Author:
Bistre Dabbah, Sharyn

Title:
Investigating the effect of individual attributes on dogs' performance in medical detection tasks

Investigating the effect of individual attributes on dogs' performance in medical detection tasks

Sharyn Bistre Dabbah

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of PhD in veterinary sciences in the Faculty of Health Sciences.

September 2022

Word count: 76,196

ABSTRACT

Medical detection dogs (MDD) are utilised to aid human health; Bio-detection dogs are trained to identify ex-situ conditions' odours, and Alert Assistance dogs to alert crises in people with chronic illnesses (Rooney et al., 2013). Individual variation is associated with detection dog performance (Lazarowsky et al., 2020). However, it is unexplored in MDD. Identifying how MDD characteristics relate to task success will aid the selection of optimal dogs.

This project aims to identify traits relevant to MDD and explore how these vary across tasks and training stages. It seeks to develop a test battery to investigate these traits in an MDD sample.

A practitioner survey revealed traits important for MDD and those that differ most from ideal. Some characteristics varied significantly across tasks. In total, 27 relevant attributes were derived.

Trainers from the charity Medical Detection Dogs® rated these 27 traits and overall ability of 58 MDD at multiple time-points. Ratings showed low consistency over training for most traits. Some characteristics were significantly associated with derived success measures.

Success measures were explored and derived for subsequent studies: training outcome, composite total ability score, scent sensitivity, and specificity.

The dogs were tested with a test battery designed to quantify the most relevant MDD traits. The variables measured were clustered into 11 components. Some were associated with dogs' demography, some with impulsivity scores from the Dog impulsivity assessment scale, and some with success measures.

A cognitive bias test assessed the dogs' tendency to make decisions over ambiguity. The dogs' latency to approach ambiguous locations was significantly linked with some test battery components and success measures.

Results indicate that certain traits in MDD are associated with different performance levels and vary across tasks and training stages. The test battery may be useful to predict MDD ability. Further exploration is necessary.

Dedicated to my parents with love. Thanks for everything and all your love and support always.

To my mentor, Moisés Heiblum Frid. Wherever you are. For your teaching and friendship and your commitment to improving dog lives in Mexico.

Acknowledgements

I wish to thank my excellent supervisors, Dr. Nicola Rooney and Prof. Mike Mendl, for their invaluable assistance and dedication during my PhD. I would not have been able to do it without you. I also dedicate a special thanks to Toby Knowles for his statistical advice.

Many thanks to the University of Bristol for all their support in making my project possible. With their workshops and all their resources to enrich my professional progress. I especially appreciate their aid during the pandemic times, when advancing my research was challenging. I am grateful to my Langford's Animal Welfare and Behaviour group colleagues for the learning opportunities I enjoyed, their advice and their friendship.

A special thanks to the Charity Medical Detection Dogs®, Claire Guest, Chris Allen, Rob Harris and all the fantastic staff for their contribution and support in my project, for facilitating and participating in my data collection and for all the outstanding work they do to help people. Also, thanks to their lovely dogs for their enthusiastic participation in my test.

This project is possible thanks to my scholarship from CONACyT (Grant 472257) and the financial support from the University of Bristol.

I am forever thankful to my alma mater, the Universidad Nacional Autonoma de México, for forging my identity in animal care and public service responsibility. Particularly thanks to my colleagues from the DEFSAL Department of ethology, wild animals and laboratory animals for their teaching and friendship and for helping me apply for my PhD: Prof. Anne María del Pilar Sisto, Prof. Alberto Tejeda, Prof. Carlos González-Rebeles, Prof. Eduardo Santurtún, Prof. Francisco Galindo, Prof. Marcela González de la Vara and Prof. Moises Heiblum.

Thanks so much to all my friends and colleagues who helped me with proofreading: Ann Goldstein, Ian Staniland, Izzie Philpotts, Sarah Kappel and Zoe Parr-Cortes. I am incredibly grateful to Sheila Brill for her dedication when correcting my writing, Helena Hale for her advice, Alejandra Escalera for her help with editorial design, and Guillermina Hernandez and Marco Antonio Ramírez for their technical support. A special thanks to Jacobo Herrera for his assistance in my project.

I dedicate my thesis to my parents, Sylvia and Jacobo, for all their love and support every step of the way. For always being there in all my pursuits and encouraging me to go forward.

I also dedicate this project to my mentor and dear friend Moises Heiblum Frid, my animal behaviour hero. He inspired and encouraged me to commit my life to help and understanding animals through science. I hope he knows I am here, thanks to him wherever he is.

I want to thank all my family:

My siblings, Celia, Sonia, Elias and Karen, for always supporting me, bullying me, missing me and spending long hours on the phone from the other side.

My nieces and nephews, Marquitos, Jacobo P., Dan, Emily, Jacobo H., Aline, Jack, Adina and baby Leo, for always waiting for my return, doing incredible things together and making me laugh. I hope you all get as far as possible to make your dreams come true.

My grandparents, Isaac and Samia, for all their values and stories and for giving us a chance to have a home in Mexico. Especially my Abuelita for inspiring me to study further.

My aunts “Las manas” Gloria, Rosa, Linda and Toña and my uncle Toño for all the support, care, and encouragement, their spoiling, and great food.

My family from Tabasco, especially my Tía Blanca and my cousins Zaira and Niza, was great growing up together, and I love you and miss you a lot.

My dogs have been family and my wisest teachers throughout my life. They inspired me to pursue this path. I particularly thank my dear boy Enzo, who helped me practice some of my tests and is waiting for me to come home.

I am forever grateful to all my friends:

Ann Goldstein, and Debbie and Ian Staniland, my family in the UK and the most incredible support, and for giving me a sense of home away from home.

My amazing friends and roommates, Marco, Ale and Karina, always there for me. We shared the most fantastic time. I could not imagine a better company for the lockdown. Also, to Guille for very good coffee afternoons.

My dear friend and colleague Sarah Kappel. Our writing retreats have been a creative shelter, helping me to complete this dissertation. I wish you the best of luck in yours.

To my beloved friends for life, especially: Berni, Cherno, Chikis, Dalit, Danny, Eliane, Ilan, Ingrid, Karla, Mica, Mimi, Peretz, Natalia, Raquel, Renato,Tofy and Yafa. “Las piedras rodando se encuentran….”

Finally, I am incredibly grateful to my country, which forged my identity and gave my family a home. *“Mexico Lindo y Querido… si muero lejos de ti”*.

COVID STATEMENT

The COVID-19 pandemic disrupted my research. From March 2020, The Medical Detection Dogs® Charity, where I performed my data collection, closed due to the lockdown. Therefore, I could not complete dog testing, missing five already scheduled dogs, and prevented potentially testing new dogs.

Before COVID-19 restrictions, trainers rated their dogs in-situ at four training time-points, using paper rating forms to generate subjective data. It was feasible to track them and collect the forms as they were centralised in the charity. However, with the lockdown, I continued with data collection remotely, which took much longer as I must generate electronic rating forms. Many trainers were on furlough, and others were working from home. Therefore, tracking their responses became very difficult.

A few trainers continued working on-site with some dogs. The training was interrupted for many, and most charity's efforts were concentrated on training the dogs to detect COVID-19.

For these, among other complications, amid having to adapt to work from home, I had significant delays in my PhD progress and must modify the project plan. Some of these included:

a. The dog sample was smaller than what was initially planned.
b. As the training of many dogs was interrupted, I could not determine if they passed the training programme. Hence, I considered dogs to have succeeded in training if they remained in the system at the end of data collection and did not fail, although I did not know their ultimate fate.
c. The ratings for each dog were incomplete as the trainers could not fill the rating sheets for each assessment time-point. Hence only multiple assessments were considered, and the data was comprised in fewer training time-points to increase the number of assessments for each. This resulted in a small dog sample for the respective analysis.
d. Due to the time limitations and the complications already mentioned, I must restructure my thesis plan. Initially, in addition to testing medical detection dogs with a test battery, I was going to try naive dogs. This study was not conducted. Also, two studies in Chapter 3 (Study A and B) were to be separate chapters. I comprised them in one as these were related. I planned to analyse video recordings of dogs evaluated with an internal test of the charity to compare their observed performance with assessors' ratings. However, time was insufficient.
e. Due to time and logistics limitations, I could not perform reliability assessments of the Test battery.
f. An advantage was that I could include scent sensitivity and specificity data from the dogs in the COVID-19 detection project, which was quite enriching.

Details on changes due to the pandemic are described throughout the dissertation.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others is indicated as such. Any views expressed in the dissertation are those of the author.

Sharyn Bistre Dabbah

SIGNED: DATE: 09 / 09 / 2022

[redacted]

Table of Contents

Table Index

Figure Index

ABBREVIATIONS TABLE

Abbreviation	*Meaning*
MDD	Medical detection dog
DAD	Diabetes alert dog
S+	True positives
S-	True negatives
VOCs	Volatile organic compounds
DIAS	Dog impulsivity assessment scale
PANAS	Positive and negative activation scale
C-BARQ	Canine Behavioral Assessment and Research Questionnaire
ICC	Intraclass correlation coefficient
PCA	Principal Component Analysis
DR	Data reduction
COR	Correlated with another variable ≥ 0.70
KMO	Kaiser-Meyer-Olkin Measure
OA	Overall ability
Disc	Discrepancy score
ODS	Overall discrepancy score
CTAS	Composite total ability score
U	Mann Whitney U Test
R^2	Coefficient of determination
ΔR^2	Adjusted R2
B	Unstandardised regression coefficient
F	Compared to an F-distribution
SE	Standard error of the coefficient
$\chi 2$	Chi square
p	p value
CI	Confidence interval
LL	Lower limit
UL	Upper limit
β	Standardised coefficient
Wald	Wald test
df	Degrees of freedom
CBT	Cognitive bias test
CB	Cognitive bias
P	Positive location
NP	Near positive location
M	Midle location
NN	Near negative location
N	Negative location

Chapter 1. General introduction

1.1 Overview

The use of medical detection dogs (MDD) is increasing due to their ability to detect biological organic compounds. Bio-detection dogs are trained to discriminate remote odours such as cancer, malaria and now COVID-19, and medical assistance dogs to alert upcoming crises to a human partner with chronic conditions such as diabetes or epilepsy. However, how performance varies between individuals, which factors are associated with this variation and how these differ across disciplines remains unexplored. Further knowledge is needed to optimise dog selection, reduce time wastage and economic losses, and train the dogs best suited to the task (Rooney et al., 2004; Cobb et al., 2015).

This chapter seeks to introduce the main aims of my PhD project and the different studies involved. In addition, it navigates the theoretical background behind this research, based on an extensive literature review on the main findings of MDD research, individual differences in detection dogs and other working dog roles and their relevance for MDD.

1.2 Background empirical and theoretical framework

The domestic dog (*Canis lupus familiaris*) can be trained as an aid for humans, particularly in scenarios that benefit from their highly sensitive sense of smell (Moulton et al., 1960; Browne, 2006; Johnen et al., 2017). Detection dogs are trained to locate an extensive range of items such as narcotics, explosives, weapons, human victims and illegal products (Furton and Myers, 2001; Gazit et al., 2003; Oesterhelweg et al., 2008; Adamkiewicz et al., 2013; Riezzo et al., 2014; Diverio et al., 2016), and wildlife species (Oldenburg et al., 2016), oestrus detection in cows (Fischer-Tenhagen et al., 2011) and insects (Masini et al., 2017).

The use of dogs' olfaction has recently been extended to the medical field (Edwards et al., 2017). Research has investigated the effectiveness of dogs for the diagnosis of conditions such as cancer (e.g. Cornu et al., 2011; Walczak et al., 2012; Seo et al., 2018) and other diseases, as well as functioning as an alert system for health events like hypoglycaemic states in patients with type 1 diabetes (e.g. Chen et al., 2000; Rooney et al., 2013; Petry et al., 2015). Although they show variability in their performance (Gonder-Frederick et al., 2014; Jezierski et al., 2015; Rooney et al., 2019), they have clear potential to provide a non-invasive diagnosis option for some diseases (Walczak et al., 2012) and to improve the quality of life of patients with chronic conditions (Rooney et al., 2013).

1.2.1 Medical detection dogs

1.2.1.1 What does an MDD do?

Scientific studies of guide dogs' behaviour date from 1934 (Humphrey and Warner). The diversification of dogs used for medical support, including diagnosis of different diseases (Bijland et al., 2013) and assistance for patients with physical or cognitive conditions (Audrestch et al., 2015), occurred relatively recently.

MDD can detect volatile organic compounds (VOCs) from different diseases (Willis et al., 2004; Bijland et al., 2013). Some animals develop this ability spontaneously (e.g. Williams and Pembroke, 1989), while for others, it is reinforced through training (e.g. Willis et al., 2004; Walczak et al., 2012).

There are different roles within MDD: Bio-detection dogs work remotely in controlled conditions, discriminating disease scent samples from healthy controls (Jezierski et al., 2015), while Medical alert Assistance dogs, like glycaemic alert dogs, work with the patient and alert them before a crisis (Rooney et al., 2013).

1.2.1.2 Measuring MDD scent sensitivity and specificity

A key concept in understanding MDD performance is that of scent detection sensitivity and specificity. The scent detection ability of MDD (and other kinds of detection dogs) is measured objectively from scent sensitivity and specificity levels during *in vitro* trials or real-life searches. Signal detection theory provides a framework to determine the accuracy of a discrimination tool (McNicol, 2005). Regarding detection dogs, the target scent is considered a signal (S+ condition), while other stimuli (non-target scents) are considered noise (S- condition). Dogs are trained to alert to the presence of S+ and to ignore the presence of S- (See Mahoney, 2012). When the dog correctly indicates an S+, it is referred to as a 'Hit', meaning a true positive, but if they alert to S- incorrectly, it is considered a 'False alarm' or false positive. If the dog correctly ignores an S- it is referred to as a true negative, but if a dog misses an S+, it is a 'Miss' or a false negative (Lalkhen and McCluskey, 2008; Mahoney, 2012; Walczak et al., 2012).

Scent sensitivity refers to the number of true positives divided by the total number of scent presentations (Lalkhen and McCluskey, 2008). It indicates how effective the dog is in alerting a target scent correctly.

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

Scent specificity refers to the number of true negatives divided by the total number of scent presentations. It shows the dog's ability to ignore the noise, meaning blank samples or control scents.

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

Both high sensitivity and specificity are necessary for optimal detection performance, but each balance will depend on the task's requirements (Gadbois and Reeve, 2016; Lazarowski et al., 2020a). Dogs are referred to by Gadbois and Reeve (2016) as 'Liberal' or 'Conservative' regarding their error bias on scent discrimination. Liberal dogs had a higher tendency for false alarms and minimising false negatives to increase the chance of a hit, while more conservative dogs are inclined to minimise false alarms but may also fail to indicate S+ (Gadbois and Reeve, 2016). These tendencies are subject to perception differences of the information, associated with variations in physical characteristics, sensory processing (i.e. olfaction threshold), personality attributes, discrimination learning, and sensitivity to decision consequences, i.e. reward vs non-reward or punishment (Hiby et al., 2004; Walker et al., 2006; Lynn and Barrett, 2014; Lazarowski et al., 2020a), and external factors such as reinforcement schedule and the information presented, e.g. the arrangement and similarities and attenuation of S+ and S- on a scent trial, misleading cues from the participants on search trial and environmental elements (Nevin, 1969; Mahoney, 2012; Lynn and Barrett, 2014; Concha et al., 2019; Lazarowski et al., 2020a).

In MDD, high sensitivity is critical since failing to detect S+ may jeopardise human lives (Reeve and Koivusalo, 2018). However, high specificity is also desirable since an increased tendency of false alarms reduces the feasibility of MDD and leads to unnecessary spending of resources. These relative values and risks, however, vary with the task. For instance, a 'liberal' bias may be preferred in dogs that alert hypoglycaemic crisis since failing to do so may risk the client's health. However, a conservative approach may be desirable in cancer detection dogs for higher diagnostic accuracy. These parameters may be compared against a gold standard diagnostic test for a given disease to assess MDD performance effectiveness (Edwards et al., 2017).

1.2.1.3 MDD scent detection training

The charity Medical Detection Dogs® (Medical Detection Dogs, 2020), located in Milton Keynes, UK, where this research was mainly conducted, trains Bio-detection dogs to discriminate odours for disease diagnosis and Assistance dogs to aid people with chronic conditions. During bio-detection training, scent samples are presented to the dogs in a line or circle arrangement (Jezierski et al., 2015), interspersing S+ among S- samples. Dogs are reinforced when indicating S+ (e.g. Concha et al., 2019; Davies et al., 2019). Training difficulty gradually increases by reducing the concentration of S+ (e.g. Concha et al., 2019), increasing the number of S- or adding S+ into mixtures with other chemical components (e.g. Davies et al., 2019). Testing frequently involves double-blind testing where the trial's composition is unknown to all participants (Lazarowski et al., 2020a), from which sensitivity and specificity data are quantified. For some tasks, subsequent training requires the dog to indicate a scent in people or an outdoor environment, with the goal of dogs eventually identifying the scent in hospitalised patients or at a port of entry (e.g. Bomers et al., 2012; Sakr et al., 2022).

Medical alert Assistance dogs' training starts with *in vitro* trials with S+ (from the prospective client), and subsequently, the dog is conditioned to identify S+ hidden on the client. The dog is trained to alert the scent with particular behaviours, in some cases retrieving a blood testing kit, although they may perform additional behaviours to demand the client's attention (Wilson et al., 2020).

1.2.1.4 Bio-detection dogs

Bio-detection dogs discriminate between (presented) scent samples containing S+ of VOCs from a health condition and those that are S-. They have been mainly utilised for the detection of human cancer (Jezierski et al., 2015). However, there are also ongoing experimental projects into other conditions (e.g. Davies et al., 2019; Guest et al., 2019; Sakr et al., 2022).

More than 30 years ago, Williams and Pembroke (1989) reported the first case of a dog detecting melanoma in its owner. Since then, research has focused on dogs' effectiveness at discriminating different kinds of cancer, including prostate (Cornu et al., 2011; Elliker et al., 2014), skin (Pickel et al., 2004), breast (McCulloch et al., 2006; Walczak et al., 2012), colorectal (Sonoda et al., 2011), ovarian (Horvath et al., 2008), lung (McCulloch et al., 2006) and bladder (Willis et al., 2004). Currently, training dogs for cancer diagnosis is primarily experimental and still not frequently applied in health services (Moser and McCulloch, 2010; Jezierski et al., 2015). However, it is a non-invasive method with relatively low costs, is easy to interpret, and may aid in diagnosing these conditions. (Jezierski et al., 2015).

Many studies revealed high levels of sensitivity and specificity, surpassing those of an electronic nose, a device that consists of gas sensors for the recognition of a range of VOCs (Scott et al., 2006; Bijland et al., 2013). However, these levels vary across conditions. Willis et al. (2004) showed dogs had a mean sensitivity of 41% in a bladder cancer detection trial, significantly higher than chance (14%). Pickel et al. (2004) revealed that dogs had a success rate of 75–85.7% when detecting melanoma biomarkers in hospitalised patients. Cornu et al. (2011) measured excellent sensitivity levels (91%) for prostate cancer identification. However, these studies included small numbers of dogs (from one to six). Small samples make it challenging to validate this evidence for practical purposes since it diminishes statistical effect, producing unreliable findings and affecting study replicability (Button et al., 2013).

Bio-detection dog research has extended to other conditions. A proof of principle study reported high levels of mean sensitivity (100%) and specificity (94%) in a dog detecting *Clostridium difficile* in human stool samples and thirty hospitalised patients (sensitivity= 93%, specificity= 97%) (Bomers et al., 2012). Even though only one dog was trained, the results are promising as it successfully transitioned from training scent trials to active employment in a health centre.

Five dogs detected bacteriuria remotely from different pathogens having a mean sensitivity of almost 100% and specificity of > 90% (Maurer et al., 2016). Two dogs successfully discriminated malaria samples (sensitivity= 73·3%, specificity= 91·0%) from socks worn by children from a high prevalence community in The Gambia (Guest et al., 2019). In the future, using malaria detection dogs may be a non-invasive screening method at ports of entry. Four dogs showed high sensitivity (94.2%) and specificity (98.5%) when detecting *Pseudomonas aeruginosa* from controls that the dogs were previously exposed to, although it was slightly lower when presented in a mixed culture with other microorganisms (sensitivity= 86.5%, specificity= 84.1%). Early detection of this bacteria may aid in preventing or managing cystic fibrosis (Davies et al., 2019).

The novel coronavirus disease 2019 (COVID-19) caused by the Severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) brought unprecedented challenges. Dogs were trained experimentally to discriminate COVID-19 VOCs in several countries to aid the urgent necessity of mass testing to help reduce its spread (e.g. Jones et al., 2020; Grandjean et al., 2021; Mendel et al., 2021).

These dogs are a potential non-invasive and fast screening tool in public areas and ports of entry (Sakr et al., 2022). The studies trained between four and 20 dogs and achieved high sensitivity (>80%) and specificity (>90%). Still, there are challenges with training enough dogs to scale up this screening for use in real-life scenarios (Sakr et al., 2022). However, in some countries, including UAE (Sakr et al., 2022), USA (Mendel et al., 2021) and Finland (Hielm-Björkman and Niuro, 2021), dogs trained to detect COVID-19 were trialled in airports and showed high success rates (>80% sensitivity and specificity). These successful trials evidence that using Bio-detection dogs in real-life situations is feasible and encourages future research.

1.2.1.5 Medical alert Assistance dogs

Medical alert assistance dogs (referred to as Assistance dogs throughout the dissertation) are recognised as service dogs by Assistance dogs UK (Assistance Dogs UK, 2022a) and trained by certified organisations like Medical Detection Dogs® (Medical Detection Dogs, 2020) and Support Dogs® (Support Dogs UK, 2022b).

Assistance dogs are paired with patients with chronic conditions like Diabetes mellitus Type 1, Postural Tachycardia Syndrome (PoTS), Addison's disease, intense allergies, epileptic seizures and various endocrine and neurological disorders.

Diabetes alert dogs (DAD) perform an assistance role similar to guide dogs, mobility or hearing support dogs. However, they differ since their tasks involve using olfaction to detect changes in VOCs from perspiration or breath, which predict drops below the normal range or hypoglycaemic episodes. Potentially, dogs may also react to behaviour changes or tremors in their owner (Tauveron et al., 2006).

