, the physical contacts with the owner are correlated with a lower accuracy of choice in the pointing test. This could be explained with a similar reasoning: dogs that have more physical contact with the owner might have quite exclusive personalities, or have been trained to obey and listen to the owner but ignore others. In which case they would be likely to grant less attention to the experimenter. Indeed, during the pointing test, it was hard to obtain some dogs' attention, as they were constantly trying to look at their owner. Moreover, studies have found that dogs grant inter-specific attention depending notably on the nature of the dog-human relationship (Mongillo et al., 2010), and familiarity was found to have an effect on the reaction of dogs to various test situations (Kerepesi et al., 2015; Shimabukuro et al., 2015). This second paper investigated the effect of familiarity in problem-solving for the same breeds that I studied, but in that case, either a familiar or unfamiliar person was present during the test, whereas both were present in my test. During that study, they did find breed difference but the methods were very different, as each dog had several sessions with the test setup.

The hair cortisol levels correlated negatively with the latency of solving the first problem in the problem-solving test. Even though this measurement might not reflect the cortisol level during the test, it's an indication of the overall levels of energy or stress. Therefore, these results could be due to cortisol and stress making dogs more focused. Indeed, studies in children affected by attention-deficit disorder found they have lower cortisol levels (Isaksson et al., 2012).

However, the difference between breeds in both intensity and cortisol reflect the difference of profile of the breeds: the American and Swedish Kennel Clubs define Labradors as very active, athletic and high-spirited dogs, while the German Shepherds are defined by their confidence and steadiness. The German Shepherd dogs escaped more than Labradors and they had lower hair cortisol levels. A negative correlation was found between cortisol and time spent outside the tent. Therefore, I hypothesize that escaping the tent could be considered as a coping strategy. Coping strategies are defined as different ways of coping with a situation, commonly a stressor. According to the personality of the individual, he/she might adopt a more tangible and bold way of dealing with the stressor. This corresponds to the fight-or-flight concept and is defined as proactive coping style. The reactive coping style is a more passive one, like a freezing response (Koolhaas et al., 1999). Studies have shown that different coping styles lead to differential activation of the HPA-axis; this results in higher levels of corticosterone in reactive coping animals and inversely proactive coping animals have lower levels of corticosterone (Carere et al., 2010; Boulton et al., 2015). Hence, the dogs which are not escaping would present reactive coping mechanisms, enduring high stress without acting on it while the dogs who are escaping show more proactive coping mechanisms and are less stressed.

Generally, I observed a relatively low difference between breeds. I hypothesize that this low difference could be due to the predominance of companion dogs in the study. Again, this seems to correlate with the study indicating recent breed history has more influence than more ancient roots (Svartberg, 2006). We can speculate that breeding for companionship has a stronger influence on individual dogs than the breed (although we cannot rule out the influence of training and life experience, having no genetic analysis coupled with the study). This is also confirmed by the other dog thesis on Labrador types, which found clearer differences between types of one breed bred for different purposes than I did between two breeds.

Sex differences were not the focus of this study, but the tendency towards more physical contact, here with the owner, for the females corroborate findings from recent studies using the very same experimental setup (Persson et al., 2015). It is also endorsed by the fact that females spent more time in the owner zone while males spent more time in the test zone and interacted more with the problem-solving test setup.

5.1 Societal & ethical considerations

Not without reason is the dog called man’s best friend. Studies show the impact of pet ownership and contact with animals on human wellbeing. Many owners go as far as to consider their pet dog as a family member and to treat them as such. This is to say how much dogs matter in our society. Shops sprout in cities selling dog clothing, dog hotels and spas appear to provide similar services as humans have access too. Overall, this anthropomorphic drift should not make us forget that dogs are not humans and should not be treated as such, for their own wellbeing. This is why animal and dog research is important, to understand them better and provide them with a better living environment.

Protecting animal welfare starts from the very beginning of their life. In the wild, animal breeding is controlled by the interaction of many factors like mate competition, natural selection notably with predation or simply the death of weaker animals due to resource competition or to organ or system failure in presence of deleterious mutations. All these factors, when the populations are big enough, enable to conserve enough genetic variation and to eliminate deleterious mutations. In domestic animals, most of these parameters play no role; humans' choices rule the breeding. And depending on the aim and focus of humans, this can lead to neglecting important factors in the animals' welfare. This is especially important in animal agriculture, where the aim is to maximize profit, e.g. produce bigger animals quicker. Easy examples of this would be broiler chickens selected to grow a lot in a short time, developing leg and heart conditions as a result. The cattle breed Belgian Blue has about double the muscle mass of most cattle breeds as it was selected for muscle hypertrophy. The animals of that breed are not able to give birth naturally anymore and have to undergo caesarean section for every birth, and this is only one of many welfare implications of the muscle hypertrophy. Dairy cows also suffered from the selection for increased milk production, their enlarged udders putting physical strain on their bodies, and increasing risks of mastitis. This is where applied ethology research comes in to try and find a balance between the productivity aims and the standards of animal welfare that consumers and governments expect and request.