Hypoglycaemic episodes are a common and potentially dangerous consequence of Type 1 diabetes (McAulay et al., 2001; Fanelli et al., 2004). Signs include sweating, tremors, anxiety, irritability, fading, blurry vision and tachycardia (McAulay et al., 2001; Fanelli et al., 2004). These can lead to neuroglycopenia with subsequent seizures, coma or abrupt death (McAulay et al., 2001).

Asymptomatic and mild hypoglycaemia can be recurrent and eventually cause hypoglycaemia unawareness. Short intervals between mild hypoglycaemic episodes and the inability to identify them may lead to severe hypoglycaemic crises, threatening the patient's life (Fanelli et al., 2004; Tauveron et al., 2006). Early awareness of hypoglycaemia symptoms allows patients to take action and prevent further complications (McAulay et al., 2001).

DAD indicate potential drops in patients' glucose levels by changing their behaviour (Bijland et al., 2013). There are reports about animals spontaneously alerting their owners to a crisis without previous training (e.g. Chen et al., 2000; Tauveron et al., 2006; O'Connor et al., 2008). Of 212 dog owners with Type 1 diabetes, 138 subjects reported variations in their dogs' behaviour during a hypoglycaemic episode (i.e. vocalising, intensive licking, jumping or nuzzling). Reports indicated that 40.8% of the dogs reacted almost simultaneously to the owner becoming aware of the glycaemia imbalance, 33.6% responded earlier, 19.2% of the dogs responded after the owner had taken action to balance their glycaemic levels, and 6.4% after the owner had fallen unconscious (Wells et al., 2008).

There is an increasing awareness of these animals' potential to detect such episodes (Wells et al., 2008). Hence, dog organisations and private trainers perform specialised Assistance dog training (Tauveron et al., 2006; Rooney et al., 2013; Petry et al., 2015). Studies based on owners' reports indicate high satisfaction and improved quality of life since obtaining the dog. For instance, a survey exploring the ability of DAD, was answered by 36 owners and indicated a significant decrease in hypoglycaemic episodes, diminished worries about them and higher chances of maintaining a normal lifestyle (Gonder-Frederick et al., 2013). These studies gave valuable evidence on DAD potential. However, they are insufficient, based on subjective retrospective data, and not associated with objective measurements of dogs' alerting performance.

Several studies were based on objective measures. Hardin et al. (2012) assessed the effectiveness of four DADs when discriminating in vitro perspiration samples. This showed high sensitivity variation (20-100%), although specificity rates were more homogeneous (71-88%). In a similar test, three dogs discriminated skin swab samples from hypoglycaemic vs euglycaemic patients, and dogs performed accurately on half the occasions (Hardin et al., 2013). A third in vitro test with a slightly larger dog sample (N=6) revealed lower sensitivity variation (50%- 87.5%) (Hardin et al., 2015). However, these experiments assessed small dog samples and only focussed on scent discrimination accuracy.

Other studies have examined both owners' reports and objective measures. Rooney et al. (2013) interviewed 17 owners with T1D, investigating changes in quality of life since they lived with a DAD. They compared the records of their dogs' alerting behaviour with changes in the patient's blood levels. Of the nine clients analysed (eight were excluded from analysis due to inaccurate records) The study found that eight dogs indicated significantly more frequently when glucose levels were out of normal range and five had a reduced incidence of hypoglycaemia episodes since obtaining the dog. In addition, owners reported enhanced quality of life, greater independence, reduced hypoglycaemic episodes and less need for paramedic services. The two dogs that did not significantly indicate glucose changes were not certified and presumably had training problems.

In research assessing alerting frequency of 18 DADs, dogs alerted an average of 59.2% of hypoglycaemic and 56.1% hyperglycaemia episodes in their owners with T1D (some dogs generalised and alert hyperglycaemia as well), but with high variability (33% to 100% of hypoglycaemic episodes) (Gonder-Frederick et al., 2017). However, in Los et al. (2017), alert rates were lower (36%), although owners also showed high levels of satisfaction and confidence in their dogs' alerting performance. Rooney et al. (2019) associated owner reports with blood records and alert accounts and showed higher reliability in alerting than Los et al. (2017), with an average sensitivity of 79%. However, these studies differed in methods and the parameters to quantify correct indications. For instance, Los et al. (2017) assessed a small and diverse dog sample (eight dogs), while Rooney et al. (2019) studied a larger dog sample (28 dogs) from a single training organisation. These differences make comparisons between studies challenging, although the latter's' findings may be more reliable as the sample was larger and more homogeneous.

DAD studies have primarily focussed on owners' satisfaction. Dog alerting behaviour is mostly assessed through owners' reports, and the dog's sensitivity is recorded from the client's self-monitoring of glucose levels. However, experiments recorded the behaviour of eight DAD with a CCTV system in their owner's household, paired with the data from a glucose monitoring scanner (Wilson et al., 2019; Wilson et al., 2020). All DADS alerted significantly more frequently to hypoglycaemia levels than in-range levels, and six also indicated hyperglycaemia states. Dogs' sensitivity varied widely across individuals when alerting out-of-range episodes (Mean= 55.9%, Min=33.3 Max =91.7%) (Wilson et al., 2019). All dogs significantly altered at least one behaviour when the owner shifted from normal glycaemia to out-of-range glycaemia. In addition, dogs tended to display individual behavioural patterns related to attention-seeking and activity, and the variation of three behaviours changed for all dogs during out-of-range episodes, i.e. 'Playing with Owner', 'Pawing Owner' and 'Barking' (Wilson et al., 2020).

The role of Assistance dogs in supporting patients with epilepsy has also been studied (e.g. Strong et al., 1999b; Dalziel et al., 2003; Brown and Goldstein, 2011). Epilepsy alerting dogs perform different tasks: 'Seizure Alert Dogs' indicate an oncoming seizure, while 'Seizure

Response Dogs' act in a specific way during or after an epileptic episode to assist the patient (Kirton et al., 2004; Brown and Goldstein, 2011). Since the early 1990s, there have been reports of dogs showing spontaneous behavioural changes towards owners' epileptic episodes (Pflaumer, 1992; Edney, 1993). Later research showed the potential to train dogs to detect seizures (Strong et al., 1999b; Strong et al., 2002).

The dogs' ability to alert epileptic seizures was studied through surveys. A qualitative questionnaire identified nine out of 29 dogs which reacted spontaneously to their owner's epileptic events (Dalziel et al., 2003). A later survey assessing the spontaneous reaction of 62 untrained pet dogs living with children with epilepsy saw more than a third showing some response to seizure episodes (Kirton et al., 2004). Most of these dogs were large-sized and mixed breeds, the most frequent response was licking the child's face and other attention-seeking behaviours and 41% of the dogs alerted the seizure before its onset. However, in a questionnaire focused on trigger factors and early symptoms before epileptic seizures (Pinikahana and Dono, 2009), only 26 of 187 pet-owning respondents indicated that their pets responded to seizures, although these also included cats. The most common pet behaviour was sitting and staring at the owner (48.6%). A survey evaluating a training programme for seizure alert dogs showed that all respondents (N=22) considered their dogs' alerting behaviour reliable and that their own quality of life improved since having the dog. Also, 13 dogs developed spontaneous alerting behaviours and 6 dogs, trained to activate alert emergency systems, performed this reliably (Kirton et al., 2008).

Other studies focussed on dog-patient video electroencephalographic observations reported limited dog alerting during seizures (Ortiz and Liporace, 2005; Doherty and Haltiner, 2007). However, these were inconclusive since samples were small (based on single dog cases) and included patients with non-epileptic seizures.

Research indicates that an odour change might occur before seizures (Brown and Goldstein, 2011; Davis, 2017; Maa et al., 2021). A study comparing odour profiles from 11 subjects with epilepsy and healthy participants using gas-chromatography/mass-spectrometry methods showed significant differences. It also identified that four dogs could discriminate differences in biomarkers from patients with epilepsy during scent training trials (Davis, 2017). A further study confirmed the dogs' ability to distinguish an epilepsy scent (Catala et al., 2019). This assessed five dogs' (pre-selected for service tasks) ability to discriminate peri-ictal sweat samples (during or right after a seizure) collected from five patients with epilepsy from control scents of the same subjects (collected when exercising and resting). Results showed high levels of sensitivity (Mean=86.8%, Min=67% Max =100%) and specificity (Mean=98%, Min=95% Max =100%). A more recent study indicated that trained dogs could identifyspecific epilepsy pre-ictal (before a seizure) and peri-ictal (during a seizure) sweat samples with a 93.7% accuracy (Maa et al., 2021). Nevertheless, there was no significant association with those from patients with non-epileptic seizures. These findings suggest epilepsy-specific scent bio-markers could help target dogs' training more accurately.

1.2.1.6 How effective are MDDs in their roles?

Proof of principle evidence suggests that MDD can effectively detect VOCs of different conditions. However, several factors may affect MDD performance, such as dogs' individual variation due to genetics, behavioural characteristics and background; small dog sample sizes and external factors, i.e. training methods and scent sample management. These factors affect the feasibility of employing these dogs widely within the health sector (Reeve and Koivusalo, 2018).

a. Dog samples sizes and background

Most studies have assessed small dog sample sizes. This is frequently related to insufficient availability of specialised dogs or some not progressing to more complex training stages due, for example, to the development of unwanted behaviours, lack of searching skills or low tendency to generalise scents. Even if some dogs can be highly proficient in a search task and samples may be good enough for proof of principle studies, replication with a larger number of animals is needed to improve research validity before employing them regularly (Lazarowski et al., 2020a).

Dogs' genetics and background/ experience influence the dogs working performance (Jamieson et al., 2017; Konno et al., 2018). In these studies, most dogs were not specifically reared or selected for medical detection purposes. Studies mainly assessed companion animals from varied breeds (Willis et al., 2004; Rooney et al., 2013) and dogs from working lines trained initially for other tasks (Cornu et al., 2011).

Walczak et al. (2012) found significant differences among dogs trained to detect cancer in the number of trials to achieve scent indication. Younger dogs tended to learn significantly faster in the initial training phases than older dogs. Still, older dogs required fewer commands and had higher sensitivity than younger dogs during more complex training trials.

Ideally, studying dogs bred for MDD in standardised conditions may help understand variation and homogenise performance (Elliker et al., 2014). However, this is difficult since the introduction of MDD is relatively recent, and there is limited availability of MDD (Jezierski et al., 2015). However, several studies have assessed dogs from an institution specialising in MDD training, the charity Medical Detection Dogs® which has standardised socialisation and training protocols (e.g. Concha et al., 2019; Guest et al., 2019; Rooney et al., 2019; Wilson et al., 2020).

b. External factors

MDD ability may be influenced by external factors. Instructors' ratings suggested that DAD performance was affected by the client's age, the severity of their condition, communication with their instructor, household size and inclination to reward their dog's alerts (Rooney et al., 2019).

Dogs' training methods are variable and may be related to the MDD ultimate performance. In Bio-detection dogs, some studies investigated the association between training complexity and performance and found the dogs' ability to discriminate cancer odours decreased as the task became more complex (Walczak et al., 2012; Elliker et al., 2014). When assessing the ability of dogs trained to detect prostate cancer, of ten dogs initially trained, only three progressed to the second training stage, and none were able to discriminate S+ in double-blind testing sessions (Elliker et al., 2014). The high failure rates were attributed to behavioural variation, a potential tendency to memorise odours instead of achieving generalisation, and the reward schedule used. In one study, dogs were trained to detect COVID-19 from armpit samples. Of 18 dogs, only eight got to the last stages (Grandjean et al., 2021).

The number of samples, types, sources and management may affect bio-detection training (Walczak et al., 2012). Types of cancer are variable. Hence different sample types are used for training and testing. Urine samples are used for bladder and prostate cancer detection (e.g. Willis et al., 2004; Cornu et al., 2011), breath samples for lung and breast cancer trials (e.g. Willis et al., 2004; Walczak et al., 2012) and tissue and blood for ovarian cancer (Horváth et al., 2008). COVID-19 studies used diverse sample sources, e.g. human sweat from armpits, worn clothing or masks, saliva or urine samples. However, the scent sample variation limits the ability to compare studies directly (Jezierski et al., 2015).

c. Individual variation in performance

In both Bio-detection and Assistance dogs, studies have shown a wide variation of scent sensitivity and specificity. For instance, for lung cancer detection, sensitivity ranged from 56% to 99% and specificity from 32% to 90% (See Jezierski et al., 2015). In a study assessing dogs detecting COVID-19, mean sensitivity was high (82.63%) and showed wide variation (From 67.9% to 95.2%), while specificity was higher (96.35%) and hardly varied (92.4% to 98.9%). These dogs may tend to ignore S- correctly but overlook S+ more often, which was attributed to dogs' individuality, diverse background and short training periods (seven days), and the nature of the samples (Jendrny et al., 2020).

Research showed wide ranges in sensitivity and specificity also in DADs (e.g. Hardin et al., 2012; Los et al., 2017), suggesting that the dogs' reliability in alerting is variable. For instance, in Rooney et al. (2019), dogs had a 79% median sensitivity, higher than required for initial certification (75%). However, sensitivity was below this parameter for ten dogs, and for three dogs, it was below 50%. The authors suggested that although the majority of dogs were reliable in alerting, for several dogs, their post-certification performance may be affected by different factors, such as clients' additional health issues or dog training problems.

The association of behavioural traits and performance in MDD has seldom been explored in contrast with other more established working roles such as explosives, drug detection dogs, and guide dogs (1.2.6). In theory, different dogs will likely have different propensities to perform well in their roles, and this needs to be adequately explored in MDD as in other working dogs.

In Bio-detection dogs studies, it was mentioned that 'Motivation', 'Acuity of the sense of smell' and 'Trainability' might be more important than dogs' breed for cancer detection (Jezierski et al., 2015). Walczak et al. (2012) attributed a decrease in performance of young cancer detection dogs to low motivation in unrewarded trials. It has also been considered that fatigue, boredom and distractibility may affect the performance of Bio-detection dogs as they are in a room performing a relatively "monotonous" task in contrast with other roles that involve working outdoors (Reeve and Koivusalo, 2018).

In Assistance dogs, ratings from human-dog partnership instructors of behavioural factors influencing performance, such as 'Dog's Motivation/enjoyment of the task', 'Dog's Willingness to try new behaviours and get it wrong' and 'Strength of dog's indication', correlated with increased sensitivity to changes in glucose levels and fewer incorrect responses in owner-collected data (Rooney et al., 2019; Wilson et al., 2019).

Recently, in a survey of owners of untrained epilepsy alerting dogs (N=72), dogs reported to alert to seizures spontaneously had higher ratings on 'Motivation', ' Training focus', and 'Amicability' and lower for 'Neuroticism' compared to dogs not alerting seizures. However, ratings were not compared with objective measures (Catala et al., 2020).

Similar to other service dog tasks, Assistance dogs may need to be highly adaptable to different environments and focus on their owner, but also to have high olfactory skills. As in other detection tasks, Bio-detection dogs may need to be active, interested in the environment, and have high searching ability (Jamieson et al., 2017; Bray et al., 2021b). However, it is still necessary to investigate the effect of individual attributes on MDD performance to better profile Bio-detection and Assistance tasks and improve selection.

Bio-detection dogs could potentially aid disease diagnosis, and Assistance dogs help prevent life-threatening crises, improve patients' quality of life and save lives. Despite the low number of dogs and the high variability in experimental methods, the research today is valuable since it shows that disease detection is possible. Nonetheless, more research is needed, with larger sample sizes and greater standardisation in methodologies. Investigating how individual differences affect task aptitude can provide vital information to increase success rates and the feasibility of using dogs' olfaction within the public health sector.

This dissertation addresses the association between individual differences, the functional ability of Assistance and Bio-detection dogs trained for the above conditions, and the link with different proficiency measures in their tasks.

The following sections address existent research on the link between dog individuality and performance of diverse working roles, evaluation methods to measure this, and main findings on traits associated with dogs' operational success, potentially relevant for MDD.

1.2.2 The importance of studying working dogs' behaviour: individual variation

Knowledge of the influence of individual differences on task performance is crucial to choose the best-suited individuals for a given task (Cobb et al., 2015; Bray et al., 2021b). This could reduce the number of potential mistakes that could jeopardise human lives, such as failure to detect an explosive device or alerting to a hypoglycaemic crisis in medical detection dogs (Rooney et al., 2016; Rooney et al., 2019).

Breeding, rearing and training a working dog is time-consuming (Goddard and Beilharz, 1983) and represents a high economic expense. Previous research estimated that training working dogs might cost between $20,000 and $50,000 ten years ago (Brown and Goldstein, 2011). However, for some disciplines, approximately 50%-70% of dogs are rejected mainly due to behavioural reasons (Serpell and Hsu, 2001; Cobb et al., 2015; Bray et al., 2017a).

These losses cause inefficiencies in the working dog industry, both in economic resources and time investment (Sinn et al., 2010; Cobb et al., 2015). If sub-optimal dogs are selected, they may have a shorter working life (Cobb et al., 2015). A dog unable to cope with the high-stress levels from challenging situations that specialised dogs are constantly exposed to, might respond with fear or aggression, increasing the risks to those around it and compromising its welfare (Slabbert and Odendaal, 1999; Haverbeke et al., 2010b; Rooney et al., 2016).

1.2.2.1 Research into individual variation in working dogs

Research into dogs' individuality is gaining high relevance. This is probably due to a higher acceptance of dogs showing consistency in behaviour traits equivalent to personality in humans and the ability to display emotions (Jones and Gosling, 2005; Kubinyi et al., 2009b; Burman, 2014; Riemer et al., 2014c).

Studies have addressed the link between dogs' attributes and their working ability in diverse roles (e.g. Rooney et al., 2004; Maejima et al., 2007; Batt et al., 2008); some are related to temperament, and others to cognitive aspects (e.g. Bray et al., 2017b; MacLean and Hare, 2018), as well as external factors (e.g. Jezierski et al., 2014).

Temperament refers to fundamental behavioural tendencies that appear early and are relatively consistent throughout life. They regulate behavioural expressions and are increasingly influenced by experience and environment (Goldsmith et al., 1987), while personality is defined as “those characteristics of individuals that describe and account for temporally stable patterns of affect, cognition and behaviour” (Gosling, 2008). These definitions are the subject of debate (Jones and Gosling, 2005). Both concepts have similar implications and are regularly used as equivalents (McCrae et al., 2000).

Temperament is shaped by an intricate combination of both internal and external factors: from genetic, neurophysiological, endocrinological and developmental aspects to early experience and learning throughout the animal's lifetime (Stur, 1987; Saetre et al., 2006; Reale et al., 2007; Riemer et al., 2014b). Examples of temperament traits are sociability, fearfulness, confidence and playfulness (Serpell and Hsu, 2001; Jones and Gosling, 2005).

Cognition refers to "mechanisms by which animals acquire, process, store and act on information from the environment" (Shettleworth, 2009, p. 4). Cognition influences performance allowing flexibility and innovation in behaviour (Bray et al., 2017b). Cognitive functions in dogs include problem-solving, decision-making, social referencing, memory and self-inhibition (Bray et al., 2017b; MacLean and Hare, 2018).

Personality traits are consistent over time and contexts and should be measurable by scientific methods (Fratkin et al., 2013). Some characteristics that showed high consistency are playfulness, sociability, exploratory activity, impulsivity and distractibility (Svartberg et al., 2005; Riemer et al., 2014a; Riemer et al., 2014b). In working dogs, consistency in behaviour brings predictability and helps infer how a dog will respond to different situations (Fratkin et al., 2013).

Several behavioural assessments have explored temperament aspects by assessing the dogs' response to startling stimuli or novel situations (e.g. Slabbert and Odendaal, 1999; Svartberg, 2002; Harvey et al., 2016). However, current research has also focussed on cognitive functions, evaluated through, e.g. problem-solving, memory and social referencing tasks (e.g. Miklosi et al., 2004; MacLean and Hare, 2018) (1.2.6.9). It has also examined cognitive bias, where an animal may show 'Optimistic' or Pessimistic' tendencies when making decisions (e.g. Mendl et al., 2009; Mendl et al., 2010a) (1.2.6.10).

Studies examining individual differences in detection dogs have mainly assessed explosives and narcotic detection tasks (e.g. Rooney et al., 2004; Maejima et al., 2007) but seldom evaluated recently developed tasks such as MDD.

The present review summarises the main findings across scientific literature about individual attributes associated with scent detection and other tasks, the methodologies applied and their potential relevance for medical detection tasks.

1.2.3 Reliability, validity and feasibility of methodologies to assess dogs' behavioural traits

Researchers have attempted to validate methodologies to identify how individual attributes- affect working dog performance. Dog evaluation methods ideally require three main aspects to be effective: reliability, validity and feasibility (Jones and Gosling, 2005; Svartberg et al., 2005).

1.2.3.1 Reliability

Reliability refers to how consistent the reports within and between observers are about the animal's performance during the same or different tests (Taylor and Mills, 2006; Sinn et al., 2010).

Reliability measures include a. intra-observer reliability - the tester's consistency when scoring an animal; b. inter-observer reliability - the agreement between two or more observers when evaluating the same dog; and c. test-retest reliability - the dog's consistency when tested on different occasions. (Jones and Gosling, 2005; Taylor and Mills, 2006; Fratkin et al., 2013).

Some researchers have reported reliability measures for working dog assessments, e.g. in guide dogs (Batt et al., 2008; Harvey et al., 2016), explosives detection dogs (Rooney et al., 2007b) and military dogs (Sinn et al., 2010). However, reliability reports are limited, seldomly applying quantitative measures and significant statistical variation (Knol et al., 1989; Jones and Gosling, 2005; Sinn et al., 2010). When developing a selection test for juvenile guide dogs, Harvey et al. (2016) found high intra-observer and test-retest reliability for most of the traits assessed. However, they did not consider inter-observer reliability since a single person did all the behaviour coding. Rooney et al. (2007b) paired four trainers to rate 26 explosives detection dogs over training. The ratings were significantly positively associated with objective records in a search task and showed high intra and inter-observer reliability. Test-retest reliability is critical since temperament traits should be consistent across time. However, this can vary between individuals due to intrinsic factors or variability in response to novelty, or it may be that the test was not conducted the same way across repeated testing (Fratkin et al., 2013; Henriksen et al., 2020). When assessing the predictive value of a military dog test, Sinn et al. (2010) found generally good test-retest reliability. However, this decreased with longer intervals between tests.

1.2.3.2 Validity

Validity refers to how useful and adequate a test is and the extent to which it measures the object of evaluation (Taylor and Mills, 2006; Meagher, 2009). It includes different criteria:

a. 'Construct validity'–the extent to which a particular measure evaluates a specific attribute; b. 'Criterion validity'- quantifies the association between measures of dog behaviour and an external standard that is supposed to assess the same measure. c. 'Convergent validity' – the extent to which a test provides a measure that theoretically agrees with another (validated) test assessing the same construct; d. 'Divergent' or discriminative validity' – the extent to which the test does not measure a construct it is not designed to measure, and. e. Predictive validity– the extent to which the test can predict behaviour indicative of the same construct in other related contexts (Jones and Gosling, 2005; Taylor and Mills, 2006; Meagher, 2009).

Predictive validity is fundamental in working dog training programmes as it indicates the functionality of the evaluation tool to predict dog task success (Sinn et al., 2010). This is of particular interest for this research since it assesses how MDD individual attributes aid in predicting successful task performance.

1.2.3.3 Feasibility

Feasibility considers how practical it is to administer and repeat the test in different contexts. Tests that are long or highly complex may be impractical to duplicate, as well as those using costly materials or performed in complicated access settings (Taylor and Mills, 2006).

1.2.4 Methods for quantifying behavioural differences in dogs

Different dog evaluation methods regularly involve coding –a relatively 'objective' measure, i.e. a specific behavioural response according to an ethogram, or subjectively rating a range of individual attributes desired for the job (Jones and Gosling, 2005; Sinn et al., 2010; McGarrity et al., 2016). Both codings and ratings have been shown to be predictive of dog training outcomes (McGarrity et al., 2016). The most utilized assessments are test batteries, questionnaires and behavioural observations.

1.2.4.1 Test battery

The test battery is usually composed of several subtests designed to measure traits potentially linked with dog performance. It conventionally records and scores or rates a dog's response to a specific stimulus based on an established rating or coding scheme (Serpell and Hsu, 2001; Jones and Gosling, 2005). It may combine coding behavioural responses (from an ethogram) and subjective ratings (e.g. Svartberg, 2002; Polgar et al., 2019). Examples of subtests include facing the dog with a startling or novel stimulus, such as opening an umbrella near them or assessing their behaviour in an unknown room (Campbell, 1972; King et al., 2003; Bray et al., 2017b.

A subtest commonly measures single or multiple behaviour items, later correlated through statistical modelling and gathered in composite factors (Sinn et al., 2010)

This method allows measuring the animals' responses more 'objectively' than questionnaires, and most test batteries do not require any training. However, dogs are frequently tested only once or twice, showing a glimpse of the animal's responses but not necessarily reflecting their general behaviour. Testing a dog once could not demonstrate construct consistency unless it has been previously validated.

Some authors consider that replicating real contexts when testing dogs improve the test's predictive value (Taylor and Mills, 2006; King et al., 2012). For instance, in drug detection dogs, Jezierski et al. (2014) tested dogs' ability in simulated search scenarios, and Harvey et al. (2016) evaluated the behavioural response of prospective guide dogs to real-like conditions in different sites, i.e. villages and community centres. However, tests are unlikely to cover all the possible eventualities in real situations. Some subtests are highly impractical, such as taking the dog to a woody area surrounded by human-sized paper figures (Wilsson and Sundgren, 1997b). However, most tests are performed in a controlled environment, allowing greater standardisation and fewer potential distractions (e.g. Bray et al., 2017b; MacLean and Hare, 2018).

Some of the subtests of temperament tests can be highly intense for the animal, eliciting excessive fear and the potential to sensitise them (King et al., 2003; Rooney et al., 2016), e.g. response to a gunshot or a stranger striking the dog with a rag (Campbell, 1972; Slabbert and Odendaal, 1999). However, more recent testing methods tend to utilise less intense stimuli or involve cognitive measurements (e.g. Bray et al., 2017b; MacLean and Hare, 2018; Tiira et al., 2020) to obtain information about the dog's behavioural predispositions in different contexts.

1.2.4.2 Questionnaires

Questionnaires allow a quicker study of larger samples without geographic limitations (Serpell and Hsu, 2001; Rooney et al., 2004; Kubinyi et al., 2009a). These instruments have shown high validity (Duffy et al., 2008), but they depend on the subjective opinion of individuals close to the animal (Clark et al., 2020b).

The Canine Behavioral Assessment and Research Questionnaire (C-BARQ©) has demonstrated high reliability and validity. Serpell and Hsu (2001) developed this psychometric scale to investigate temperament traits in juvenile guide dogs and (Hsu and

Serpell, 2003) adapted it for pet dogs. This scale has been widely applied, e.g. to examine aspects of behaviour in different animal populations (e.g. Duffy et al., 2008) for criteria validation of guide dog tests (Harvey et al., 2016) and to investigate differences between search and rescue and pet dogs' behaviour (Hare et al., 2018).

Other surveys were designed explicitly for working dog roles. A dog handler and trainer questionnaire investigating the most important traits of explosives and drug detection dogs (Rooney et al., 2004) was later adapted to examine differences between handlers and trainers of two different breeds of drugs and explosives detection dogs (Adamkiewicz et al., 2013). Another questionnaire was developed to assess traits related to certification in prospective guide dogs showing high-reliability consistency and predictive validity for several traits (Harvey et al., 2017).

Questionnaires have also focused on specific aspects of behaviour. For instance, the Dog Impulsivity Assessment Scale (DIAS) was initially validated for pet dogs (Wright et al., 2012) and later applied in different companion and working dog studies (e.g. Riemer et al., 2014a; Fadel et al., 2016; Brady et al., 2018). The Positive And Negative Affect Scale (PANAS) (Sheppard and Mills, 2002) measures diverse dogs' emotional tendencies. It has been used to assess police dogs' long-term success related to the dogs' emotional states (Brady et al., 2018).

Questionnaires allow gathering reliable and quantitative data while still relying on human opinion and revealing a broader picture of the animal (Rooney et al., 2007b). Yet, respondents may have diverse experiences and no training on the scale items and may interpret them differently (Bray et al., 2021b; Clark and Rooney, 2021).

1.2.4.3 Behavioural observations

Behavioural observations involve recording and coding behaviours in a specific context, frequently in their regular activities, based on an ethogram. This allows non-invasive data collection (Dalla Villa et al., 2013). Sometimes test batteries include behavioural observations (e.g. Wilsson and Sundgren, 1997a). Research assessing dog welfare frequently includes observations of dog body language potentially indicative of elevated stress.

Scientists have studied housing effects and environmental stress on kennelled companion and shelter dogs' by correlating behaviour observation with physiological indicators (Beerda et al., 1998; Beerda et al., 1999; Beerda et al., 2000; Dalla Villa et al., 2013). Some working dog studies measured dogs' reactions during walks in public areas (Goddard and Beilharz, 1984; Wilsson and Sundgren, 1997a). Haverbeke et al. (2008) recorded dog body postures during training sessions and found an association between lower postures in dogs trained with aversive methods.

1.2.4.4 Integrating measures

Evaluations may correlate objective and subjective measures through association and data reduction statistical analysis, i.e. correlations and principal component analysis (PCA) (e.g. Svartberg, 2002; Tiira et al., 2020). For instance, Svartberg (2002) identified four out of five components as part of the Boldness-Shyness continuum: 'Playfulness', 'Curiosity/Fearlessness', 'Chase proneness', 'Sociability' excluding 'Aggressiveness' and 'Boldness' were related to good performance at different working tasks.

McGarrity et al. (2016) compared the predictive value between ratings and codings for the training outcomes of 52 prospective explosives detection dogs. They found strong associations between both types of measure and training success, suggesting that each measure can provide different information. Trainers' ratings were more consistent over time, while coding measured more context-specific behaviours. Hence, combining both may provide more robust data on the dogs' performance.

In an assessment for explosives detection dogs, Rooney et al. (2007b) compared subjective ratings from trainers with quantitative results from a standard search exercise, finding a significant correlation. In a guide dog test battery, scores from an 'Isolation' subtest were linked significantly with excitability rates from the C-BARQ questionnaire (Bray et al., 2017b). Tiira et al. (2020) assessed explosives detection dogs with a cognitive test battery. They correlated the resulting variables with the dogs' performance in a search task and the handlers' ratings finding significant correlations between them. The experiment indicated that dogs with better inhibitory control were more successful in their tasks.

The current research derived the most relevant traits for MDD performance through a questionnaire for practitioners (Chapter 2) and applied them to develop the test battery, combining objective and subjective measures (Chapter 4).

1.2.5 Measures of dogs' task success

Many studies assess the predictive validity of an assessment by correlating different behavioural measures with success in a working dog training programme (Sinn et al., 2010). Nonetheless, there are no clear measures of working dog operational success. These vary significantly between studies and working dog agencies.

The training programme outcome is a frequent measure of dog performance, a binary measure of 'pass or fail', regularly relying on the criteria of the training agency. Training outcomes can be highly subjective and may not reflect a broad picture of a dog's ability. For example, (Batt et al., 2008) correlated behavioural measurements and cortisol levels with certification in the guide dog training program. Still, trainers selected the dogs according to the organisation's criteria without relying on objective measures. Similarly, Bray et al. (2017b) assessed the value of a test that integrated temperament and cognitive measures with a binary outcome (success/release). However, there was no clear indication about how the dog agency selected the animals.

Performance is multifaceted and influenced by several factors (Bray et al., 2021b; Rooney and Clark, 2021). Therefore, it may be restrictive to measure dogs' task success only by fitting them into a pass-fail training outcome. Multiple measures may capture more diverse information on dogs' abilities. In Rooney et al. (2007b) trainers rated the overall ability (OA) of explosive detection dogs on a scale from one to five, and the dogs' discrepancy from ideal levels of 12 traits on different occasions. The study integrated and derived three performance measures: 'Weighted mean overall ability', 'Weighted mean discrepancy from ideal' and 'Overall ranking'. Subjective measures mostly correlated with measures of dogs' searching ability from a search trial. McConnell et al. (2022) calculated a systematic measure of livestock guarding dogs' aptitude, integrating handlers' ratings for dogs' OA and the difference between ratings on ideal levels, which reflected both the dogs' abilities and deficiencies. These measurements may show a broader range of dogs' aptitude for the task than a restrictive binary measure alone, such as training outcome.

Scent sensitivity and specificity are objective measures for evaluating detection dogs' searching ability (*1.2.1.2*). Several experiments in explosives and detection dog tasks have assessed dogs' behaviour against scent sensitivity and specificity (e.g. Maejima et al., 2007; Rooney et al., 2007b; Tiira et al., 2020), while in MDD, these associations have seldomly been explored *(1.2.1.6)*.

Most dog assessments are performed pre-certification of the training programme. However, long-term dog performance is not frequently investigated (Brady et al., 2018). Dogs tend to be young when certified (Bray et al., 2021b). Although some dog personality traits remain stable over time, others may change with age (e.g. Slabbert and Odendaal, 1999; Batt et al., 2009; Fratkin et al., 2013), potentially affecting working performance. Hence the importance of ongoing monitoring of the dogs' behaviour over their working life towards a successful and long career. Few studies have assessed operational dogs. Brady et al. (2018) analysed factors related to long-term success or failure in already-certified detection dogs from the handler's opinion on items associated with dogs' impulsivity and affective elements. They showed that dogs in service were rated significantly higher in 'Responsiveness' and 'Energy and interest' than those withdrawn early. However, this was assessed retrospectively from the handler's recollection of the retired animal's behaviour, which may be negatively biased if the dog was problematic. Another study investigated behavioural differences of trained drugs detection dogs across breeds against their performance in a search trial (Ganitskaya et al., 2020). It revealed significant associations with 'Play', 'Sociability' and average 'Activity' levels.

Overall, research into operational dogs' performance is scarce and requires further exploration to evidence the long-term predictive validity of selection methods and tests. Combining different subjective and objective success measures aids in increasing assessment validity. Therefore, in this research, I explore and derive diverse success measures (Chapter 3) to investigate how MDD behaviour is associated with task proficiency.

1.2.6 Individual 'personality' traits tested in working dogs

The following section looks at scientific findings on the behaviour of different working dogs. The traits discussed here are those that, after an extensive review of previous research, were identified to be potentially relevant to MDD. Each attribute is discussed separately. However, personality traits are multifaceted, interconnected with other characteristics, and frequently vary across contexts. Therefore, the link with different characteristics is occasionally referred to.

Chapter 2 investigated which traits found relevant for working tasks (described next) are important to MDD. In Chapter 4, these traits were tested in an MDD sample. Some procedures useful to measure these traits in previous studies were selected.

1.2.6.1 Confidence and fearfulness

Confidence and fearfulness are among the most studied traits across working dog tasks (e.g. Svartberg, 2002; Haverbeke et al., 2010b; Wilsson and Sinn, 2012; Overall et al., 2016). Confidence has been described as a trait involving an evaluative process based on past and current experiences, which provides certainty about the favourable outcome of future events (Stankov et al., 2009). Confidence is a dynamic construct that may change over time, as it is based on past knowledge and may update from the outcome of actions conducted over uncertainty (Bandura, 2000; Guennif, 2002).

A confident dog shows a positive, adaptative response to different situations (Bray et al., 2021b). A quick recovery and a calm disposition, when faced with an unknown stimulus are essential for the daily performance of dogs working in public sites, such as Assistance dogs (Bray et al., 2021b) or dogs seeking drugs or illegal items on people (e.g. Rooney et al., 2004). Proactive detection dogs work in extensive areas, e.g. searching for explosives, and should be confident enough to explore novel environments independently (e.g. McGarrity et al., 2016). Military and patrolling dogs who frequently face potentially startling situations like gunshots or explosions are expected to be bold and recover easily (e.g. Svartberg, 2002).

Several test batteries have measured confidence by scoring or coding the animals' approach-avoidance responses or body postures when presented with a novel situation. These have frequently shown that dogs with a higher tendency to approach or adapt quickly to a new stimulus have a higher chance of success in their roles. For instance, Svartberg (2002) assessed companion dogs of breeds generally considered suitable for detection and protection tasks with a personality test where dogs faced potential fear-eliciting situations such as loud noises and the appearance of a dummy. They found that higher boldness in new situations was correlated with better performance in function trials, i.e. searching and protection. In a 'startle' test, dogs who approached faster to a stranger appearing abruptly were more likely to qualify as police dogs.

In questionnaires, Handlers and trainers of explosives and drug detection dogs rated 'Boldness' (Rooney et al., 2004; Adamkiewicz et al., 2013) and absence of fearfulness (Adamkiewicz et al., 2013) to be highly required for good performance. In a study assessing prospective explosives detection dogs, the component 'Nerve stability' from trainers' ratings was linked with 'Confidence' codings from a test battery and was significantly associated with dogs' training outcome (McGarrity et al., 2016).

Confidence in specific conditions, such as kennelling, travelling and physical manipulation, may also be critical to the task(Sinn et al., 2010; Broach and Dunham, 2016). Military and detection dogs may remain in kennels for extended periods; considering dogs' social nature, this can be highly stressful (Rooney et al., 2004; Hiby et al., 2006; Rooney et al., 2007a; Gaines et al., 2008; Broach and Dunham, 2016). Ease of adaptation to kennelling was desirable for trainers and handlers of explosive and drug detection dogs (Rooney et al., 2004). The frequency of stress-related body postures in kennels was significantly correlated with reduced success in a scent search assessment (Rooney et al., 2003a). Transportation by land, air or sea could be distressing for working animals. During an avalanche simulation (Diverio et al., 2016), search and rescue dogs were transported in a helicopter, and their physiological stress parameters increased after landing. However, these returned to basal levels relatively quickly, and there was no significant effect on performance. One aspect related to confidence is tolerance to physical contact. Dogs should be able to cope with physical examinations. Service dogs often require wearing harnesses and service coats. Increased body sensitivity may result from health conditions affecting behaviour (Mills et al., 2020). Studies of guide dogs have shown that successful dogs tended to have lower body sensitivity or be less reactive when manipulated than rejected dogs (Harvey et al., 2017; Bray et al., 2019).

Fearfulness is described as a personality trait that denotes how an individual responds towards threatening situations (Boissy, 1995). Though needed for survival, fear is a negative affective state (Panksepp, 2004). Not being managed appropriately from a young age frequently leads to behaviour problems (Flint et al., 2017). In working dogs, fearful responses may disturb task performance, which is potentially dangerous to society and the dog's welfare (Foyer et al., 2016).

Fear-related behaviours in guide dogs are considered the main reason for failure (Goddard and Beilharz, 1984). When validating the C-BARQ, non-social fear was significantly linked with guide dogs' lacking confidence and fear of thunder, traffic and noise (Serpell and Hsu, 2001). In several studies based on test batteries or behavioural observation, guide dogs showing fearful responses toward salient social and non-social stimuli had a higher propensity to fail training (e.g. Goddard and Beilharz, 1984; Harvey et al., 2016; Harvey et al., 2017). Search and rescue dogs that achieved certification had lower scores on fearfulness-related items in the C-BARQ (Hare et al., 2018).

Assistance dogs often live with a human that may be unpredictable physically (e.g. may collapse during a hypoglycaemic episode or epileptic seizure) and emotionally since people with chronic conditions may suffer from comorbid mental health issues such as anxiety and depression (Beyenburg et al., 2005; Berg, 2011; Naranjo et al., 2019). Patients also may show cognitive alterations during or after a health episode (McAulay et al., 2001). Dogs exposed to uncertain conditions, particularly those without specialised training or a higher fearfulness tendency, may have compromised welfare and their reactions may be unpredictable when handled (Strong and Brown, 2000; Rooney et al., 2016). A dog may display fear-driven responses when confronted with a situation perceived as threatening, including fleeing, fighting or freezing (Gray, 1987). Spontaneous seizure detection dogs were reported to show anxiety-related behaviours when confronted with a seizure, including restlessness, escaping, showing conflict behaviours and reacting aggressively towards the patient or other people (Edney, 1993; Strong et al., 1999b; Strong and Brown, 2000). Assistance dogs’ specialised training aids in developing a positive conditioned response to the occurrence of health crises (Strong et al., 2002; Rooney et al., 2013). However, dogs must be carefully selected for the task, considering confident animals who can recover quickly in different situations and without fearful tendencies (Rooney et al., 2016; Bray et al., 2021b). Appropriate habituation and training should be provided to build positive associations with the conditions the dogs are exposed to when working and reinforce and train appropriate responses (Strong et al., 1999a; Rooney et al., 2016; Bray et al., 2021b).

1.2.6.1 Sociability

Dogs' sociability with people depends on genetic predisposition and social experience (Topal et al., 1997). Different working tasks may require varying levels of reliance on humans (Bray et al., 2021b). Dog-handler cooperation is one of the most critical factors influencing detection proficiency (Lesniak et al., 2008). Working dogs (i.e. military, protection, detection) with more experienced handlers and increased positive interactions were more trainable, less aggressive and performed better (Svartberg, 2002; Lefebvre et al., 2007; Haverbeke et al., 2008). However, dogs who work remotely from the handler, such as explosives detection dogs, are desired to be more independent than dogs who remain close, such as passive drug detection dogs (Rooney et al., 2004; Rooney et al., 2007b).

The tendency to bond with and to pay attention to a human is particularly relevant for service animals, who spend a long time with their partner (Serpell and Hsu, 2001; Brown and Goldstein, 2011; Rooney et al., 2013). In a test battery, German Shepherd Dogs trained as guide dogs scored higher on 'Ability to cooperate with humans' than police dogs (Wilsson and Sundgren, 1997b). In an epilepsy alert dogs’ review (Brown and Goldstein, 2011), the ability to bond with people was described as a key factor for dogs' enrolment in the epilepsy alert dogs’ training programme by Support Dogs UK® (Support Dogs UK, 2022b). In a survey, owners of dogs that alerted epileptic seizures indicated having a stronger bond with them than owners of non-seizure alerting dogs (Catala et al., 2020).

However, extreme levels of attachment or independence can negatively affect performance. Too high independence may disrupt the animal's obedience to the handler when working (Rooney et al., 2004). In contrast, excessive attachment to one person is undesirable when multiple people handle the dogs, e.g. when different handlers manage search dogs or service dogs transition from foster care/training to living with a client. In addition, it may be linked with separation-related behaviours detrimental to dogs’ welfare. In a strange situation test, guide dogs remained calmer when they stayed alone in a room compared to police dogs and pet dogs. However, they presented high cardiac frequency, similar to pet dogs and higher than police dogs. This suggests that guide dogs might be able to control their motor responses due to increased training, but they experience high stress when separated from their attachment figure (Fallani et al., 2007). Bray et al. (2017b)evaluated guide dogs' responses when isolated for two minutes. They found that increased vocalisations were associated with higher excitability rates in the C-BARQ and higher cortisol levels. Another study showed that search and rescue dogs with a greater propensity for separation anxiety were inclined to fail certification (Hare et al., 2018).

Friendliness to unfamiliar people is essential for dogs that are more often exposed to strangers, such as service dogs, drug body search dogs or search and rescue dogs (Jones et al., 2004; Rooney et al., 2004) than for those that work in empty buildings or outdoor natural areas, such as explosives dogs, or conservation dogs (Rooney et al., 2004; Beebe et al., 2016). Several test batteries measure the animals' response to unfamiliar individuals. In Svartberg (2002), dogs that greeted the experimenter more intensely performed better in working trials. However, excessive friendliness might distract the dog while working. In an attachment trial for military dogs, lower involvement with the experimenter was correlated with better searching skills (Rooney et al., 2003a).

1.2.6.2 Tendency to be aggressive towards people

Aggressive behaviours towards people and other dogs are undesirable for most working tasks. It is one of the leading causes of rejection (Serpell and Hsu, 2001), and it is unacceptable in dog roles with more significant public contact, such as Assistance dogs (medicaldetectiondogs.org.uk). Aggressive reactivity causes a risk for society and is a sign of compromised welfare (Netto and Planta, 1997; Duffy et al., 2008). Aggression has been associated with fear and frustration (Serpell and Hsu, 2001; McPeake et al., 2019) and has a high prevalence in some working breeds (Wilsson and Sundgren, 1997a; Duffy et al., 2008; Serpell and Duffy, 2014; van der Borg et al., 2017).