Dog breeding welfare can also be questioned from a critical point of view: it started from a purpose of getting more specialized dogs, able to tackle various tasks to help humans in everyday life i.e. hunting, herding, guarding. Historical dog breeding for specialized functions has probably had no major welfare impact on the dogs, the selection for a function implying the necessity of obtaining healthy dogs. But this has shifted since the Victorian era and the appearance of dog shows, beauty competitions and purebred standards. It is now widely known that most purebred "show" lines suffer major health issues related to inbreeding and selection for purely subjectively aesthetic characteristics. Many airplane companies refuse nowadays to transport molossoid dog breeds since their shortened snout puts them at higher risk of cardiovascular and respiratory accidents. Dog insurance companies have different price quotes according to the breeds, purebred ones always higher than for mutts. German Shepherds, one focus of our study, are commonly known to be at higher risk of hip dysplasia. It should be noted though, that this seems to be only valid for pet and show lines, whereas working and army German Shepherds do not have this predisposition. Working dogs are often physically much closer to their ancestors from before the shift in selection, when show and pet lines of all breeds are often physically very different from their ancestors. Studies underlining the impact of recent history of the breeds opposed to the historical background also give us hope, showing that with good decisions and regulations in dog breeding, these issues could be resolved, providing that enough awareness is given to the problem. Indeed, promoting health issues and painful conditions in dogs only for the sake of keeping a fleeting aesthetic standard and genetic purity is questionable from an ethical point of view. Not only does it concern dog welfare, but the owner's welfare too, with the anguish of seeing one's pet suffer. Another issue related to human welfare is the occurrence of behavioural disorders with inbreeding that can lead to aggressiveness in dogs.

Nevertheless, this particular study has no ethical conflict, considering the dogs were recruited from private owners that agreed to participate in the research project. The dogs were not restrained or physically harmed in any way, at worst they could get some frustration from the tests, but they were rewarded and pet afterwards. The ethical permit for this study was approved by the committee for ethical approval of animal experimentation in Linköping, approval no 51-13.

5.2 Conclusions

In conclusion, I did reveal breed differences in both the problem-solving test and in the pointing test. These differences could be due to intrinsic lower trustfulness of German Shepherd dogs in strangers. The latency of choice, especially, was correlated to the intensity but also to the body posture and the physical contacts with the experimenter. Overall, the latency of choice could be an indicator of the level of trust of the dog (trust in its choice, in the experimenter, and self-confidence). The dogs were also less accurate in the pointing test when having many physical contacts with their owner. This might be linked to training or personality allowing less focus on other humans than the owner. I also conclude that this study and the parallel one on Labrador types give strong indication of the impact of recent breeding for different purpose, and show that types of a same breed can differ as much or more than two different breeds.

6 Acknowledgement

I would like to express my gratitude first to Pr Per Jensen and Dr Lina Roth, my supervisors, for their help and availability, and for giving me the opportunity to do my thesis under their supervision.

I also want to thank PhD student Mia Persson for the extra help with the project, with the testing and the statistics.

I am grateful Ann-Charlotte Svensson for setting me up in the lab for the hair preparation for cortisol analysis and to Dr Åshild Olsen Faresjö for letting me follow and help with the hair cortisol extraction and analysis, and for her work performing the extraction and analysis.

My thanks to the AVIAN team in general for welcoming me as a member of the team.

Thank you to Lise-Lotte Halldén for helping set up for the testing, and for the moral support and help throughout the thesis. It was really nice being able to discuss our projects together and have someone to relate to.

I want to thank all dog owners who gave their time to participate in the study, and all dogs for their enthusiasm.

And finally thanks to my classmates whose presence and jokes kept me sane, and my family for all the support.

7 References

Bennett, A., Hayssen, V., 2010. Measuring cortisol in hair and saliva from dogs: coat color and pigment differences. Domestic Animal Endocrinology 39, 171-180.