Aggressive reactivity is frequently assessed through test batteries by measuring the dogs' reaction to a potentially aversive stimulus (Duffy et al., 2008), such as the threatening intrusion of a stranger, handling the dog roughly or using an artificial hand to manipulate the dogs' food (e.g. Netto and Planta, 1997; Wilsson and Sundgren, 1997b; Barnard et al., 2012). This exposure can be highly intense, affecting the dogs' well-being (King et al., 2003).

In questionnaires assessing detection dogs, aggression towards humans was rated to be unwanted by handlers and trainers in both explosives and drug detection dogs but thought highly important when selecting a dog (Rooney et al., 2004). But, it was considered more important for dogs screening people for drugs than those searching for explosives since the former are more exposed to public environments (Rooney et al., 2004; Adamkiewicz et al., 2013).

Conditioned aggression is desired in patrol or guarding dogs, and those with higher scores on aggression tests tend to be successful (e.g. Wilsson and Sundgren, 1997b; Sinn et al., 2010). Nevertheless, if trained aggression gets out of control, it could be dangerous for the dog handler and society (Haverbeke et al., 2009).

1.2.6.3 Playfulness with people

Play indicates a positive affective state (Panksepp, 2005). Interspecific play is highly characteristic in dogs and involves an integral part of human-animal interaction (Rooney et al., 2000). Playfulness in dogs is commonly tested by eliciting play engagement with toys and has been related to improved performance for different tasks, i.e. protection and detection dogs (e.g. Svartberg, 2002; Ganitskaya et al., 2020). A strong play motivation is essential for several detection tasks, such as conservation dogs, since play is used as a reward during training (Beebe et al., 2016). For instance, play drive and object orientation were the main selection parameters for grizzly bear tracking dogs (Wasser et al., 2004). Handlers of explosives detection dogs rated off-duty playfulness among the essential traits for explosives detection dogs' performance (Rocznik et al., 2015). Human rescue dogs are trained by olfactory searching for hidden toys or playing "hide and seek" with the handler. Hence they should focus more on playing than on other distractors (Jones et al., 2004). In a test battery, more playful German Shepherd Dogs were faster in a drug searching trial (Ganitskaya et al., 2020). However, excessive playfulness may be disruptive in tasks that require dogs with a tendency to remain calm and focussed, such as service dogs.

1.2.6.4 Trainability

Trainability is an underlying trait that implies a predisposition to learn and be willing to comply with cues. Obedience is the behavioural exhibition influenced by underlying trainability and the dogs’ training (Serpell and Hsu, 2005; Jamieson et al., 2017; Bray et al., 2021b). Trainability and obedience are essential across dog working tasks (Jamieson et al., 2017; Bray et al., 2021b).

In test batteries, obedience is frequently tested by coding or rating the ability to perform basic obedience tasks in diverse situations or to learn new ones during training (e.g. Svartberg, 2002; Harvey et al., 2016; Lazarowski et al., 2018; MacLean and Hare, 2018; Wilson et al., 2020). Test batteries have shown significant associations between dogs' performance in an obedience task and training success in roles such as protection/ detection (Svartberg, 2002), drug detection (Maejima et al., 2007) and guide dogs (Harvey et al., 2016).

In studies based on trainers' and handlers' ratings, 'Obedience to human commands' and 'Ability to learn from being rewarded' were highly relevant for explosives and drug detection tasks. However, 'Obedience' was rated higher for explosives dogs (Rooney et al., 2004). Explosives dogs working remotely from the handler may need better control than passive drug detection dogs that are handled with a leash. (Rooney et al., 2004; Rooney et al., 2007b). In a different study, expert observers scored 'Trainability' as the ease and speed of learning new tasks during training trials and showed that dogs more proficient in detecting a person wearing explosives were more trainable than standard explosive detection dogs (Lazarowski et al., 2018).

Trainability is especially important for service dogs exposed to regular public contact and being relied on by a vulnerable person who might not handle them as consistently as a professional handler. In studies based on guide dog trainers' ratings, trainability was significantly associated with achieving guide dog qualification (Harvey et al., 2017; Bray et al., 2019).

However, for some tasks, independent decision-making is also necessary (Jamieson et al., 2017). For instance, guide dogs must learn to perform without following obedience cues, e.g., avoiding obstructions and traffic (Knol et al., 1988). Occasionally, dogs must show "selective disobedience" by acting contrary to their owner's wishes when their indications may jeopardise them, like crossing a road when it is unsafe to do so (Knol et al., 1988; Audrestch et al., 2015). Assistance dog trainers have anecdotally mentioned that when a patient experiences a crisis, the dog should be able to persist in alerting the owner, who might be in an altered state of mind and ignore the dog or ask it to stop (Personal communication from interviews with trainers of Medical Detection Dogs®).

Internal and external factors may influence the dogs' training performance. The dogs' ability to learn may vary with age. Harvey et al. (2016) evaluated juvenile guide dogs aged five months and again at eight months. Scores from an obedience subtest showed low test-retest consistency attributed to the dogs' immaturity. Although it may be an effect of age, measurements from only two time points may be insufficient to determine this, as factors other than age may have changed as well.

The choice of training method and schedule can influence the dogs' task performance. Each task may differ in the training requirements (Meyer and Ladewig, 2008). Positive reinforcement has been linked with less presentation of aggressive and attention-demanding behaviours (Blackwell et al., 2008) and training based on punishment, with a higher incidence of behaviour problems (Hiby et al., 2004). Working dogs trained with positive reinforcement-based methods and interacting more frequently with the handler tend to show improved performance and less fear-related postures than those punished (Haverbeke et al., 2010a). A study assessed 14 dog-human dyads performing human search and rescue tasks. In a human-dog relationship survey, dogs reported to be handled without physical punishment made fewer mistakes in a searching task, to find a hidden person, than those that were physically punished (Hournmady et al., 2016).

The dog’s reward preference should be considered when reinforcing training (Lazarowski et al., 2020a). External reward usually includes food, play, verbal praise, tactile reinforcement or a combination (e.g. Hournmady et al., 2016). Their use might vary depending on the task, the trainers' inclinations and the dogs' predilections since they can under or over-value rewards (Hayes et al., 2018). In a cognitive test battery for assistance and detection dogs, initially, the dogs were allowed to choose between a food treat and a toy presented simultaneously over repeated trials. The reward chosen most frequently was used during the test as reinforcement (MacLean and Hare, 2018).

1.2.6.5 Motivation when working

Motivation, frequently referred to as ‘Drive’, has been described as innate incentive to perform a behaviour (Brownell, 2002). High motivation is crucial for dogs' performance across working roles (Brownell, 2002; Beebe et al., 2016; Jamieson et al., 2017). It changes over time and is influenced by external factors (Jamieson et al., 2017). This is frequently measured by scoring or rating the level of motivation in behavioural tests or during working trials (e.g. Maejima et al., 2007; Rooney et al., 2007a; McGarrity et al., 2016) or through questionnaires. For instance, handlers rated ‘General drive’ to be critical for explosives detection dogs (Rocznik et al., 2015). In police dogs, the DIAS factors 'Energy and Interest' and 'Responsiveness' were rated higher in successful than in dogs withdrawn early (Brady et al., 2018).

For detection dogs, there is specific interest in the dogs' motivation to search, often referred to as ‘Search drive’ or ‘Hunt drive’. This integrates both motivation and olfactory ability (1.2.7), referred to as a dog's 'Desire to search' (Cablk and Heaton, 2006) and is described as a critical factor for wildlife tracking dogs (Beebe et al., 2016; Jamieson et al., 2017). An "almost obsessive" performance is considered optimal for this task (Jamieson et al., 2017). However, if the tendency to hunt turns into chasing and killing actual prey, it would be a disqualifying factor (Beebe et al., 2016).

'Search drive' was rated highly important for explosives detection dogs (Rocznik et al., 2015). In Rooney et al. (2004), 'Incentive to find an object which is out of sight’ was considered amongst the ten traits that handlers and trainers of explosives and narcotic detection dogs required at higher levels.

A lack of motivation may disrupt scent discrimination training (Lazarowski et al., 2020c). In search and rescue dogs, a high 'Hunt drive' will keep the dog seeking a toy without expecting help from the handler (Jones et al., 2004), which is later applied to human search. In an experiment on explosives detection dogs, motivation to search significantly decreased when dogs were continuously exposed to an explosive-free path. This was attributed to low reinforcement (Gazit et al., 2005). During a breast cancer study, dogs that did not show enough searching motivation throughout training were rejected (McCulloch et al., 2006).

1.2.6.6 Concentration

Concentration allows faster learning and is one of the main factors when selecting explosives and drug detection dogs (Maejima et al., 2007; Adamkiewicz et al., 2013; Rocznik et al., 2015). An animal able to engage its attention in a task will be less affected by the influence of the environment. Maejima et al. (2007) found that ratings by drug detection dog trainers for concentration during training were significantly associated with the component 'Desire to work' derived from a behavioural test, which correlated with training success. In a test battery for guide dogs, dogs’ tendency to sustain attention on the tester against salient toys and treats was negatively correlated with excitability in the C-BARQ (Bray et al., 2017b). Suggesting that guide dogs with a higher ability to focus may have less tendency to be excitable, which can be disruptive to their performance (1.2.6.9).

Conversely, distractibility is an interfering factor in learning and working performance. This was rated by trainers and handlers as undesirable for drug search dogs (Rooney et al., 2004; Adamkiewicz et al., 2013) and has been associated with training failure for this task (Maejima et al., 2007). In guide dogs, high distractibility is among the main reasons for rejection since dogs must deal with multiple distractions when working (Arata et al., 2010; Bray et al., 2021b). For guide dogs, pulling the lead towards other dogs and jumping more in a distraction test was linked with a higher failure probability (Batt et al., 2008). In a further study, the component ‘Distraction’ from trainers' ratings was the main factor associated with failure in guide dog training (Arata et al., 2010). Distractibility showed high individual consistency when prospective guide dogs when tested at five and eight months old (Harvey et al., 2016). This suggests that the tendency to be distracted might be stable across time, potentially a sound selection criterion.

1.2.6.7 Stamina and agility

Stamina is an essential attribute in active detection since these animals should be able to perform long-duration searches without affecting their functionality (Rooney et al., 2007b). This attribute combines motivation, energy and physical ability elements (Jamieson et al., 2017). It is frequently tested by scoring or coding behaviours related to the dogs' endurance in a specific test or a regular working situation (e.g. Svartberg, 2002).

Trainers and handlers rated stamina as one of the top attributes of explosives and drug detection dogs (Rooney et al., 2004; Rocznik et al., 2015). It was also described as crucial for wildlife conservation dogs since they must follow animal tracks in challenging conditions for long periods, and sometimes, scent intensity is very low (Cablk and Heaton, 2006; Jamieson et al., 2017).

Some highly energetic companion dogs left in shelters due to their difficulty to cope with a sedentary lifestyle are considered good candidates for search and rescue tasks as they may fulfil the stamina requirements (Jones et al., 2004). Agility is vital for dogs which operate in complicated areas, such as explosives detection (Rooney et al., 2004; Rocznik et al., 2015), search and rescue (Diverio et al., 2016) and wildlife tracking (Wasser et al., 2004; Arandjelovic et al., 2015). In detection dogs, speed is necessary to complete a search quickly without becoming exhausted (Jezierski et al., 2014). This also relies on physical features such as size and breed, which are relevant when selecting dogs (Hurt and Smith, 2009; Helton, 2010; Jamieson et al., 2017).

1.2.6.8 Calmness and Excitability

Calmness and, conversely, excitability are inherent but influenced by factors such as dog breed, size, age, sexual status, health condition and the environment (Serpell and Hsu, 2001; Miklósi, 2014; Serpell and Duffy, 2014; Bray et al., 2015). These are frequently tested by scoring the response of a dog to a potential arousal-eliciting situation, e.g. greeting a new person (Harvey et al., 2016), their behaviour in a new room (Batt et al., 2008), or being called with a high-pitched voice tone (Bray et al., 2015).

The propensity to remain calm in different contexts is vital for service tasks. However, calmness when not working was also rated as desirable for explosives search dogs (Rocznik et al., 2015). Guide dogs are subject to long periods of inactivity throughout their owner's routine (Fallani et al., 2007), and the ability to remain still is crucial. In a test battery, guide dogs that took less time to rest and spent longer lying in a room had a higher probability of obtaining qualification (Batt et al., 2008). This can be similar in Assistance dogs, in which a calmer disposition is expected for public access.

In the C-BARQ, the factor 'Excitability' describes a dog's tendency to react with high arousal to stimuli, (Serpell and Hsu, 2001). An excitable dog is referred to be highly energetic, and with difficulty to settle down (Serpell and Hsu, 2001). It has been negatively associated with low control inhibition and problem-solving in pet dogs (Wright et al., 2012; Bray et al., 2015; Bray et al., 2017b). In guide dogs, high excitability rated by trainers in questionnaires has been significantly correlated with training failure (Arata et al., 2010; Duffy and Serpell, 2012; Harvey et al., 2017). In the C-BARQ, excitability was associated with guide-dog training programme rejection (Duffy and Serpell, 2012), and in Bray et al. (2017b) was linked with higher reactivity when confronted with novel stimuli in a test battery and higher arousal when left in isolation.

In detection dogs, relatively high excitability and energy are favourable for highly active tasks like wildlife tracker dogs (Jamieson et al., 2017) but less for passive tasks such as body searching narcotic detection (Rooney et al., 2004; Rocznik et al., 2015). When screening people for drugs, a highly excitable dog may be challenging to hold. Certain diseases like Malaria or Covid-19 may require Bio-detection dogs to scan people in public places, and they may need to be less excitable than dogs that perform extensive searches.

1.2.6.9 Cognitive functions

The study of cognition in working dogs is still in early development (Bray et al., 2021b). Cognitive tests evaluate the dogs' capacity to adapt activity patterns to changing spatial and situational contexts, problem-solving skills, working memory, and communication with humans (Miklosi et al., 2004; MacLean et al., 2017).

MacLean and Hare (2018) applied a cognitive test battery developed for pet dogs (MacLean et al., 2017) to explosives, detection and service dogs. Other studies have focussed on specific cognitive dimensions of explosives search dogs (Lazarowski et al., 2019a; Lazarowski et al., 2020b; Tiira et al., 2020). However, an experiment that combined cognitive and temperament measures in guide dogs found better predictive value than purely behavioural or cognitive tests (Bray et al., 2017b). Recently, a cognitive test battery was implemented to disability service dogs aged approximately 9 weeks (mean = 9.20 weeks, min= 7.86, max=10.43 weeks) and again when dogs were under two years old (mean = 1.79 years, min= 0.99, max=2.01 years) and found that several cognitive traits were present early in life and improved with development (although this may be related with other factors other than age like genetic or environment). Several showed high stability over time, suggesting that early assessment of cognitive characteristics may have a good predictive value for working dog selection (Bray et al., 2020; 2021a).

There is evidence suggesting social referencing is valuable for different tasks. Human-dog communication depends on dogs' capacity to rely on human cues (Hare and Tomasello, 2005; Hayes et al., 2018). Dogs can recognise facial expressions and act upon pointing gestures and human demonstrations, using this information to solve problems (Pongracz et al., 2001; Soproni, 2001; Hare and Tomasello, 2005; MacLean et al., 2017). Also, they surpass chimpanzees when inferring the location of a reward hidden inside opaque containers by relying on human gestures, as dogs selected the target container more following human referencing than chimpanzees (Hare and Tomasello, 2005).

Service dogs' success has been attributed to their predisposition to rely on human cues, as they constantly work with their owners (MacLean and Hare, 2018; Bray et al., 2021b). This ability was evidenced in service dog candidates as early as 8-10 weeks old with social referencing tasks but improved with development (Bray et al., 2021a). Graduated service dogs displayed greater social referencing in an unsolvable task (the dog was presented with an inaccessible food reward) than failed dogs. In contrast, explosives detection dogs had a negative correlation between social referencing and training success, possibly because they need to be more independent. Service dogs also tended to gaze at humans for longer when a cooperative activity was interrupted than explosives detection dogs (MacLean and Hare, 2018).

Detection dogs that found a hidden reward by relying on the handler's arm pointing more frequently succeeded in detection training (MacLean and Hare, 2018). In addition, the dogs' tendency to gaze at a human in an unsolvable task increased with age and was predictive of detection dogs' training success (Lazarowski et al., 2019b).

Task persistency has been studied in service and detection dogs (e.g. MacLean and Hare, 2018). Explosives detection dogs with better inhibitory control were more persistent in an unsolvable task (Tiira et al., 2020). However, in Lazarowski et al. (2019b), persistency in this task was stable over time but not related to detection training outcome. The persistence of disability service dogs at an unsolvable task was highly stable over development (Bray et al., 2021a). Assistance dogs should be insistent when alerting a client in an episode, and Bio-detection dogs should ideally maintain the search without giving up, so this trait may also be relevant to MDD.

Dogs apply independent problem-solving, for instance, when making decisions on alerting to an odour or when a guide dog leads their owner in traffic. Prospective guide dogs that solved a multistep puzzle faster and had lower social referencing to the experimenter were generally more successful in their training programme (Bray et al., 2017b). Task speed was also associated with higher trainability and fewer separation-related behaviours in the C-BARQ.

A good memory may influence working roles' proficiency since detection dogs should have a good recollection of the target scent and may be required to identify several scents. Assistance dogs should remember how to act in different situations. Detection dogs performing better in a short memory task had a higher probability of succeeding in their training (MacLean and Hare, 2018). Guide dogs with higher scores in a memory exercise were rated as less excitable in the C-BARQ (Bray et al., 2017b). Impulsivity is defined as the lack of response inhibition when confronted with a relevant cue (Fadel et al., 2016). It has been associated with less tolerance to delayed rewards (Wright et al., 2012; Riemer et al., 2014a; Fadel et al., 2016) and decreased levels of serotonin and dopamine (Wright et al., 2012). When presented with aggression, it has been related to higher reactiveness and a lack of warning of aggressive responses (Fatjo et al., 2005). Impulsivity is usually assessed in choice paradigms involving a delayed reward (Wright et al., 2012; Riemer et al., 2014a; Fadel et al., 2016), motor inhibition in detour tasks (e.g. Bray et al., 2014; Tiira et al., 2020) or through owner or handler ratings with the DIAS (Wright et al., 2011; Wright et al., 2012).

In pet dogs, high DIAS scores have been significantly associated with low tolerance to reward delay and low self-inhibition and have shown to be a reliable tool for assessing impulsivity over time (Wright et al., 2012; Riemer et al., 2014a). High DIAS scores were linked with early retirement in police dogs (Brady et al., 2018). However, the link between DIAS and working dog performance has seldom been assessed.

Inhibitory control is the capacity to withhold impulsive behaviour that may have initially satisfactory outcomes but may be ultimately detrimental (de Ridder et al., 2011). Working dogs should be able to avoid their immediate inclinations and focus on their task. However, the required level of self-control depends on the task (Tiira et al., 2020), with less need for self-control in more active tasks than in service roles.

The link between impulse control and cognitive functions has been evidenced through different tasks. The cylinder task assesses the dog's tendency to inhibit an immediate response to remove a treat inside a clear cylinder (e.g. bumping into it or scratching it) and perform a detour instead (Bray et al., 2014). Detection dogs with fewer mistakes in this task were more successful in an explosive search trial and less likely to give up searching (Tiira et al., 2020). Guide dogs showing greater flexibility in a different detour navigation task had higher confidence when tested in a new situation and had lower cortisol levels (Bray et al., 2017b). In a longitudinal study of service dogs, performance in the cylinder task was highly consistent over development (Bray et al., 2021a). A study compared the ability of disability service dogs and pet dogs to perform a detour of a transparent obstacle when confronted with an increasingly arousing situation (experimenter calling dog with low voice vs high pitched voice). Arousal was measured from video coding of dogs' tail wagging. They found that service dogs had higher basal calmness (less tail wagging during the familiarisation phase) than pet dogs, and their self-inhibition improved as the arousing stimulus increased, while pet dogs were less able to make a detour (Bray et al., 2015). This was attributed to service dogs' preselection to be less excitable and have higher self-control, and to higher individual variation in pet dogs.

The association between MDD performance and impulsivity has not been tested. MDD with a higher tendency to respond impulsively may hasten to indicate odours and have a higher rate of false alarms. Therefore, as part of my research, I evaluated the association of impulsivity and control inhibition with MDD performance (4.4.4). Further details about impulsivity and the DIAS are described in the study (4.1.5).

1.2.6.10 Cognitive bias

Cognitive bias refers to 'the influence of affect on information processing in animals' (Mendl et al., 2010a; Mendl et al., 2010b). Research in human psychology has revealed a link between affective states and cognitive elements such as attention, memory processes and judgement (Mendl et al., 2010b). Positive emotions have been identified in humans as a highly relevant factor for resilience and related to greater adaptation in daily contexts and throughout traumatic events (Bonanno, 2005; Ong et al., 2006; Bonanno and Mancini, 2011).

Judgement bias refers to "the propensity of a subject to show behaviour indicating anticipation of either relatively positive or relatively negative outcomes in response to affectively ambiguous stimuli" (Mendl et al., 2009). One type of judgement bias assesses whether animals tend to make optimistic or pessimistic decisions about the outcome of an ambiguous cue (Harding et al., 2004). Cognitive bias test (CBT) results in companion and shelter dogs suggest that the occurrence of behaviour problems in these animals may be associated with pessimistic decision-making (Mendl et al., 2010a; Karagiannis et al., 2015; Barnard et al., 2018).

In the original CBT adapted for dogs, they were first trained to discriminate between a food bowl in either a 'positive' baited (food) or a 'negative' (empty) location. Then their latency to approach the bowl was assessed when it was placed in intermediate, and hence ambiguous, positions. Dogs displaying more separation-related behaviours took longer to reach the ambiguous locations, suggesting that they judged the likely contents of the bowl 'pessimistically' (Mendl et al., 2010a). Subsequent studies adapted this paradigm to assess the association of CBT with different internal and external factors. An experiment in companion dogs evaluated the association between CBT and behavioural measurements from the C-BARQ questionnaire and a temperament test. It suggested that more friendly dogs with lower rates of non-social fear took less time to reach the ambiguous location.

Dogs displaying more fear or aggressive behaviour towards other dogs and human separation responses showed more pessimistic judgements (Barnard et al., 2018)

Karagiannis et al. (2015) applied a cognitive bias test to five companion dogs displaying separation-related behaviours before, during and after behavioural therapy in conjunction with fluoxetine treatment. They found that the experimental subjects initially took longer to approach ambiguous locations than the control group. However, in the sixth week of treatment, they showed shorter (more 'optimistic') latencies to approach the ambiguous bowls than at baseline and improved clinical signs.

Cognitive bias paradigms are seldom implemented in working dogs. They could reveal important information about the effect of emotions on performance. As cognitive bias may influence the animal's decisions, this might affect the dog's ability to indicate a scent presence. For instance, a dog with a predominantly pessimistic approach might be more 'conservative' (1.2.1.2), less willing to search for odours (e.g. Gazit et al., 2005), not expecting to find the target scent, and then reduce seeking attempts or miss true positives. Meanwhile, a highly optimistic dog may tend to have a 'liberal' approach (1.2.1.2) and generalise the presence of a scent leading to false indications. There might thus be an optimism-pessimism balance that affects detection accuracy. Therefore, the current project explored the association of these tendencies with scent sensitivity and specificity (Chapter 6).

1.2.7 Olfactory detection skills

A detection dog requires acute olfaction and an innate tendency for searching odours (Rooney et al., 2004). 'Acuity of sense of smell' was regarded as essential for explosives and drug detection success (Rooney et al., 2004; Adamkiewicz et al., 2013). Dogs' olfactory thresholds may vary in their receptivity to different VOCs (Walker et al., 2006). Searching ability is influenced by behavioural, physical, genetic, and environmental factors (Phelan and Barnett, 2002; Lesniak et al., 2008; Lazarowski et al., 2020a). Isolating and quantifying the specific role of olfaction receptivity in search tasks is complicated and has not been fully achieved (Jezierski et al., 2015). It is unknown what particular chemicals or chemical combinations the dogs react to, and dogs trained to detect a specific VOC may struggle to achieve generalisation when mixed with other elements (Jezierski et al., 2015).

The ability to generalise odours refers to the dog's propensity to identify scents with a similar composition within different conditions, not only the one it was initially trained to detect (e.g. Walczak et al., 2012; Elliker et al., 2014). Scent detection training frequently begins by presenting a single target scent. Generalisation is achieved by increasing difficulty with sample quantity variations and mixtures (e.g. Williams and Johnston, 2002; Walczak et al., 2012). The importance of scent generalisation varies with the task. It is beneficial, e.g. when DAD indicate hyperglycaemia states when initially trained to identify hypoglycaemia (Rooney et al., 2013). Generalisation is also essential in cancer detection dogs since VOCs may be present in different concentrations, mixed with patients' body scents and affected by sample management procedures (e.g. Walczak et al., 2012; Elliker et al., 2014).

In a cancer detection training programme, increased training repetitions with a small number of S+ and a greater quantity of S- were related to greater accuracy and fewer guessing attempts (Walczak et al., 2012). A prostate cancer training study (Elliker et al., 2014) found that dogs which could distinguish between known prostate cancer odours and controls in initial training stages stopped their progress when new scent targets were added. The authors concluded that dogs could remember many scents instead of generalising a single odour. However, the tendency to generalise scents may be affected by the dog's personality as well as their training.

1.2.8 Other internal factors

1.2.8.1 Dogs' Sex

In some studies, dogs' sex was significantly associated with their working ability (e.g. Svartberg, 2002; Rooney and Bradshaw, 2004), but in others, it was not (e.g. Maejima et al., 2007; Sinn et al., 2010). Some organisations prefer males (Rooney and Bradshaw, 2004), probably due to tradition or the belief that they surpass females when working. Certain studies reported that males scored higher for confidence than females (Wilsson and Sundgren, 1997b; Svartberg, 2002). In cognitive research, males outperformed females in a spatial learning test (Fugazza et al., 2017). Other studies suggested that males may display more aggression than females in drug and explosive detection (Rooney and Bradshaw, 2004) and in guide dogs (Goddard and Beilharz, 1982). Maejima et al. (2007) reported a significant link between neutering with lower distractibility in drug detection dogs. Labradors scored higher for cooperation with humans after neutering than before (Wilsson and Sundgren, 1997b).

1.2.8.2 Dogs' Breed

There is a great physical and behavioural variance among over 400 existing dog breeds (Bradshaw et al., 1996). Dog breed has been associated with differences in behaviour (Mehrkam and Wynne, 2014). However, no breeds are developed explicitly for detection tasks. Therefore, search dog agencies tend to utilise those traditionally intended for other working roles (Rooney and Bradshaw, 2004). Breed preferences differ across roles, working dog organisations and individuals. This will depend partly on traditions and the opinion of those working with the dogs (Hayes et al., 2018). Different studies have suggested that some breeds may be more suitable for various tasks than others (e.g. Wilsson and Sundgren, 1997a; Svartberg, 2002; Rooney and Bradshaw, 2004). The two breeds most frequently reported in working dog research are Labrador Retrievers and German Shepherd Dogs (Jones and Gosling, 2005). Studies have found significant differences between both breeds in temperament and performance at specific tasks. For instance, when tested, German Shepherd Dogs showed higher levels of ‘Sharpness’ (Tendency to use aggression) and ‘Defence drive’ in comparison to Labrador Retrievers and were suggested to be more appropriate for police and protection tasks, while Labrador Retrievers were more cooperative with humans than German Shepherd Dogs and were considered more suitable for guide dogs tasks (Wilsson and Sundgren, 1997b). In Adamkiewicz et al. (2013), handlers and trainers rated Labrador Retrievers trained for explosives detection as less aggressive towards other dogs and having lower object possessiveness than German Shepherd Dogs but less adaptable to the kennel environment than German Shepherd Dogs.

Despite several breeds being initially intended for work purposes, their behaviour has been altered by the growing tendency to select aesthetic traits for dog showing (Svartberg and Forkman, 2002; Svartberg et al., 2005; Mirko et al., 2012). In an assessment of military dogs, gundog lines performed better than show lines in scent-searching tests (Rooney et al., 2003a). Specific trait levels related to work suitability, such as impulsivity, may vary between breeds. For instance, when assessed with the DIAS, working Collies performing herding work were significantly more impulsive than working Labradors specialised in gun work. This may be related to their breeding. However, this difference was not significant between the show lines of both breeds. Overall, working dogs had higher impulsivity than show dogs (Fadel et al., 2016).

Research has regularly centred on specific breeds (Jamieson et al., 2017). Exploring the association between dog breeds and working performance may help improve working dog breeding and selection. However, this relationship is still understudied.

1.2.8.3 Health

Health is not a temperament trait but is affected by internal and external factors (Mills et al., 2020). It is essential for optimal and long-lasting working dog performance (Rooney et al., 2004; Cobb et al., 2016; Bray et al., 2021b). Health detriment can damage the animal's quality of life, being a cause for early retirement (Cobb et al., 2015; Bray et al., 2021b). It can affect a dog's willingness to investigate a scent (Walczak et al., 2012). Breeding to achieve attractive features increases inherited health conditions, which reduce dogs' working longevity (e.g. Banfield et al., 1996). Pain may disturb animals' performance and affect behaviour (Panksepp, 2005; Mills et al., 2020). The dogs' immunity may be affected by chronic stress (Koolhaas et al., 1999). Cardiac, respiratory or orthopaedical alterations are likely to affect the dogs' ability to perform their task effectively (Bray et al., 2021b). The olfactory epithelium function can be affected by trauma, infection and metabolic diseases, causing hyposmia or anosmia, which is especially detrimental for detection roles (Jenkins et al., 2016).

1.2.9 Limitations and future steps

Although research into detection dogs and other working roles is growing rapidly, many gaps should be filled to improve working efficiency. Whereas particular challenges vary with the task, a shared priority is a necessity to produce highly reliable animals that perform their role accurately. However, using a "generic dog fit" would fail to incorporate individual diversity across tasks (Rooney and Clark, 2021). Hence, there is a need to characterise role profiles and to develop task-specific selection instruments highly predictive of working performance to improve dog training, operational cost-effectiveness and welfare (Jones and Gosling, 2005; Cobb et al., 2015; Rooney and Clark, 2021).

Most working dog test batteries have focused on temperament measures based on startling response-eliciting tests (e.g. Goddard and Beilharz, 1984; Slabbert and Odendaal, 1999). Currently, there is promising research in dog cognition which has shown relevant correlations with behavioural responses in pet dogs and other animal species, tending to be consistent over time. However, there is limited research on working dogs. Therefore more investigation on the association between cognitive functions and cognitive bias and working dog performance would aid in understanding the role of dogs' internal processing on decision-making and behavioural flexibility in their roles.

Several studies have assessed dogs' behaviour using codings or ratings only (e.g. Podberscek and Serpell, 1996; Wilsson and Sundgren, 1997a). However, some studies have associated scores from behavioural tests with handlers' ratings (e.g. Harvey et al., 2016; Bray et al., 2017b) or with physiological measures (e.g. Batt et al., 2009; Wright et al., 2012), and few works have correlated temperament and cognitive measures (Bray et al., 2017b). Linking different measurements has shown more representative measures of task performance. An integrative approach is necessary to increase the predictive validity of behavioural assessments.

Several studies use training outcome which, is a binary measure which simply compare pass and fails, so are unable to elucidate factors correlated to optimal performance, and so limits its measurement (e.g. Batt et al., 2008; Bray et al., 2017b). Test predictive validity should not rely exclusively on a binary outcome but on systematic objective and subjective criteria that capture performance diversity (Rooney et al., 2007b). Future behavioural assessments would benefit from deriving different success measures to assess distinct facets of task aptitude.

Working dogs often graduate when still immature. Their performance may change over time or be affected by external factors. More evidence is needed to understand these changes in operational dogs and which factors may be related to maintaining optimal performance. Post-certification monitoring should be reviewed periodically, with evaluations adequate for the role and easy to perform by dog agency staff.

Studies assessing the interaction between behavioural factors underlying scent discrimination accuracy are still limited. Those that have addressed this have focussed more on temperament traits (Svartberg, 2002; Maejima et al., 2007) and less on cognitive functions (MacLean and Hare, 2018; Tiira et al., 2020). There are several detection tasks, but most research has focussed on explosives and drug detection (Rooney et al., 2004; Rocznik et al., 2015). Research into individual variation in more recent tasks, i.e. MDD, is needed to develop task-specific assessment methods. Exploring these is key to achieving better scent discrimination in potentially life-threatening situations.

1.2.10 Relevance to MDD

Research into MDD is affected by small animal samples, low standardisation and a lack of replicability across tasks and studies. Procuring these animals is complicated since MDD is a relatively recent dog role, and there are fairly few animals within training organisations to recruit large samples (Jezierski et al., 2015). There is a rather limited MDD demand from public institutions, probably because of insufficient scientific evidence on their reliability for clinical services.

However, medical detection dog charities are growing, and there is greater public awareness about MDD's positive effect on patients' quality of life. In addition, the COVID-19 emergency showed how joint efforts from scientists and governmental organisations might help develop research rapidly and employ Bio-detection dogs to save lives, which may extend to other conditions.

Evidence is scarce about the influence of individual behavioural characteristics on MDD performance (1.2.1.6). Still, MDD traits may influence their operational effectiveness, differing across bio-detection and assistance tasks.

My PhD project sought to identify and measure individual behavioural characteristics in MDD and to relate these to task performance. I tested how temperament and cognitive traits varied in a sample of these dogs (N=58) and their association with several performance measures. To my knowledge, this is the first research exploring this subject in-depth. I expect to generate valuable evidence to accurately profile MDD across disciplines and improve selection methods that would aid in producing effective operational animals for human health benefit.

1.3 Project aims

This research aims to identify dog traits associated with good performance in medical detection tasks and how these differ across bio-detection and assistance disciplines. It seeks to design a test battery for MDD to explore how temperament and cognitive traits vary in a sample of MDD and explore which of these make them more suitable to perform their tasks. The project also examines how MDD individual attributes are associated with performance at different stages of their working life: it aims to assess if trainee dogs' behaviour is related to success in training. It not only focuses on a binary outcome (pass/fail) but also on how different dog traits may associate with varying ability levels over their training. The tendency to alert scents in operational MDD may vary with their decision-making in discrimination trials. Therefore, this research also examines if trained dogs' attributes are linked with different scent sensitivity and specificity levels in their scent detection projects.

It is hypothesised that:

1) Different dog traits (as tested) are linked with different levels of success in their detection roles.

2) Specific traits in trainee dogs are associated with differences in performance over the training programme and outcome.

3) Specific traits in dogs (as tested) will vary between dogs trained for bio-detection or assistance tasks: these dogs may be selected based on trainers' judgements of their traits and may be reared and trained differently.

4) Different dog traits are associated with different levels of sensitivity and specificity in their detection roles.

1.4 Dissertation structure

The dissertation is divided into seven chapters. (The methods for each are described in each chapter independently since these differ across studies, although some analyses are similar).

1.4.1 Chapter 2. Survey on the importance of individual differences: What do professionals believe are the most important traits for MDD dogs?

To investigate relevant behavioural characteristics of MDD, I first examined which individual attributes were viewed by practitioners as being the most important for the successful performance of these tasks. I generated a questionnaire to identify characteristics relevant to MDD performance. This was answered by 62 trainers and researchers (N=62) from 42 organisations in 16 countries. The survey aided in identifying the most important traits according to the participants and how these varied across bio-detection and assistance dogs. Those traits rated as furthest from ideal in the dogs the participants worked with may indicate potential areas that need improvement. In total, 27 main traits were derived and applied further to the project's later phases.

1.4.2 Chapter 3. How do we measure medical detection dogs' ability? Deriving meaningful measures of MDD performance

Working dogs are frequently selected based on their training accomplishments. However, these decisions are often subjective and not systematic. Hence, I analysed what makes an MDD successful over training and in their daily work to identify what indicates optimal performance and to derive measures of task success used in this project.

Dog trainers rated their dogs at different training time points for prospective bio-detection and assistance tasks. I investigated which dog trainer-rated traits from the survey were associated with their ratings of dogs’ overall ability at different training stages and how consistent they were when rating their dogs. I also explored scent sensitivity and specificity measures provided by the dogs' trainers and the charity Medical Detection Dogs® to derive objective performance measures for trained dogs. These analyses aided in determining measures of success to be used in subsequent analyses of links between behavioural characteristics and performance in Chapters 5 and 6.

This chapter also assesses an internal test from the charity utilised to evaluate MDD aptitude for their prospective task. I investigated the level of agreement between the assessors when rating the dogs, how much variation the traits ratings showed and if the test was predictive of training outcome.

1.4.3 Chapter 4. Development of a test battery to measure individual attributes in MDD dogs: Which variables can be measured? And how do they associate with each other?

I designed a test battery based on the survey's findings (Chapter 2) to assess the most important behavioural traits for medical detection tasks in a sample of 58 MDD from the charity. The test battery included 18 cognitive and temperament subtests. I applied the test battery to trainee dogs for bio-detection and assistance tasks (N=39) and trained dogs enrolled in bio-detection projects (N=19). The test produced 98 variables. After Principal Component Analysis (PCA), these were clustered into 11 components. I assessed their association with the dogs' demography, including their sex, age and training stage, and whether these were linked with the dogs' DIAS scores.

1.4.4 Chapter 5. Associations between medical detection dogs' behaviour in the test battery and their performance: How do these vary for different training stages and tasks?

I explored how the dog behavioural components from the test battery were associated with measures of success. In trainee dogs (N=39), I investigated their relationship with their training outcome and their total ability over training and if these varied according to the specific tasks that the dogs were being trained for. In trained dogs (N=19), I examined if their behaviour in the test battery was associated with their total ability and their levels of sensitivity and specificity in their scent discrimination projects.

1.4.5 Chapter 6. Cognitive bias in medical detection dogs: Does the outcome of a Cognitive Bias test associate with dog performance in MDD tasks?

I assessed the same dogs with a modified version of the Cognitive bias test (Mendl et al., 2010) to investigate whether they were 'Optimistic' (anticipating a reward) or 'Pessimistic' (expecting a negative outcome) when making decisions in an ambiguous situation. I investigated whether dogs' cognitive biases have implications for how they make decisions when trained or in scent discrimination tasks. I explored if these were linked to the behavioural components from the test battery and their DIAS scores.

1.4.6 Chapter 7. General Discussion

This chapter integrates and discusses the most relevant findings from the previous experimental studies to identify the most relevant individual attributes for MDD tasks and how these vary across training stages and disciplines. I propose what their significance is for improving MDD selection, enhancing training and maintaining optimal working performance.

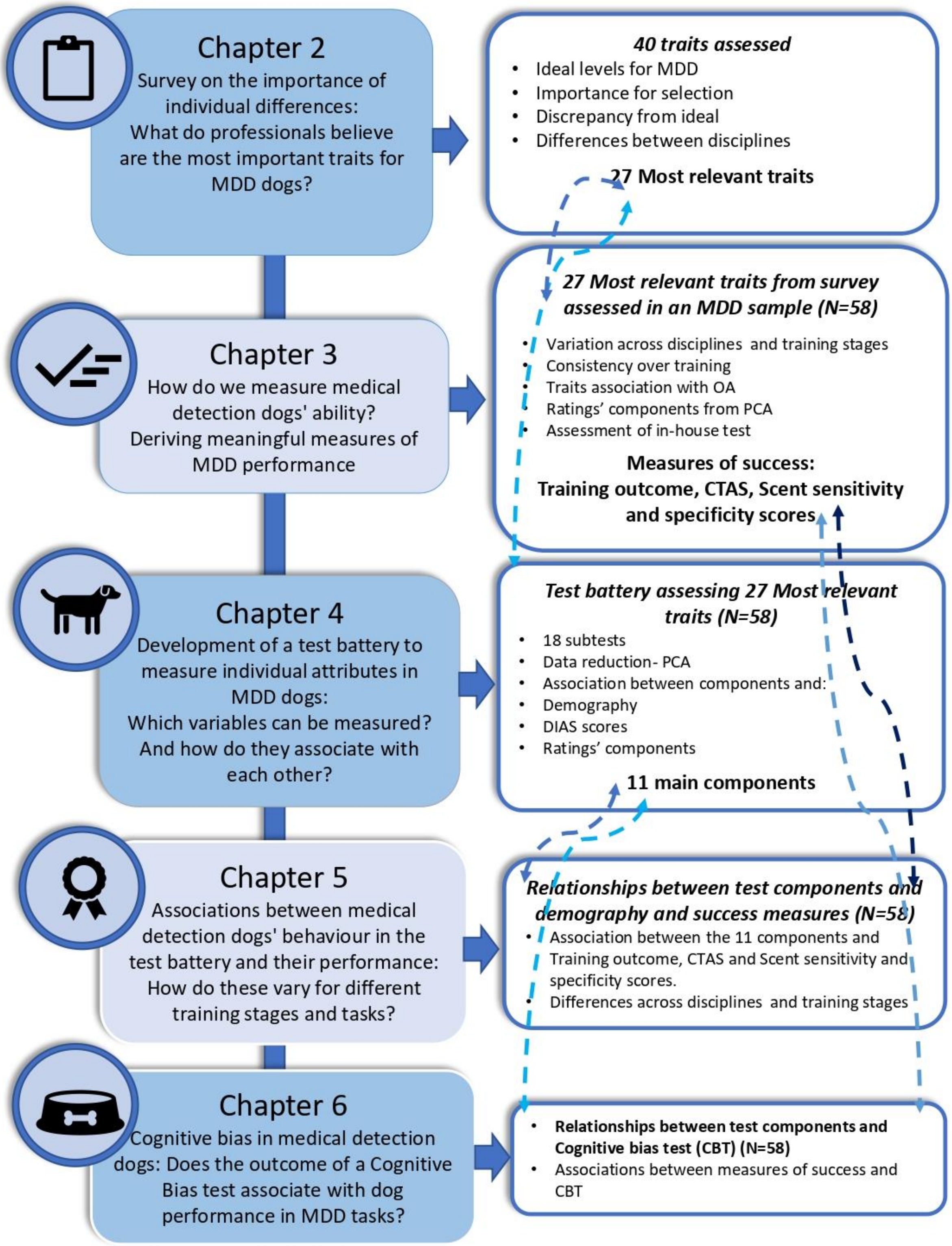


Figure 1. 1 *Flow diagram of experimental chapters timeline, main procedures and outcome measures. Dashed arrows indicate measures derived from a chapter applied in another.*

Chapter *2*. Survey on the importance of different traits for medical detection dogs' performance:

What do professionals believe are the most important traits for medical detection dogs?

Abstract

Medical detection dogs (MDD) are gaining importance in the diagnosis of diseases and in assisting patients with chronic conditions. There has been little research on the behavioural traits linked to good performance for dogs carrying out medical roles.

A survey was developed based on (Rooney et al., 2004) to investigate which behavioural attributes were most important in MDD and also examined how the attributes vary between dogs used for Bio-detection tasks and medical alert assistance tasks.

The survey was completed by 62 professionals from 16 different countries. Participants considered 40 behavioural traits for each rating: the ideal level in MDD, the level of importance when selecting them and the level in a dog with which they had recently worked. Overall, 62.9% of the respondents rated Assistance dogs, and 37.1% rated Bio-detection dogs.

The most important traits for MDD performance from participants' ratings Importance of the traits on a rating scale from 1 to 5 were 'Level of motivation when working' (mean=4.66, SD=0.51), 'Health' (mean= 4.6, SD=0.62), 'Ability to learn from being rewarded' (mean=4.57, SD=0.57), 'Concentration' (mean=4.57, SD=0.60) and 'Acuity of sense of smell' (mean=4.5, SD=0.76). Trait importance differed significantly between the two roles; 'Level of Attachment to human partner' (U=685, p=0.001), 'Confidence in different environments' (U=604 p=0.002) and 'Travel ability' (U=576, p=0.018) were deemed significantly more important for Assistance dogs. 'Acuity of sense of smell' (U=280, p=0.004), 'Tendency to be distracted when working' (U=279.5 p=0.009) and 'Ability to solve problems when working' (U=310 p=0.009) were more important for Bio-detection dogs.