Boulton, K., Couto, E., Grimmer, A.J., Earley, R.L., Canario, A.V.M., Wilson, A.J., Walling, C.A., 2015. How integrated are behavioral and endocrine stress response traits? A repeated measures approach to testing the stress-coping style model. Ecology and Evolution 5, 618-633.

Bryan, H.M., Adams, A.G., Invik, R.M., Wynne-Edwards, K.E., Smits, J.E.G., 2013. Hair as a Meaningful Measure of Baseline Cortisol Levels over Time in Dogs. Journal of the American Association for Laboratory Animal Science 52, 189-196.

Carere, C., Caramaschi, D., Fawcett, T.W., 2010. Covariation between personalities and individual differences in coping with stress: Converging evidence and hypotheses. Current Zoology 56, 728-740.

Dorey, N.R., Udell, M.A.R., Wynne, C.D.L., 2009. Breed differences in dogs sensitivity to human points: A meta-analysis. Behavioural Processes 81, 409-415.

Duffy, D.L., Hsu, Y.Y., Serpell, J.A., 2008. Breed differences in canine aggression. Applied Animal Behaviour Science 114, 441-460.

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E.J., Silva, P.M., Galaverni, M., Fan, Z.X., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vila, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E.A., Marques-Bonet, T., Wayne, R.K., Novembre, J., 2014. Genome Sequencing Highlights the Dynamic Early History of Dogs. Plos Genetics 10, 12.

Gacsi, M., Gyori, B., Viranyi, Z., Kubinyi, E., Range, F., Belenyi, B., Miklosi, A., 2009a. Explaining Dog Wolf Differences in Utilizing Human Pointing Gestures: Selection for Synergistic Shifts in the Development of Some Social Skills. Plos One 4, 6.

Gacsi, M., McGreevy, P., Kara, E., Adam, M., 2009b. Effects of selection for cooperation and attention in dogs. Behavioral and Brain Functions 5, 8.

Gaunet, F., 2010. How do guide dogs and pet dogs (Canis familiaris) ask their owners for their toy and for playing? Animal Cognition 13, 311-323.

Hare, B., Plyusnina, I., Ignacio, N., Schepina, O., Stepika, A., Wrangham, R., Trut, L., 2005. Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. Current Biology 15, 226-230.

Hare, B., Rosati, A., Kaminski, J., Brauer, J., Call, J., Tomasello, M., 2010. The domestication hypothesis for dogs' skills with human communication: a response to Udell et al. (2008) and Wynne et al. (2008). Animal Behaviour 79, E1-E6.

Hare, B., Tomasello, M., 2004. Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. Animal Behaviour 68, 571-581.

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

Horowitz, A., 2011. Theory of mind in dogs? Examining method and concept. Learning & Behavior 39, 314-317.

Hsu, Y.Y., Serpell, J.A., 2003. Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. Journal of the American Veterinary Medical Association 223, 1293-+.

Isaksson, J., Nilsson, K.W., Nyberg, F., Hogmark, A., Lindblad, F., 2012. Cortisol levels in children with Attention-Deficit/Hyperactivity Disorder. Journal of Psychiatric Research 46, 1398-1405.

Kaminski, J., Neumann, M., Braeuer, J., Call, J., Tomasello, M., 2011. Dogs, Canis familiaris, communicate with humans to request but not to inform. Animal Behaviour 82, 651-658.

Karlén, J., Ludvigsson, J., Frostell, A., Theodorsson, E., Faresjö, T., 2011. Cortisol in hair measured in young adults - a biomarker of major life stressors? BMC Clinical Pathology C7 - 12 11, 1-6.

Kerepesi, A., Doka, A., Miklosi, A., 2015. Dogs and their human companions: The effect of familiarity on dog-human interactions. Behavioural Processes 110, 27-36.

Kirchhofer, K.C., Zimmermann, F., Kaminski, J., Tomasello, M., 2012. Dogs (Canis familiaris), but Not Chimpanzees (Pan troglodytes), Understand Imperative Pointing. Plos One 7, 7.

Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H., De Jong, I.C., Ruis, M.A.W., Blokhuis, H.J., 1999. Coping styles in animals: current status in behavior and stress-physiology. Neuroscience and Biobehavioral Reviews 23, 925-935.

Lakatos, G., Soproni, K., Doka, A., Miklosi, A., 2009. A comparative approach to dogs' (Canis familiaris) and human infants' comprehension of various forms of pointing gestures. Animal Cognition 12, 621-631.