Ideal levels also differed between the roles, e.g., Bio-detection dogs were thought to require higher levels of 'Tendency to search by smell alone' (U=277.5 p=0.006) but lower levels of 'Attachment to human partner' (U=277.5 p <0.001) than Assistance dogs. When comparing ratings for the current dogs to the perceived ideal levels, there were significant deviations in several traits; i.e., dogs were generally rated as having higher than desired levels of 'General excitability' (mean deviation=1.02) and 'Tendency to be distracted when working' (mean deviation=1.00). Also, several traits significantly differed between tasks in the deviation from ideal levels, such as 'Tendency to vocalise' (U=230 p=0.003) and 'Tendency to be aggressive towards other dogs' (U=212 p=0.03), which were higher than ideal in Bio-detection dogs compared to Assistance dogs.

Paying particular attention to traits that are important and differ greatly from ideal levels when selecting MDDs may aid in improving task success. There are differences between MDD disciplines, and this needs to be considered during selection if subsequent performance is to be optimised.

2.2 Introduction

The demand for MDD as a technology to identify VOCs from different diseases is increasing (Jezierski et al., 2015). Research into detection dogs' behaviour has focused on other search tasks but not on MDD (1.2.2.1). More knowledge of how their behavioural attributes and demography affect their performance is needed to optimise selection, reduce time wastage and economic losses, and train the dogs best suited to the task (Rooney et al., 2004; Cobb et al., 2015). Hence, it is necessary to investigate which traits are relevant for the performance of these tasks. MDD behaviour may differ across disciplines as observed for other detection roles (1.2.1.6). Past studies reviewed differences across guarding, detection and service tasks (Bray et al., 2021b). Others investigated variations across detection roles, i.e. passive and active drug and explosives detection dogs (Rooney et al., 2004; Adamkiewicz et al., 2013). Some traits associated with good performance differed even with subtle task variations. Similarly, for MDD, Bio-detection and assistance tasks may require similar levels of some characteristics, although others may vary notably. Exploring these differences may aid in improving MDD profiling.

Questionnaires are shown to be a valid method to obtain valuable information about dog behaviour (Serpell and Hsu, 2001; Duffy et al., 2008). They sample the opinion of those that know the dog well, asking them to rate a series of items about the animal using established rating scales (1.2.4.2).

Questionnaires have effectively assessed behavioural problems in companion dogs (e.g. Serpell and Hsu, 2001) and dog breed differences (e.g. Bradshaw et al., 1996). Also, specific behavioural dimensions such as aggressiveness (Podberscek and Serpell, 1996; Duffy et al., 2008), impulsivity (Wright et al., 2012), frustration (McPeake et al., 2019), hyperactivity (Vas et al., 2007) and affective states (Sheppard and Mills, 2002). In working dogs, questionnaires have been used to investigate links between individual attributes and intrinsic factors such as dogs' sex and breed relevant to task performance(e.g. Rooney and Bradshaw, 2004; Adamkiewicz et al., 2013). They have been used in explosives and drug detection (Rooney and Bradshaw, 2004; Rooney et al., 2004; Adamkiewicz et al., 2013), livestock guarding (McConnell et al., 2022) and guide dogs (Hsu and Serpell, 2003; Duffy and Serpell, 2012). However, to date, there has been no similar research on MDD.

Therefore, this study consisted of a survey aimed to investigate the most important traits for MDD success, according to the opinion of experienced professionals in the field and to explore how these vary across MDD disciplines.

2.2.1 The survey method

The MDD survey methodology was based on Rooney et al. (2004), modified from Serpell (1996), which investigated which traits dog trainers and handlers considered the most relevant to the performance of explosives and drug search dogs and how these differed across tasks. Subsequent research modified this paradigm for different working dog populations (Rooney et al., 2007b; Adamkiewicz et al., 2013; Rooney and Clark, 2021; McConnell et al., 2022).

This method allowed us to systematically derive information on traits relevant to a specific role by investigating the ideal levels of different traits, how important these are for dog selection and to what extent the behaviour of actual dogs differs from what is ideal for the role.

The current study adapted this paradigm for MDD to investigate the relevance of 40 traits previously identified to be relevant for MDD from interviews with professionals in the field and the literature. Specifically, it investigated:

a. What level of each trait is ideal for MDD roles to determine required trait parameters?

b. Which are the most important traits for MDD selection? It is necessary to differentiate between the ideal level and its importance since a trait could ideally have very low levels, e.g. aggression but still have significant implications when selecting a dog.

c. What level of these traits is present in current dogs that the participants worked with, and how do these vary with their demography?

d. The discrepancy between trait levels in an actual medical detection dog and what is ideally expected. Establishing how different the dogs trained for this role are from what is required of them.

e. How each traits' importance, desired levels and discrepancy from the ideal, vary between bio-detection and assistance tasks?

f. Which are the most relevant traits to each discipline?

Although all traits assessed may somehow influence dogs' performance, it is not feasible to evaluate many items when assessing a dog (Rooney and Clark, 2021). Hence, I determined criteria to select traits more relevant for MDD tasks to use in the subsequent studies in this thesis.

An online questionnaire survey including 40 traits (Table 2.1) was designed to investigate these questions and distributed internationally amongst professionals working with MDD to collect a broad range of data on MDD behaviour.

This study also seeks to determine the main characteristics of MDD, to aid the design of subjective rating methods to evaluate MDD over their training and operational performance (Chapter 3), and the development of a test battery to investigate the variability of these traits in current MDD (Chapter 4).

2.3 Methods

2.3.1 Ethical statement

The study was approved the 13/11/2018 by the University of Bristol Faculty of Health Sciences Research Ethics Committee (Ref 76041).

2.3.2 The survey development

The Survey development process is summarised in figure 2.1.

Preliminary research to develop the MDD survey was carried out with the staff of Medical Detection Dogs® (Medical Detection Dogs, 2020).

Initially, I visited the charity on four occasions to familiarise myself with the organisation, staff and MDD tasks. I observed the Bio-detection and Assistance dogs' training, the induction process of potential clients being partnered with Assistance dogs, and a range of dog selection assessments.

In subsequent visits, I interviewed staff members from Medical Detection Dogs® to determine which traits to include in the survey and to derive the appropriate vocabulary to ensure the relevance and understanding of the questions for the prospective participants.

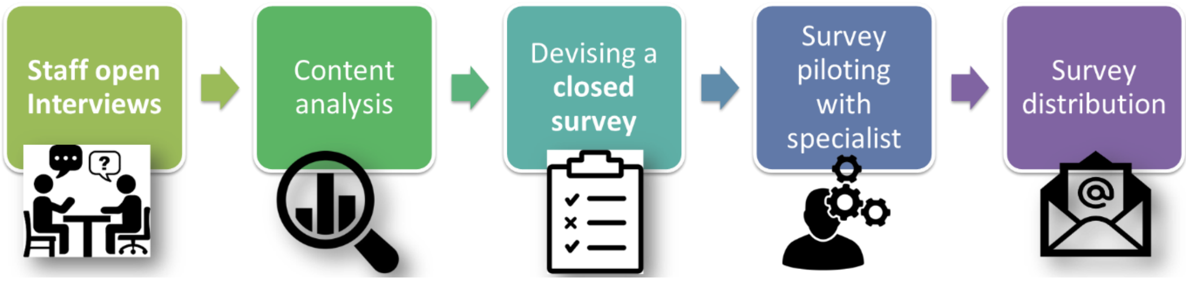


Figure 2. 1 *Flow diagram of the survey development timeline.*

2.3.2.1 *Staff interviews*

The open staff interviews occurred during three visits to Medical Detection Dogs® between May and June 2018. Participants were interviewed individually, each on a single occasion, in an isolated area of their office to avoid the influence of other individuals and distractors. The interviews lasted 29.2 min on average (min= 17, max= 39 min). The participants were twenty staff members with different roles and degrees of experience with MDD. Thirteen were female. Only four participants worked in the Bio-detection area and 16 in the assistance alert section. Staff roles included one behaviourist (assesses dogs with behaviour problems), four dog-client support instructors (supervises and guides client-dog partnerships), and ten dog trainers (handles and trains the dog at different training stages; their activities vary with their training stage and task). Three subjects had management positions. They had worked with MDD for 4.16 years on average (min=1 week, max=10 years). The interviews used an open questionnaire consisting of ten questions: These explored the participants' background with working dogs and MDD, which traits they considered were needed to make a dog best suited for MDD, a description of the dog they worked with, and which traits handlers deemed important to later use them for the survey development. Other questions investigated which traits they deemed undesirable for performance. In addition, they answered questions on dog breed and sex preferences to assess whether they considered demography aspects linked with MDD performance (Full interview questionnaire in Appendix 1).

2.3.2.2 Deriving the survey traits

I used the interview to derive a list of traits, all mentioned in multiple interviews, using content analysis. I clustered similar mentions into broader constructs and determined traits' nomenclature based on Rooney et al. (2004) and previous research on characteristics relevant to working tasks (1.2.8). This amounted to forty traits to be included in the survey (Table 2.1).

Table 2. 1 *Traits in the survey from the vocabulary that the participants (N=20) used in the interviews.*

Trait name as described in the survey	Examples of the participants' vocabulary in the interviews (extracted from cuotes)	Percentage of participants that mentioned the trait (%)
1. Acuity of sense of smell -sharpness of nose	Good nose, Good sense of smell	75
2. Tendency to search by smell alone	Scent hunting, Interest in odours, Using nose	80
3. Tendency to investigate humans by sniffing	Focus on human scent, Human body search	65
4. Tendency to explore areas by sniffing	Exploring physical area, Outdoor's search	25
5. Level of motivation when working	Motivation, Enthusiasm, Enjoys working, High drive	100
6. Willingness to try new behaviours even if they are wrong	Offering different behaviours/ trying again even if getting wrong	35
7. Persistence when alerting the presence of a target odour	Persistency, Robustness, Clear indication	80
8. Tendency to remain specific to the target odour - only respond to precise odour trained	Odour specific, Accuracy	50
9. Tendency to generalise	Generalising scents	25

alerts to similar scents		
10. Consistency of behaviour from day to day	Reliability, Consistency	60
11. Tendency to become frustrated when working	Frustration tolerance/control	40
12. Health - likelihood of having a long healthy working life	Healthy, Longevity	35
13. Stamina - endurance when working	Stamina, High energy/ activity	45
14. General excitability - tendency to become highly aroused	High excitability/ arousal, Hyperactivity	45
15. Ability to remain calm when not working	Calm, Settle, Steady	45
16. Confidence in different environments	Confidence, Boldness, Quick recovery after negative experiences.	100
17. Travel ability	Travel ability	15
18. Ease of adaptation to crate or kennel	Adaptation to crate, Settling when alone	25
19. Friendliness towards new people	Friendly/sociable/ enjoy been with people	85
20. Friendliness towards other dogs	Sociable with dogs	20
21. Level of attachment to human partner	Strong attachment/bond with human	45
22. Obedience to human commands	High obedience/ trainability, Willing to please	75
23. Ability to learn from being rewarded	Wanting to learn	25
24. Motivation to obtain food	Food motivation	70
25. Motivation to play with toys	Toy motivation	45
26. Willingness to bring an object back to a person	Good retrieval	15

27.	Motivation to retain possession of an object	Possessive, Lack of retrieval ability, Food stealing	35
28.	Independence -ability to work without constant guidance	Independent thinking/ Forward thinking, Learned disobedience	60
29.	Ability to solve problems when working	Problem solving	65
30.	Ability to concentrate during a trained task	High focus, Concentration	25
31.	Tendency to vocalise in public places	Excessive Barking/Vocalising	40
32.	Tendency to seek human attention	Excessive attention seeking	25
33.	Tendency to chase an object	Chasing	30
34.	Tendency to be distracted when working	Distractibility	45
35.	Fear of specific things (e.g. litter bags, brooms)	Fearfulness, nervousness, Environment sensitivity	100
36.	Reaction to sudden loud noises	Noise fear/ sensitivity/ phobias	30
37.	Body sensitivity - reactivity to touch and contact with objects	High Touch/ body sensitive	25
38.	Impulsivity - tendency to make hasty choices when working	Lack of impulse control/ impulsivity	40
39.	Tendency to be aggressive towards people	Aggression towards people	60
40.	Tendency to be aggressive towards other dogs	Aggression towards dogs	65

2.3.3 The MDD survey

An initial version of the survey was piloted by six dog experts, i.e. three researchers specialising in companion animals, two professionals experienced in MDD and a scientist specialist in dog behaviour. I made subtle changes to the wording based on their feedback.

The survey was produced in an online format on Google forms. From February 2019, the link to the survey was distributed via email to staff members of 56 different medical detection dogs' organisations, including charities, dog training institutions, universities and independent trainers. Also, it was advertised on dog science and training Facebook pages (Appendix 2). After two months, 62 participants (N=62) from 16 countries completed the survey.

The survey was divided into six sections. It was concluded from piloting that it could be completed in approximately 20 minutes. It commenced with a Participant Information Sheet (2.2.3.1). In the three following sections, the participants rated the same 40 traits derived from the interviews according to

a. The traits' ideal levels (2.2.3.2)

b. Their importance for selection (2.2.3.3)

c. Their levels in a dog they worked with (2.2.3.4).

Finally, they answered general questions about their professional background (copy of the survey in Appendix 3).

2.3.3.1 Participant information sheet

The first section contained a general overview of the study and explained that participation was anonymous and voluntary and that participants could withdraw. It also stated the inclusion criteria: being over 18, being a member of an organisation working with MDD (Bio-detection or Assistance) and having trained or worked closely (currently or in the past) with at least one dog. In addition, it had instructions on how to fill out the survey. Finally, the participants were asked to consent to the study.

2.3.3.2 Ideal Medical Detection Dog

The following section displayed the 40 dog traits potentially relevant for MDD tasks (derived from the informal interviews) related to searching ability, motivation, confidence and adaptation to different environments, sociability and human attachment, activity levels and reactivity (Table 2.1). The participant was asked to indicate what they considered to be the level of each trait in the ideal dog of their discipline using a scale of 1 to 5: 'As high as possible' (5), 'High' (4), 'Intermediate' (3), 'Low' (2), 'As low as possible' (1). An example was provided.

2.3.3.3 Importance of different traits

The same traits appeared in the subsequent section. However, here the participant was asked to rate how important they considered each trait for dog selection, independent of how they had rated it in the previous section. The scale ranged from 'Vitally important' (5), 'Important' (4) ', Intermediate' (3), 'Slightly important' (2) and 'Not at all important' (1). The instructions emphasised that the ideal trait levels in the previous section were not the same as their importance for selection. I provided an example of how these differ to promote a different scale used for each section.

2.3.3.4 Information about a medical detection dog that the participant was working with

The participant selected a dog they were working with or had recently. First, there were questions about the dog's task, if it was Bio-detection or Assistance, and the target scent it was trained to find. The participant provided the dog's details: their name, age, sex, reproductive status, origin, time of acquisition and health status. They were asked the duration they had been working with the dog, the dog’s trained indication behaviour, and the reward used to reinforce. Finally, they indicated their level of satisfaction with the dogs' performance; 'Very satisfied', 'Satisfied', 'Neutral', 'Not very satisfied', 'Not at all satisfied'.

2.3.3.5 Behavioural trait Rating for participants' dog

Next, the participant scored the same 40 traits in their dog by indicating the level of each they considered the dog to show: 'extremely high' (5), 'high' (4), intermediate (3), 'low' (2), and 'extremely low' (1).

2.3.3.6 Participants' demography and professional background

The last section enquired about the participant's details, including their gender, age, country, employment place and role. In addition, they were asked about their duration of working with MDD and other dog tasks, the number of dogs they had worked with, and the medical conditions they had trained dogs for. Finally, the participant indicated which breeds they had worked with and whether they had a preferred sex and breed for MDD tasks. They were asked to suggest three traits they considered the most important for MDD, and they could add characteristics not listed in the survey.

2.4 Data analysis

I analysed the data with IBM SPSS software® for the whole sample and then for Bio-detection and Assistance dogs separately.

I examined whether the dogs' trait levels varied with their sex using Mann-Whitney U tests. Regarding dog breeds, only Labrador Retrievers (46%) and German Shepherds (15.4%) included more than 10% of the dogs (Rooney and Bradshaw, 2004). Since the former exceeded the latter considerably, I did not compare trait levels between dog breeds.

I calculated a 'discrepancy score' (Rooney et al., 2004) for the whole sample, the mean difference between the dogs' ideal trait levels and their levels in the dogs they worked with. I squared and calculated the square root of the outcomes to get only positive numbers regardless of the direction of the difference.

Subsequently, I ranked the traits' mean scores according to each category, their ideal levels, importance and the dogs' discrepancy from ideal for all dogs and each discipline.

I performed Mann–Whitney U tests for traits' importance, ideal levels and discrepancy from ideal to investigate differences between disciplines.

2.4.1 The most relevant traits

From the results, I identified the most relevant MDD traits to apply in the subsequent studies within my project, based on two criteria (The traits included need to fulfil at least one):

1. The trait must appear in the top ten for importance for the whole sample or at least one task (Bio-detection or assistance).

2. The trait must be within the top ten highest discrepancies between their ideal level and that rated in an actual dog, but with a mean importance rating higher than 3 out of 5.

3. Traits that did not fulfil the above criteria but were of particular interest for this research due to their relevance from previous research in working dogs and potentially for MDD. (1.2.6).

2.5 Results

2.5.1 Survey participants

In total, there were 64 survey entries. However, two were incorrectly completed and hence excluded. Most participants came from the United Kingdom (32.3%), and the US (33.9%), followed by Canada (6,5%), Finland (4.8%), Germany (3.2%), Czech Republic (3.2%), Belgium (3.2%) and single participations from Italy, Poland, Spain, New Zealand, South Africa, Mexico, Sweden, The Netherlands and Australia.

Of the participants, 77.4% were female. There were 16.1% aged 21 to 29 years, 24.2% aged 30 to 39 years, 25.8% aged 40 to 49 years, and 17.7% age 50 to 59 years. 12.9% were 60 or older, and 3.2% preferred not to say their age.

More than half had experience with companion dogs (58.1%), and 37.1% with dog sports and show disciplines. The majority (93.5%) had worked with diverse dog tasks, including service dogs, e.g. guide dogs, detection and military or guarding duties. They worked with working dogs for an average of 16.8 years (SD=11.53, Min=0.75, Max=50) and with MDD for an average of 6.1 years (SD=5.45, Min=0.16, Max=35).

At the time of the survey, most performed multiple roles, defining themselves as dog trainers (79%), client-dog instructors (51.6%), dog handlers (22.6%), dog walkers (9.7%) and animal care assistants (4.8%). Others were lecturers or researchers (29.6%), dog behaviour specialists (35.5%), veterinary nurses (8.1%) and one veterinary surgeon (1.6%). Several participants had management positions (46.8%), and many carried out dog selection and procurement (35.5%), while some performed external communication functions, including client support (27.4%), public representation (25.8%) and fundraising (12.9%).

In total, 59.7% of the respondents worked mainly with Assistance and 30% with Bio-detection dogs, although some worked with both disciplines (9.7%). At the survey time-point, they were working with a mean of 5.8 MDD (SD=6.40, Min= 0, Max=30) and in the past with an average of 18 (SD=22.17, Min= 0, Max=100).

They had worked on MDD tasks with various dog breeds. Most with Labrador Retrievers (79%), mixed breeds (53.2%), Golden Retrievers (35.5%), German Shepherds (24.2%), Cocker Spaniels (24.2%), English Springer Spaniels (16.1%), Border Collies (14.5%), Poodles (11.3%), Belgian Malinois (9.7%) and Beagles (3.2%). There were single reports of less traditional breeds such as Shetland Sheepdog and Siberian Husky. The Labrador Retriever was the preferred breed for MDD tasks for almost half of respondents (40%), although a significant number did not have a preference (37.7%).

When participants were asked their preferred sex, the majority were indifferent (74.2%), while some selected neutered males (11.3%) and spayed females (9.7%), and a few preferred any male (3.2%) or female (1.6%).

2.5.2 Participants' institutions

The survey responses came from 44 dog training organisations and eight universities and ten were from private self-employed trainers. These had trained an average of 30.7 MDD dogs (SD=37.16, Min= 2, Max=250).

2.5.3 Dogs rated by the participants

Of the dogs, 51.6% were male, and 93.7% were neutered/spayed. Their ages ranged from one to twelve years old (mean 3.2). The majority were large-sized (71%), 13 dogs were medium-sized (21%), and the rest were small. Most dogs were pure breed (62.9%), including Labrador Retriever (46%), German Shepherd Dog (15.4%), Golden Retriever (5.1%), Beagle (7.7%), Cocker Spaniel (5.1%), Belgian Malinois (2.6%), English Springer Spaniel (2.6%), and four dogs each of different breeds. The remaining dogs were mixed breeds (24.2%) or first generation crosses between two breeds (12.9%).

The mean age of the dogs when obtained was 6.07 months (Min=1 month, Max=24 months, SD=5.53). Half of the dogs came from a breeder, 19.4% from a dog-rehoming shelter, 14.5% were reassigned from a previous working task they were not suitable for, 9.7% were personal donations, and one was the participant's dog. The origin of the remaining dogs was unknown.

Regarding the dogs' discipline, 39 were trained on assistance tasks (62.9%), while 23 were Bio-detection dogs (37.1%). In total, 69.2% of the Assistance dogs worked with a single condition, and the rest with more than one. Most Assistance dogs alerted glycaemia alterations from diabetes mellitus (66.7%). Others assisted with various health disorders. Bio-detection dogs identified samples of cancer, infectious and neurological conditions, and one was still in early training (Table 2.2).

The dogs performed different behaviours to indicate the presence of a scent, in most cases displaying more than one. The main behaviours were stand and stare (30.6%), sit (29%), scratch or paw (27.4%), nudge or nose poke (12.9%), lick the client (11.3%), jump-up (8.1%), vocalise (4.8%), spin (2.6%), lie down (1.6%), or place their head on the client's lap (1.6%). Some Assistance dogs retrieved an article to the client (6.5%), sought help (3.2%) or pressed an alert button (1.6%) when a health crisis occurred.