McKinley, J., Sambrook, T.D., 2000. Use of human-given cues by domestic dogs (Canis familiaris) and horses (Equus caballus). Animal Cognition 3, 13-22.

Miklosi, A., Kubinyi, E., Topal, J., Gacsi, M., Viranyi, Z., Csanyi, V., 2003. A simple reason for a big difference: Wolves do not look back at humans, but dogs do. Current Biology 13, 763-766.

Miklosi, A., Topal, J., Csanyi, V., 2004. Comparative social cognition: what can dogs teach us? Animal Behaviour 67, 995-1004.

Mongillo, P., Bono, G., Regolin, L., Marinelli, L., 2010. Selective attention to humans in companion dogs, Canis familiaris. Animal Behaviour 80, 1057-1063.

Morelius, E., Nelson, N., Theodorsson, E., 2004. Salivary cortisol and administration of concentrated oral glucose in newborn infants: improved detection limit and smaller sample volumes without glucose interference. Scandinavian Journal of Clinical & Laboratory Investigation 64, 113-118.

Ostojic, L., Clayton, N.S., 2014. Behavioural coordination of dogs in a cooperative problem-solving task with a conspecific and a human partner. Animal Cognition 17, 445-459.

Parker, H.G., Ostrander, E.A., 2005. Canine genomics and genetics: Running with the pack. Plos Genetics 1, 507-513.

Persson, M.E., S.V., R.L., Martin, J., Dominic, W., Per, J., 2015. Human-directed social behaviour in dogs shows significant heritability. Genes, Brain and Behavior In press

Pettersson, H., Kaminski, J., Herrmann, E., Tomasello, M., 2011. Understanding of human communicative motives in domestic dogs. Applied Animal Behaviour Science 133, 235-245.

Pollard, T.M., 1995. USE OF CORTISOL AS A STRESS MARKER - PRACTICAL AND THEORETICAL PROBLEMS. American Journal of Human Biology 7, 265-274.

Reid, P.J., 2009. Adapting to the human world: Dogs' responsiveness to our social cues. Behavioural Processes 80, 325-333.

Russell, E., Koren, G., Rieder, M., Van Uum, S., 2012. Hair cortisol as a biological marker of chronic stress: Current status, future directions and unanswered questions. Psychoneuroendocrinology 37, 589-601.

Scheider, L., Kaminski, J., Call, J., Tomasello, M., 2013. Do domestic dogs interpret pointing as a command? Animal Cognition 16, 361-372.

Shimabukuro, C., Putrino, N., Helbling, J., Tognetti, S., Bentosela, M., 2015. Individual differences in social and non-social behaviors in domestic dogs (Canis familiaris) during the acquisition, extinction and reacquisition of a problem solving task. Behavioural processes 113, 179-186.

Smith, B.P., Litchfield, C.A., 2013. Looking back at 'looking back': operationalising referential gaze for dingoes in an unsolvable task. Animal Cognition 16, 961-971.

Soproni, K., Miklosi, A., Topal, J., Csanyi, V., 2002. Dogs' (Canis familiaris) responsiveness to human pointing gestures. Journal of Comparative Psychology 116, 27-34.

Svartberg, K., 2006. Breed-typical behaviour in dogs - Historical remnants or recent constructs? Applied Animal Behaviour Science 96, 293-313.

Udell, M.A.R., Wynne, C.D.L., 2010. Ontogeny and phylogeny: both are essential to human-sensitive behaviour in the genus Canis. Animal Behaviour 79, E9-E14.

Vila, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J., Wayne, R.K., 1997. Multiple and ancient origins of the domestic dog. Science 276, 1687-1689.

Viranyi, Z., Gacsi, M., Kubinyi, E., Topal, J., Belenyi, B., Ujfalussy, D., Miklosi, A., 2008. Comprehension of human pointing gestures in young human-reared wolves (Canis lupus) and dogs (Canis familiaris). Animal Cognition 11, 373-387.

Wilsson, E., Sundgren, P.E., 1997. The use of a behaviour test for the selection of dogs for service and breeding .1. Method of testing and evaluating test results in the adult dog, demands on different kinds of service dogs, sex and breed differences. Applied Animal Behaviour Science 53, 279-295.

Wobber, V., Hare, B., Koler-Matznick, J., Wrangham, R., Tomasello, M., 2009. Breed differences in domestic dogs' (Canis familiaris) comprehension of human communicative signals. Interaction Studies 10, 206-224.

8 Appendix

8.1 Pointing orders

Two random pointing orders were generated to alternate between sides without having more than two points consecutively in the same side.