The participants had worked with the dog they rated for 23.5 months on average (SD=26.13, Min=2 months, Max=120 months). They trained them with different reinforcement methods, mostly with food (98.4%), but also with verbal praise (59.7%), stroking or petting (38.7%) and toy play (37.1%). Most respondents were very satisfied (67.7%) or satisfied (30%) with their dog's performance, only one was neutral, and none indicated being dissatisfied.

Table 2. 2 Medical *conditions that the rated dogs were trained to detect (N=62). Number and percentage of dogs trained to discriminate each condition (some trained for more than one scent)*

MDD tasks		Dogs trained	
MDD discipline	Condition/ target scent	N	%
Assistance	Hypoglycaemia	26	66.7
	Hyperglycaemia	16	41
	Epileptic seizures	6	15.4
	Addison's disease	1	2.6
	Allergies	4	10.3
	Narcolepsy	4	10.3
	Postural tachycardia syndrome (POTS)	4	10.3
	Pain episodes	1	2.6
	Migraine	1	2.6
	Change of blood pressure for Dysautonomia	1	2.6
	Anxiety or panic attacks	3	7.8
	Cancer (general)	15	65.2

Bio-detection	Lung cancer	1	4.3
	Ovarian cancer	4	17.4
	Colorectal cancer	1	4.3
	Infectious disease (general)	5	21.5
	Malaria	1	4.3
	Pseudomonas aeruginosa	1	4.3
	Parkinson's disease	3	13

2.5.4 *Ideal levels of each trait*

The MDD traits required at the highest levels were 'Health', 'Acuity of sense of smell', 'Persistence when alerting', 'Concentration' and 'Confidence'. Traits ideally expected at very low levels were 'Tendency to be aggressive towards dogs and people', 'Fear of specific things, 'Tendency to be distracted when working' and 'Reaction to sudden loud noises' (Table 2.3).

2.5.4.1 *Differences in traits' ideal levels between disciplines*

Sixteen traits were rated significantly differently between disciplines. Traits that the participants working with Bio-detection dogs rated ideally higher than those working with Assistance dogs included 'Acuity of sense of smell', 'Tendency to search by smell alone', 'Ability to concentrate', 'Independence', and 'Motivation to play with toys'.

Seven traits were ideally higher for assistance tasks than bio-detection: 'Confidence', 'Willingness to try new behaviours', 'Travel ability', 'Level of attachment to human partner', 'Obedience', 'Tendency to investigate humans by sniffing' and 'Willingness to bring an object back'. However, traits generally thought to be ideally low or very low were expected at lower levels for Assistance dogs than for Bio-detection dogs, including 'Tendency to chase an object', 'Tendency to vocalise', 'Fear of specific things' and 'Tendency to be aggressive towards other dogs'. (Table 2.3).

Table 2. 3 *Mean Ideal levels ratings for each trait for the overall sample (N=62) and each discipline, and significant task differences (*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).*

Trait	All dogs	Mean ± SD		Task differences	
		Bio-detection N=23	Assistance N=39	Mann-Whitney U	*p*
Very High					
1. Health	4.83 ±0.42	4.74 ±0.54	4.89 ±0.32	488.5	0.361
2. Acuity of sense of smell	4.77 ±0.58	4.7 ±0.56	4.63 ±0.60	330 *	0.029
3. Persistence when alerting	4.66 ±0.58	4.87 ±0.34	4.51 ±0.66	438.5	0.852
4. Ability to concentrate	4.64 ±0.55	4.83 ±0.39	4.51 ±0.61	327 *	0.032
5. Confidence	4.59 ±0.62	4.35 ±0.71	4.74 ±0.51	576.5 *	0.027
6. Ability to solve problems	4.59 ±0.53	4.74 ±0.45	4.49 ±0.56	344	0.075
7. Ability to learn from being rewarded	4.55 ±0.60	4.61 ±0.66	4.51 ±0.56	423.5	0.547
8. Level of motivation when working	4.52 ±0.66	4.57 ±0.66	4.49 ±0.66	414	0.562
High					
9. Consistency of behaviour	4.47 ±0.68	4.43 ±0.79	4.49 ±0.61	451.5	0.960
10. Tendency to remain specific to the target odour	4.38 ±0.64	4.57 ±0.59	4.26 ±0.66	337.5	0.073
11. Tendency to search by smell alone	4.34 ±0.74	4.7 ±0.56	4.11 ±0.76	277.5 **	0.006
12. Travel ability	4.33 ±0.81	4.00 ±1.02	4.54 ±0.56	557.0 *	0.033
13. Stamina	4.33 ±0.66	4.52 ±0.67	4.20 ±0.63	328.5	0.054
14. Independence	4.29 ±0.65	4.52 ±0.67	4.14 ±0.60	279.5 **	0.009
15. Level of attachment to human partner	4.19 ±1.03	3.52 ±1.24	4.63 ±0.55	677.0 ***	<0.001
16. Obedience to human commands	4.07 ±0.72	3.78 ±0.74	4.26 ±0.66	613.0 **	0.009
17. Friendliness towards new people	3.97 ±0.99	3.48 ±1.20	4.29 ±0.67	437.0	0.858
18. Willingness to try new behaviours	3.97 ±0.86	3.96 ±1.11	3.97 ±0.66	620.5**	0.008
19. Motivation to obtain food	3.97 ±0.86	4.09 ±0.79	3.89 ±0.9	407.0	0.521
20. Friendliness towards other dogs	3.86 ±0.94	3.48 ±1.16	4.11 ±0.68	496.0	0.466
21. Ability to remain calm	3.86 ±0.91	3.74 ±1.14	3.94 ±0.84	444.0	0.945
22. Tendency to investigate humans by sniffing	3.86 ±0.96	3.87 ±1.01	3.89 ±0.85	579.5 *	0.043
23. Ease of adaptation to crate or kennel	3.69 ±1.19	3.91 ±1.24	3.54 ±1.15	350.0	0.135
24. Willingness to bring an object back	3.67 ±1.16	3.04 ±1.36	4.09 ±0.78	640.5 **	0.004
Intermediate					
25. Motivation to play with toys	3.47 ±0.73	3.83 ±0.78	3.23 ±0.60	277.0 **	0.006
26. Tendency to explore areas by sniffing	3.40 ±1.15	3.70 ±1.11	3.20 ±1.16	352.5	0.144
27. Tendency to seek human attention	2.79 ±1.02	2.52 ±1.12	2.89 ±0.93	552.0	0.112
28. Tendency to generalise alerts to similar scents	2.74 ±1.41	2.61 ±1.23	2.91 ±1.52	497.0	0.469
Low					
29. General excitability	2.38 ±0.93	2.17 ±1.07	2.51 ±0.82	556.0	0.097
30. Tendency to chase an object	2.21 ±0.95	2.57 ±0.95	1.89 ±0.87	276.0 **	0.008
31. Object possessiveness	2.16 ±1.22	2.43 ±1.38	2.06 ±1.11	399.0	0.451
32. Frustration when working	1.64 ±0.83	1.52 ±0.59	1.71 ±0.96	504.5	0.372
33. Impulsivity	1.57 ±0.70	1.43 ±0.59	1.66 ±0.76	524.5	0.218

Very low					
34. Tendency to vocalise	1.38 ±0.62	1.7 ±0.76	1.17 ±0.38	281.0 **	0.003
35. Body sensitivity	1.34 ±0.55	1.43 ±0.59	1.29 ±0.52	387.5	0.269
36. Reaction to sudden loud noises	1.28 ±0.49	1.39 ±0.58	1.20 ±0.41	369.5	0.122
37. Tendency to be distracted when working	1.21 ±0.41	1.09 ±0.29	1.29 ±0.46	524.5	0.106
38. Fear of specific things	1.12 ±0.38	1.26 ±0.54	1.03 ±0.17	362.0 *	0.014
39. Tendency to be aggressive towards other dogs	1.10 ±0.36	1.22 ±0.52	1.03 ±0.17	381.5 *	0.039
40. Tendency to be aggressive towards people	1.09 ±0.34	1.17 ±0.49	1.03 ±0.17	401.0	0.104

2.5.5 Importance of considering each trait

The most important traits for MDD selection were considered to be 'Level of Motivation when working', 'Health', 'Concentration', 'Ability to learn from being rewarded' and 'Acuity of sense of smell' (Table 2.4).

When the participants indicated in their own words which traits they thought were the most relevant for MDD performance, they frequently mentioned 'Confidence' and 'Tendency to search by smell alone' (N=23), 'Motivation when working' (N=18) and 'Friendliness towards people' (N=16). Rare answers included 'Empathy', 'Work ethics' and 'Patience'. Some referred to factors that were not dog-centred but which they thought necessary, i.e. 'Suitability for the client's lifestyle' and 'Educating and supporting the client'.

2.5.5.1 Differences between disciplines of traits' levels of importance

The importance of ten traits differed significantly between tasks. 'Acuity of sense of smell', 'Tendency to search by smell alone', 'Ability to solve problems', 'Motivation to play with toys' and 'Tendency to be distracted when working' were considered more important when selecting Bio-detection dogs. 'Attachment to human partner', 'Confidence', 'Willingness to try new behaviours', 'Travel ability', and 'Friendliness towards new people' were more relevant for Assistance dogs (Table 2.4).

Table 2. 4 *Mean importance ratings for each trait for the overall sample (N=62) and each discipline and significant task differences (*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).*

	Mean ± SD and Ranked mean	Task differences

	Trait	All dogs	Bio-detection N=23		Assistance N=39		Mann-Whitney U	*P*
1.	Level of motivation when working	4.66 ±0.51	4.78 ±0.42	2	4.57 ±0.56	3	382.5	0.230
2.	Health	4.60 ±0.62	4.61 ±0.58	5	4.63 ±0.65	2	468.0	0.727
3.	Ability to learn from being rewarded	4.57 ±0.57	4.61 ±0.50	6	4.57 ±0.56	4	454.0	0.924
4.	Concentration	4.57 ±0.6	4.70 ±0.56	3	4.54 ±0.61	7	369.5	0.176
5.	Acuity of sense of smell	4.50 ±0.76	4.83 ±0.49	1	4.37 ±0.60	9	280.0**	0.004
6.	Persistence when alerting	4.47 ±0.68	4.35 ±0.83	11	4.54 ±0.56	5	498.0	0.411
7.	Confidence	4.46 ±0.76	4.09 ±0.87	15	4.69 ±0.58	1	604.0**	0.002
8.	Ability to solve problems	4.45 ±0.60	4.70 ±0.47	4	4.29 ±0.67	11	310.0*	0.022
9.	Consistency of behaviour	4.33 ±0.73	4.39 ±0.66	10	4.29 ±0.79	12	441.5	0.910
10.	Tendency to remain specific to the target odour	4.26 ±0.76	4.43 ±0.73	9	4.14 ±0.77	15	358.5	0.156
11.	Independence	4.19 ±0.78	4.52 ±0.67	7	4.06 ±0.80	18	292.0	0.013
12.	Tendency to search by smell alone	4.17 ±0.80	4.48 ±0.73	8	4 ±0.80	19	312.0*	0.032
13.	Willingness to try new behaviours	4.14 ±0.74	3.91 ±0.79	18	4.54 ±0.61	6	578.5 *	0.039
14.	Stamina	4.14 ±0.63	4.17 ±0.65	14	4.31 ±0.83	10	460.5	0.845
15.	Level of attachment to human partner	4.14 ±0.93	3.65 ±1.07	26	4.2 ±0.63	13	658.0***	<0.001
16.	Travel ability	4.09 ±1.02	3.73 ±1.12	24	4.4 ±0.88	8	576.0*	0.018
17.	Motivation to obtain food	4.09 ±0.80	4.26 ±0.69	12	4.06 ±0.87	17	393.0	0.387
18.	Tendency to be aggressive towards people	3.96 ±1.63	3.91 ±1.57	20	3.89 ±1.75	22	458.5	0.608
19.	Friendliness towards new people	3.93 ±1.01	3.65 ±0.98	27	4.17 ±0.98	14	580.0*	0.042
20.	Obedience to human commands	3.91 ±0.82	3.83 ±0.89	22	4.06 ±0.80	16	508.0	0.352
21.	Tendency to be aggressive towards other dogs	3.83 ±1.62	3.61 ±1.56	28	3.89 ±1.75	23	507.0	0.347
22.	Ability to remain calm	3.81 ±1.02	3.7 ±1.11	25	3.97 ±0.92	20	514.5	0.314
23.	Fear of specific things	3.81 ±1.46	3.96 ±1.26	16	3.6 ±1.63	27	425.0	0.716
24.	Friendliness towards other dogs	3.78 ±1.03	3.57 ±1.08	31	3.97 ±0.98	21	552.0	0.116
25.	Tendency to be distracted when working	3.78 ±1.30	4.26 ±1.01	13	3.4 ±1.44	33	279.5**	0.009

26. Reaction to sudden loud noises	3.74 ±1.45	3.91 ±1.28	19	3.6 ±1.54	28	414.5	0.601
27. Impulsivity	3.71 ±1.30	3.96 ±1.30	17	3.46 ±1.36	31	335.0	0.083
28. Tendency to investigate humans by sniffing	3.67 ±1.13	3.39 ±1.16	34	3.86 ±1.09	24	347.5	0.125
29. Body sensitivity	3.66 ±1.38	3.61 ±1.37	29	3.63 ±1.48	26	446.0	0.970
30. Motivation to play with toys	3.52 ±0.92	3.87 ±0.69	21	3.26 ±0.98	34	303.5*	0.024
31. Ease of adaptation to crate or kennel	3.47 ±1.13	3.52 ±0.90	33	3.77 ±1.11	25	475.5	0.681
32. Tendency to seek human attention	3.47 ±1.26	3.22 ±1.35	36	3.54 ±1.20	29	549.5	0.127
33. Frustration when working	3.45 ±1.39	3.61 ±1.44	30	3.23 ±1.48	36	377.5	0.282
34. Tendency to explore areas by sniffing	3.38 ±1.25	3.83 ±1.03	23	3.11 ±1.35	38	347.0	0.125
35. General excitability	3.34 ±1.26	3.35 ±1.23	35	3.49 ±1.29	30	487.5	0.558
36. Tendency to generalise alerts to similar scents	3.29 ±1.08	3.57 ±0.90	32	3.14 ±1.19	37	398.0	0.444
37. Tendency to vocalise	3.24 ±1.73	3.22 ±1.70	37	3.26 ±1.77	35	458.5	0.878
38. Willingness to bring an object back	3.12 ±1.13	2.87 ±1.29	38	3.40 ±0.95	32	562.0	0.087
39. Tendency to chase an object	2.90 ±1.24	2.87 ±1.18	39	2.91 ±1.29	39	445.5	0.964
40. Object possessiveness	2.67 ±1.47	2.83 ±1.50	40	2.77 ±1.48	40	418.0	0.646

2.5.6 Trait levels in current dogs

The trait levels in the dogs that the participants rated are listed in Table 2.5. Only 'Tendency to seek human attention' varied with the dogs' sex (U=353.5, p=0.031), being significantly higher in males (Mean=3.5, SD=0.82) than in females (Mean=3.0, ±0.92).

2.5.6.1 The discrepancy between dog trait ideal levels and actual dog ratings

The traits with the highest discrepancy scores were 'General excitability', 'Tendency to be distracted when working' and 'Object possessiveness', which were higher than ideal. 'Tendency to generalise alerts to similar scents' was lower than expected.

'Tendency to be aggressive towards people', 'Level of motivation when working', 'Consistency of behaviour', Ability to learn from being rewarded, and 'Persistence when alerting' differed little from ideal levels (Table 2.5).

2.5.6.2 Differences between tasks in discrepancy scores

The discrepancy scores of eight traits differed significantly between tasks. The divergences greater for Bio-detection were 'Willingness to try new behaviours', 'General excitability', 'Tendency to vocalise', 'Willingness to bring an object back', 'Tendency to be aggressive towards people', 'Tendency to be aggressive towards other dogs' (Too high) and 'Friendliness to other dogs' (too low). The discrepancy was biggest for Assistance dogs for 'Tendency to search by smell alone' (too low).

Table 2. 5 *Rated dog discrepancy (N=62) from ideal levels for each trait and by discipline (*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).*

Dog trait rating, discrepancy from ideal, ranking and direction (higher/lower than ideal)													
Trait	All dogs			Bio-detection N=23				Assistance N=39				Significant Task differences	
Discrepancy score highest to lowest	Dog rating	Disc.	Direction	Dog rating	Disc.	Rank	Direction	Dog rating	Disc.	Rank	Direction	Mann Whitney U	p
1. Tendency to chase an object	3.09	1.19	Higher	3.52	1.22	4	Higher	3.36	1.17	1	Higher	383.5	0.749
2. General excitability	3.12	1.02	Higher	3.17	1.43	1	Higher	3.16	0.74	14	Higher	251.0*	0.010
3. Tendency to be distracted when working	2.21	1.00	Higher	2.00	0.91	13	Higher	2.12	1.06	2	Higher	448.0	0.445
4. Tendency to generalise alerts to similar scents	2.84	0.98	Lower	2.87	0.87	17	Lower	3.17	1.06	3	Higher	463.0	0.307
5. Object possessiveness	2.75	0.98	Higher	3.23	1.00	10	Higher	2.76	0.97	4	Higher	352.5	0.367
6. Ease of adaptation to crate or kennel	3.97	0.95	Higher	3.87	0.91	14	Lower	3.80	0.97	5	Higher	429.5	0.646
7. Motivation to play with toys	3.96	0.91	Higher	4.23	0.91	15	Higher	2.24	0.91	6	Higher	423.0	0.772
8. Body sensitivity	2.12	0.91	Higher	2.35	1.00	11	Higher	4.21	0.86	8	Higher	359.0	0.461
9. Tendency to seek human attention	3.24	0.9	Higher	3.04	0.96	12	Higher	3.12	0.86	9	Higher	370.5	0.586
10. Tendency to vocalise	2.07	0.88	Higher	2.57	1.30	2	Higher	2.48	0.60	25	Higher	230.0**	0.003

11.	Willingness to bring an object back	3.65	0.86	Lower	3.36	1.22	5	Higher	2.16	0.63	22	Lower	266.0*	0.02
12.	Impulsivity	2.21	0.86	Higher	2.22	0.87	18	Higher	3.38	0.86	10	Higher	396.0	0.912
13.	Friendliness towards new people	4.09	0.86	Higher	3.61	1.04	9	Lower	3.72	0.74	15	Higher	320.0	0.160
14.	Tendency to investigate humans by sniffing	3.74	0.83	Lower	3.26	1.09	8	Lower	3.12	0.66	17	Higher	319.5	0.154
15.	Friendliness towards other dogs	3.69	0.83	Lower	2.96	1.30	3	Lower	3.24	0.51	33	Lower	230.0**	<0.001
16.	Willingness to try new behaviours	4.02	0.79	Higher	3.65	1.13	7	Higher	3.72	0.57	28	Higher	261.5*	0.015
17.	Reaction to sudden loud noises	1.90	0.76	Higher	2.13	0.91	16	Higher	2.04	0.66	18	Higher	340.0	0.281
18.	Confidence	4.02	0.72	Lower	3.74	0.61	29	Lower	3.80	0.8	12	Lower	471.0	0.236
19.	Tendency to search by smell alone	4.05	0.69	Lower	4.39	0.39	38	Lower	3.96	0.89	7	Lower	555.5**	0.008
20.	Tendency to remain specific to the target odour	4.17	0.69	Lower	4.35	0.48	34	Lower	4.16	0.83	11	Lower	512.0	0.057
21.	Health	4.24	0.69	Lower	4.13	0.78	22	Lower	4.24	0.63	23	Lower	401.5	0.986
22.	Ability to remain calm	3.78	0.69	Lower	3.91	0.83	20	Higher	4.28	0.60	26	Lower	326.5	0.181
23.	Stamina	4.16	0.67	Lower	4.22	0.74	24	Lower	2.21	0.63	24	Lower	375.0	0.631
24.	Motivation to obtain food	4.39	0.67	Higher	4.32	0.48	35	Higher	4.38	0.80	13	Higher	493.5	0.112
25.	Tendency to be aggressive towards other dogs	1.70	0.67	Higher	2.32	1.17	6	Higher	4.16	0.34	39	Higher	212.0*	0.028
26.	Tendency to explore areas by sniffing	3.74	0.66	Higher	4.04	0.52	32	Higher	1.84	0.74	16	Higher	484.0	0.149
27.	Frustration when working	2.26	0.66	Higher	2.17	0.83	21	Higher	3.83	0.54	30	Higher	332.0	0.216
28.	Travel ability	4.31	0.66	Higher	4.13	0.74	25	Higher	4.20	0.60	27	Lower	403.0	0.993
29.	Level of attachment to human partner	4.16	0.66	Lower	3.77	0.87	19	Higher	2.32	0.51	34	Lower	317.0	0.133
30.	Fear of specific things	1.74	0.66	Higher	1.91	0.65	28	Higher	3.96	0.66	19	Higher	403.0	0.993
31.	Concentration	4.30	0.60	Lower	4.50	0.52	33	Lower	4.46	0.66	20	Lower	497.5	0.088
32.	Ability to solve problems	4.23	0.59	Lower	4.23	0.70	26	Lower	4.21	0.51	35	Lower	418.0	0.780
33.	Obedience to human commands	4.11	0.57	Lower	3.95	0.61	30	Lower	4.08	0.54	31	Lower	418.5	0.772
34.	Independence	4.07	0.57	Lower	4.09	0.78	23	Lower	3.92	0.43	37	Lower	352.5	0.367
35.	Acuity of sense of smell	4.43	0.53	Lower	4.61	0.35	39	Lower	4.40	0.66	21	Lower	485.5	0.135
36.	Persistence when alerting	4.33	0.53	Lower	4.39	0.48	36	Lower	4.48	0.57	29	Lower	450.5	0.389

37. Ability to learn from being rewarded	4.61	0.50	Lower	4.64	0.70	27	Lower	4.63	0.37	38	Higher	341.0	0.256
38. Consistency of behaviour	4.26	0.48	Lower	4.30	0.48	37	Lower	4.32	0.49	36	Lower	443.0	0.456
39. Level of motivation when working	4.55	0.45	Higher	4.52	0.30	40	Lower	4.56	0.54	32	Higher	489.5	0.077
40. Tendency to be aggressive towards people	1.39	0.34	Higher	1.73	0.57	31	Higher	1.67	0.20	40	Higher	299.0*	0.028

2.5.7 *The most relevant traits*

From these findings, I identified 27 traits as the most relevant to assess in subsequent studies according to the criteria (2.3.1) regarding the traits' relative importance and discrepancy from ideal levels or their relevance for this research (Table 2.6).