Right-Left-Right-Right-Left-Right-Left-Left-Right-Left

Left-Right-Left-Left-Right-Left-Right-Right-Left-Right

8.2 Body posture scoring

A subjective score from 1-5 explaining the body posture of the dog based upon overall body posture and behaviour shown towards the experimenter and the test setup.

Score	***Description***
1	*High wagging tail and high body position. Appears confident.*
2	*Mostly high wagging tail and/or high body position. Appears fairly confident.*
3	*Mostly neutral tail and/or body position. Appears as sometimes unsecure.*
4	*Mostly low tail and/or body position. Can wag their tail. Appears unsecure and cautious.*
5	*Low tail and/or body position. Appears unsecure and very cautious.*

8.3 Feeding score scale

A score from 1-3 explaining the time it took for the dog to eat the first treat presented on the single plate.

Score	***Description***
1	*Late feeding. Needs encouragement; more than 40 s until the dog ate the first treat. All treats were consumed within 2 minutes.*
2	*Medium feeding. Some needed encouragement, between 20-40 seconds before the dog ate the first treat. All treats were consumed within 20-40 seconds.*
3	*Early feeding, No encouragement is needed; the dog ate all treats within 20 seconds of when the plate was put down.*

8.4 Intensity scoring

A score from 1-5 explaining the level of energy, activity, intensity of the dog based on pace, test interaction, jumps, passive behaviour.

Score	***Description***
1	*Very intense. Moving all the time, energetic and with high pace, high frequency of paw use, licking or nose use when interacting with test setup, jumps.*
2	*Intense. Mostly moving in a high pace, intermediate frequency of paw use, licking or nose use when interacting with the test setup.*
3	*Medium intensity. Moving around most of the time in medium/low pace. Medium frequency of paw use, licking of nose use when interacting with test setup*
4	*Low intensity. Low pace when moving around, might be passive (standing or sitting) for some time.*
5	*Very low intensity. Does not move around most of the time, passive (standing, sitting of lying down) for considerable time during the test.*

8.5 Cortisol analysis protocol

8.5.1 Preparation of the hair samples for analysis

2 mL tubes containing a steel bead each were weighted with a precision scale. Hair samples were sorted to keep only guard hair for the analysis. It was then cut from the root end in small sections of about 1 mm on a total length from the original cut end of about 3 cm. About 5 to 7 mg of these hair sections were transferred in the tubes.

8.5.2 Following protocol from Karlén et al. (2011)

Batches of 5 tubes were put in each solid aluminum holder specially made for the purpose. The holders including the samples and steel balls in the sample tubes were frozen in liquid nitrogen for 2 minutes and rapidly minced in Retch TissueLyser II at 30 Hz for 20 seconds, producing very fine hair powder. One mL pure methanol (Chromasolv, Sigma-Aldrich) was added to each tube and each tube fixed in metal holder on an oblique (5 degrees from the horizontal) plate on a horizontal shaker (Edmund Bühler, type B1) at room temperature, making sure that the steel balls continuously moved back and forth in the sample tubes for at least 10 hours. The tubes were then centrifuged and 700 µL of the supernatant moved to another sample tube for lyophilization in SpeedVac Plus SC210A (Savant) using Edwards XDS 5 vacuum pump.

Measuring cortisol in the extracted hair samples

The samples were dissolved in 150 µL 0.1 mol L^{-1} phosphate buffer, pH 7.4 containing 0.02% bovine serum albumin (BSA) and 0.01% triton X-100, and concentrations of cortisol measured as described earlier by Morelius et al. (2004). The lowest detectable concentration was 1 nmol L^{-1} and the within- and between assay coefficients of variation were 6.1% and 9.3%, respectively, at 10 nmol L^{-1}.

8.5.3 Following protocol from Morelius et al. (2004)

Antiserum was diluted in the ratio 1:40 (instead of 1:20), the radioligand was diluted until 100 mL representing 3000 counts per minute (CPM). The antiserum volume was 100 mL, radioligand volume was 100 mL (~3000 CPM), calibrator and sample volumes were each 10 mL. In order to lower the limit of detection further, the addition of 114 E. Morelius et al. radioligand was delayed and the incubation time prolonged. The samples were initially incubated with the antiserum at +4ºC for 48 h before the radioligand was added and then additionally 24 h before the bound and free fractions were separated using solid-phase, bound anti-rabbit antibodies. The assays were analysed in a gamma counter 1277 from Wallac (Turku, Finland).