Table 2. 6 *Behavioural attributes from the survey relevant to medical detection tasks according to inclusion criteria (at least for one task): 1= the ten most important traits; 2= the ten most discrepant traits with importance ≥3; 3= Not fulfil 1 and 2 but relevant to examine.*

Trait	Inclusion criteria	Trait	Inclusion criteria
1. Motivation when working	1	14. Obedience to human command	3
2. Tendency to search by smell alone	1,2	15. Impulsivity	2
3. Concentration	1	16. Independence	1
4. Confidence	1	17. Consistency of behaviour	1
5. Frustration during training	3	18. Ability to solve problems	1
6. Adaptation to crate or kennel	2	19. Ability to learn from being rewarded	1
7. Attachment to human partner	1,2	20. Willingness to try new behaviours	1,2
8. Acuity of sense of smell	1	21. Tendency to vocalise	2
9. Tendency to remain specific to the target odour	1	22. Motivation to play with toys	2
10. Persistence when alerting	1	23. Motivation to obtain food	3
11. Tendency to investigate humans by sniffing	1	24. Willingness to bring an object back to a person	2
12. Tendency to explore areas by sniffing	3	25. Tendency to seek human attention	2

13. Body sensitivity	2	26. Tendency to be distracted when working	2
		27. General excitability	2

2.6 Discussion

The current survey is the first study to address the relevance of MDD individual attributes for the performance of their roles. It identified 40 traits relevant for MDD tasks from preliminary interviews, which helped develop an appropriate vocabulary for the MDD survey participants, exploring the level of considered importance, ideal levels of each trait and extent to which characteristics differed from ideal levels that may need improvement. It identified significant differences between tasks. Using a set of criteria, 27 traits were derived for subsequent research.

2.6.1 Traits rated importance and ideal levels

The traits rated by the 62 participants as most important for MDD tasks were 'Level of motivation when working', 'Health', 'Ability to learn from being rewarded', 'Concentration', and 'Acuity of sense of smell'. These mostly coincided with the essential traits for explosives and drug detection dogs derived from studies with a similar methodology (Rooney et al., 2004; Adamkiewicz et al., 2013). Their relevance is evident since high task motivation, ability to focus, and good searching skills have been associated with increased dog success in detection tasks (e.g. Maejima et al., 2007; Rooney et al., 2007b; Rocznik et al., 2015). Good health is paramount for optimal dog performance as health disorders may be linked with behaviour alterations and prevent dogs from working (Walczak et al., 2012; Mills et al., 2020).

Some highly important traits, including 'Level of motivation when working', 'Ability to learn from being rewarded' and 'Acuity of sense of smell', also were scored very highly in the dogs the participants rated. Potentially, dog selection may target these traits, and current dogs may have these at very high levels. Conversely, the participants may be positively biased about their dogs and rated them favourably, similar to Clark et al. (2020b), where trainers of explosives detection dogs tended to be lenient when rating their dog's performance in a search task, compared with impartial raters when assessing the same dog.

There were traits that the participants ideally required at very high levels, while others at very low levels. For example, according to the participants, an ideal MDD would be healthy, with a highly developed sense of smell, persistent when alerting, highly focused on the task, and confident in different situations, but would not tend to be aggressive or fearful. However, some traits deviated noticeably from ideal in the current dogs, like 'General excitability' and 'Tendency to be distracted when working', which were rated as too high, and might be potentially disruptive (e.g. a dog that fails to alert the client due to a distractor). In previous questionnaires, similar undesirable traits in companion dogs and explosives and drug detection dogs, such as aggression, excitability and fear related traits, were differed significantly from participants' ideal levels (Serpell, 1996; Rooney et al., 2004). These traits have been repeatedly associated with working dogs' training failure (e.g. Serpell and Hsu, 2001; Batt et al., 2008).

Only 'Tendency to seek human attention' was significantly higher in males than females. Previous studies have seen that males tend to get more involved in dog-human play (Jones and Gosling, 2005), but are no different in attention-seeking. In contrast, in several studies, females displayed more social interactions with humans and engaged more frequently in gazing behaviour (e.g. Kubinyi et al., 2009b; Duranton et al., 2016; Kovacs et al., 2016).

Although all the 40 traits assessed here were valuable for MDD performance, I brought forward 27 identified as the most relevant using the criteria I posed, considering how important they are thought to be for selection and their deviation from ideal levels. These are studied further in the subsequent project phases.

2.6.2 Differences between disciplines

2.6.2.1 Traits rated significantly higher in Assistance dogs

'Confidence' and 'Travel ability' were significantly more important and ideally expected at high levels in Assistance dogs compared to Bio-detection dogs. These dogs aid their human partners in public places and must cope with different environments, whereas Bio-detection dogs work primarily in a room (1.2.1). Assistance dogs were ideally required to have higher 'Obedience to human commands' than Bio-detection dogs. These dogs must follow the cues of potentially vulnerable people and ignore distractors (1.2.6.5). Also, they were required to be more willing to retrieve objects, possibly since some dogs are trained to bring the client an emergency kit in a crisis. 'Friendliness towards new people' was considered more important when selecting Assistance dogs, probably because they are more frequently in contact with people than Bio-detection dogs. However, excessive levels may be task disruptive (1.2.6.2).

'Level of attachment to human partner' was expected to be higher and was more important for selecting Assistance dogs than for Bio-detection. Assistance dogs spend a great deal of time with the client, likely reading their body language and being aware of body odour changes to alert them to an upcoming crisis (1.2.6.2). In the initial interviews, professionals working with Assistance dogs stressed the importance of a strong bond between the dog and the client. The relevance of this relationship has been emphasised in previous research on different service tasks (Bray et al., 2021b). Cooperation and communicative skills have also been associated with successful performance in service dogs (e.g. Brown and Goldstein, 2011; MacLean and Hare, 2018).

'Willingness to try new behaviours' was highly important and ideally expected at higher levels for Assistance dogs compared to Bio-detection dogs. In the interviews and previous studies (Rooney et al., 2019), Assistance dogs professionals have linked it with higher alerting accuracy. Therefore, further investigating this trait may bring valuable insights into these tasks' performance.

2.6.2.2 Traits ratedsignificantly higher in Bio-detection dogs

'Acuity of sense of smell' and 'Tendency to search by smell alone' were believed to have higher importance and were required at higher levels for bio-detection tasks than for assistance. Bio-detection tasks sometimes require the identification of a scent at reduced concentrations (e.g. Concha et al., 2014). Therefore, a highly developed sense of olfaction and an innate tendency to use their nose over their vision are valuable traits for scent detection (Rooney et al., 2004). 'Tendency to search by smell alone' was lower than ideal and deviated more in Assistance dogs than in Bio-detection dogs. Perhaps this trait is more stressed during selection and easier to quantify objectively in Bio-detection dogs since scent concentrations are controlled in training trials.

'Ability to concentrate' and 'Independence' were ideally thought to be higher in Bio-detection dogs than in Assistance dogs. Participants expected a very high 'Ability to concentrate' for both tasks. However, Bio-detection dogs may require to concentrate intensely to learn scents faster and focus on a specific odour at small concentrations. These dogs are also expected to work mostly independently from the handler to avoid misguidance from body language. Similar to other detection tasks, such as avalanche search and rescue dogs (Diverio et al., 2017) and explosives detection (Lazarowski et al., 2019a), for which, the ability to make independent decisions when searching was linked with higher task success. 'Ability to solve problems when working' was considered more important for Bio-detection dogs than for Assistance, possibly as they must apply their problem-solving skills when making quick decisions to indicate the presence of the target scent.

'Motivation to play with toys' was also more important and required at higher levels for bio-detection tasks than assistance. Generally, detection dogs have been shown to be more energetic and playful when compared to service dogs, e.g. guide dogs, which tend to be calmer (Bray et al., 2021b). A high play motivation is useful for training and rewarding dogs for different detection tasks (e.g. Jones et al., 2004; Beebe et al., 2016). However, this MDD population were rated to have very high 'Motivation to obtain food', and they were reinforced more frequently with food rewards than toy playing. This may be partly why 'Motivation to play with toys' was not considered very important (mean importance=3.87).

Some traits, expected at very low levels, were significantly higher in Bio-detection dogs than in Assistance dogs, including 'Tendency to vocalise', 'Tendency to chase an object', 'Fear of specific things' and 'Aggressiveness towards other dogs'. Several traits had significantly higher discrepancy scores in Bio-detection dogs than in Assistance dogs. Bio-detection dogs rated by the participants tended to be more excitable, vocalise more and be more aggressive towards people and dogs than desired. These undesirable traits may be more tolerated in Bio-detection dogs since they are less exposed to public environments than Assistance dogs. Additionally, Bio-detection dogs may tend to be more active and excitable like in other detection roles (Jamieson et al., 2017) (1.2.6.9). The current selection of Bio-detection dogs may overlook the presence of undesirable traits to avoid losing dogs in which the organisation has invested resources. However, certain conditions, such as COVID-19 or malaria, may require Bio-detection dogs to work in public places such as hospitals or ports of entry. Therefore, high levels of negative traits may be more problematic.

2.6.3 Limitations and future steps

Although questionnaires may be subjective, this survey has greater clarity regarding the participants' perspectives because it rates the traits regarding their importance and ideal levels.

This practitioners' sample is diverse since the respondents come from several countries and differ in their professional experience. Although there is currently no data on the number of professionals working with MDD worldwide, the survey respondents may represent a significant proportion of those employed in this relatively recent area and provide an overview of the MDD general scenario. The participants' understanding of the surveys' constructs may have affected their ratings, even when vocabulary suitability was considered (Clark et al., 2020a; Clark and Rooney, 2021). Some may have confused the ideal levels with the importance of some traits for dogs' selection (Rooney et al., 2004). Their inferences and standards of how MDD dogs should behave might vary from their perspectives, institution criteria and professional and cultural backgrounds (Clark et al., 2020b; Clark and Rooney, 2021). Dog training organisations may standardise criteria for dog-task profiling, selection and training, and educate their staff on what they should focus when working with the dogs.(Clark and Rooney, 2021; Rooney and Clark, 2021).

In this study, most respondents were satisfied with their dog's performance. They also may have chosen a dog they thought was proficient. The participants could have been biased when rating their dogs. Dog trainers may show partiality regarding their dogs since they develop a strong bond with the animals and consider the dog's performance to reflect their professional work (Clark et al., 2020b).

The survey revealed traits that are more important for selection and their ideal levels. The traits in actual dogs far from desired levels may indicate deficiencies in the current selection and rearing practices and the need to refine them (Rooney et al., 2004). Early intervention to reduce or increase these traits when socialising, training and selecting these dogs may produce animals more suitable for the task with less economic resources and time invested (Cobb et al., 2015; Bray et al., 2021b).

The survey showed differences between disciplines in traits relevant to MDD in their expected levels, their importance and how different their ideal levels are from reality. Generating evidence of these variations for specific profiling is paramount for improving performance since basing dog selection and training on a general model of a working dog without considering differences across disciplines may contribute to a deficient operation with economic losses and dog welfare consequences.

Some of the traits' differences between tasks are similar to those previously seen in detection and service dogs (e.g. Bray et al., 2021b). However, most of these traits have not been evaluated in MDD. Therefore, there is a need to explore this further in current MDD populations. The following chapters rely on the 27 traits brought forward as the most relevant. Chapter 3 assesses these traits in a sample of MDD dogs rated by the trainers for performance during training and regular tasks. In addition, Chapters 4 and 5 explore the development of a test battery to investigate the level of those behavioural aspects in these dogs and their relation to task success.

2.7 Conclusion

Medical detection dogs are a valuable aid for diagnosing and assisting with different conditions worldwide. It is essential to explore the relationship between dogs' traits and task performance and profile the requirements for each MDD discipline to improve rearing and selection methods and reduce rejection.

This survey showed differences in traits' relevance between tasks and identified several traits whose levels were not as desirable. Increasing attention to these barriers enhances the likelihood of improved selection.

Chapter 3. Measurements of medical detection dog's ability:

Deriving meaningful measures of MDD performance

Abstract

This chapter aims to derive meaningful success measures for medical detection tasks and explore how trainers perceive different facets of dog performance in a population of medical detection dogs (MDD). The chapter is divided into two parts:

In Study A, firstly, I assessed training outcome frequencies and the reasons for dogs failing the programme. The main rejection reasons were: reduced searching ability, anxiety and poor progress. 59% of the dogs remained in the system, and 41% failed and were withdrawn.

Secondly, dog trainers rated their dogs on the traits derived from the Survey (2.4.7) and their overall ability (OA) at different time points. I analysed the distribution of their ratings, how consistent the ratings were for each task, which traits they prioritised when scoring OA and the traits discrepancy from ideal levels.

Nine traits showed wider variability than others, such as 'Frustration during training' and 'Ability to solve problems'. Assistance dogs had significantly higher ratings than Bio-detection dogs for OA and nine traits, e.g. 'Concentration' and 'Ability to learn from being rewarded'. Trained dogs had significantly higher ratings on OA scores and ten traits than trainee dogs. Only a few traits were rated consistently over time for each discipline. These were 'Impulsivity' (α= 0.76), 'Obedience' (α= 0.80), 'Willingness to bring an object back' (α= 0.80) and 'Motivation to play with toys' (α= 0.76) for Bio-detection dogs, and 'Confidence' (α= 0.80), and 'Adaptation to crate or kennel' (α= 0.89), 'Ability to learn from being rewarded' (α= 0.82) and 'Obedience' (α= 1.00) for Assistance dogs. OA was significantly linked with 'Ability to solve problems' (p <0.001), 'Willingness to try new behaviours' (p =0.002), and 'Tendency to investigate humans by sniffing' (p <0.001) ; and negatively associated with 'Body sensitivity' (p =0.028) and 'Adaptation to crate or kennel' (p =0.035). Several traits differed from ideal levels, such as 'Ability to solve problems' for Bio-detection dogs, 'Body sensitivity' for Assistance dogs and 'Tendency to get distracted' for both. I

calculated the Composite total ability score (CTAS) by integrating OA with the dogs' discrepancy from ideal as a composite measure involving both the dogs' ability and its deficiencies in their performance. The CTAS was significantly associated with training outcome ($p < 0.001$), suggesting that dogs considered to have better aptitude and show less deficiencies during training are more likely to succeed.

The trait ratings were reduced into five components to assess how these were associated with the dogs' performance in the test battery in Chapter 4.

Thirdly, I derived scent sensitivity and specificity scores from the charity's scent discrimination data during blinded trials. These analyses helped guide the choice of measures of success to test for factors predicting performance in Chapters 5 and 6, including training outcome, CTAS and scent sensitivity and specificity scores.

Study B examined how assessors rated dogs in a one-off in-house aptitude test. It investigated rating distribution, whether assessors agree when rating the dogs, their ability to distinguish traits from one another and if this test is predictive of training outcome.

Five traits were excluded due to low variability, e.g., 'Tendency to seek human attention' and 'Tendency to explore areas by sniffing'. Agreement between assessors was generally low, but two traits showed high agreement: 'Level of Motivation during training g' ($\alpha = 0.84$) and 'Tendency to search by smell alone' ($\alpha = 0.84$). Several traits were highly correlated for most assessors ($R \geq 0.70$), including 'Motivation during training' and 'Concentration'. Dogs' OA in the test was not linked with the training outcome.

Overall, the study produced measures of success for subsequent chapters. Certain traits seem more appropriate to focus on when assessing the dogs, such as those showing higher variability and those linked with ability. Staff differences in their perception of dogs' performance may be associated with lack of understanding, failure to distinguish traits or staff bias. Including clear concepts and staff training may improve future assessments.

3.1 Introduction

Working dog agencies often decide which dogs are best suited for the task based on their training ability. Whether dogs complete or fail the training programme may depend on the opinion of staff members or internal testing methods to examine performance (1.2.5). These assessments may be adequate for the organisations' needs, but is a simple pass or fail an absolute indicator of ability? Is the trainers' opinion reliable and consistent enough? Here, I explore how dogs are assessed by their trainers to understand how they consider different traits regarding their dogs' performance and to produce meaningful performance measures for subsequent project phases.

3.1.1 Training outcome

Several studies determine if their findings are associated with the dogs' suitability for a task using the institution's criteria. (e.g. Batt et al., 2008; Bray et al., 2017b). These are often based on the training agency's decision or the dog's certification from an official body (1.2.5). Research in guide dogs assessed behavioural measures from test batteries (e.g. Wilsson and Sundgren, 1997b; Batt et al., 2008; Tomkins et al., 2011), or trainers' numerical ratings (e.g. Goddard and Beilharz, 1984; Duffy and Serpell, 2012), as predictors of qualification. In a cognitive test, detection dog qualification was based on ratings from handlers of prospective search and rescue dogs (Hare et al., 2018). Maejima et al. (2007) tested whether the trainers' scores of dogs' performance in a detection trial predicted drug-detection certification. Dogs with a higher 'Desire to work' and less 'Distractibility' had a higher likelihood of completing training.

Training outcome is a useful parameter in research since subjects with less aptitude are filtered out during training. However, this decision is based on the opinion of expert trainers but can still be highly subjective and not reflect aspects of ultimate ability. To decide if a dog would achieve qualification, they regularly consider their overall performance in the training programme or their ability in internal testing trials, which offers a snapshot of their behaviour. However, a dog's ability is multi-dimensional (Rooney and Clark, 2021). A binary pass/fail training measure does not capture the complexity of the dogs’ individuality behind their performance. Therefore standardised subjective and objective parameters are needed to determine dogs’ task suitability better.

Although some traits may be relevant for many tasks, others differ in their requirements for the role (Bray et al., 2021b; Rooney and Clark, 2021). Performance should also be evaluated in operational dogs (1.2.5). Therefore, selection parameters should consider the task requirements and the dog's career stage.

3.1.2 Trainers' opinion of dogs' ability

Dog trainers often have significant expertise; they know the dogs' best and can provide comprehensive insight into the animals' behaviour (Rooney et al., 2004). Studies have been based wholly or partially on the dog trainer's view investigating aspects of the dog's performance through psychometric scales, i.e. assessing traits important for working tasks (e.g. Rooney et al., 2004; Adamkiewicz et al., 2013; Clark et al., 2020b). Trainers' questionnaires are relatively efficient for evaluating dogs since they are inexpensive, can be done remotely (Fadel et al., 2016), and have shown high predictive validity regarding training outcomes (e.g. Serpell and Hsu, 2001; McGarrity et al., 2016).

Rating dogs' OA provides a broader performance measurement than a binary outcome (1.2.5). Previous studies on explosives detection assessed trainers’ ratings of the dogs' OA on a scale from one to ten (Rooney et al., 2007b; Clark et al., 2020b). Research on dog livestock guarding integrated OA ratings with their variation from optimal levels into a subjective composite score, allowing trainers to systematically assess their performance by incorporating the dogs' general ability with their operational deficiencies (McConnell et al., 2022).

However, working dog agencies tend to consider the trainers' opinions from informal discussions or team consensus without a more systematic approach. Trainers may have different perceptions of the dogs' behaviour, and their judgment might be biased due to emotional bonds or the notion that the dogs' performance reflects their professional skills

(Clark et al., 2020b; Rooney and Clark, 2021). Also, the institution's objectives and culture may influence dog trainers' decisions. When the selection process relies on more than one person, frequently, there is no attempt to retrieve the opinions of staff members independently (Clark et al., 2020b).

3.1.3 *Scent sensitivity and specificity*

Objective measures are also necessary to determine the dogs' suitability for a task. Trained dogs' scent sensitivity (true positives / (true positives + false negatives) and specificity (true negatives / (true negatives + false positives) (1.2.1.2) are essential parameters of detection dogs' decision-making in searching trials. Research frequently utilises them when studying detection dog effectiveness (e.g. Rooney et al., 2007b; DeMatteo et al., 2018). These have shown great individual variation across studies (1.2.1.6). The olfactory threshold of each animal varies (Concha et al., 2019). Searching accuracy is impacted by training techniques, the origin and management of scent samples and adjacent odours, and environmental elements (Edwards et al., 2017; Lazarowski et al., 2020a). Also by internal factors such as the animal's olfactory capacity, anatomical characteristics, age, and emotional state (Jamieson et al., 2017; Lazarowski et al., 2020a)

3.1.4 *How does the Medical Detection Dogs® charity's training system work?*

Here I assessed how dog trainers from the charity Medical Detection Dogs® consider their dogs' behaviour during training and when operational regarding ultimate measures of task success. The charity trains bio-detection and alert Assistance dogs. Their training process is complex, and dogs are evaluated with methods that channel them to a particular task or filter them out if unsuitable as early as possible. However, the system is flexible and often adapted to the individual dog. Different staff and volunteers from the public are involved in the dogs' upbringing and training (Table 3.1)

During training, the dogs pass through several selection filters and can be rejected at any point. Based on their genetics and general behaviour, puppies are typically procured from different breeders around eight weeks of age and placed with socialiser volunteers. A socialising trainer works with the volunteers and dogs on general training and social skills. The dogs receive training from 6 months to 12 months of age. Training takes place in-situ and at the training centre, and they are evaluated multiple times on their general aptitude.

From an early age, it is established whether individuals may be best suited to training for bio-detection or assistance tasks based on the trainer’s opinion and internal testing. However, this decision can change through socialisation based on the dogs' perceived personality. At around 12 months old, dogs are tested for scent detection ability to decide if they will progress to specialised task training depending on their potential aptitude. Then, dogs destined for bio-detection training, and at 18 months, dogs intended for assistance work come into the training centre for dedicated training with a specialised trainer or instructor. This lasts approximately ten weeks but varies with each dog's progress. (During the pandemic lockdown, some dogs' training was suspended for up to 12 months). For Bio-detection dogs, the initial ten weeks of training involve discriminating scent samples of neutral targets like a Kong® toy or a tennis ball, progressively reducing the size of the scent source. After this period, suitable Bio-detection dogs continue into specific scent detection projects (1.2.1.3). However, if they don't perform as expected, they may later be changed to another, more appropriate project, used as demonstration dogs, or fully withdrawn from the programme.

Initial training for Assistance dogs involves working with an instructor with scent samples from their likely future partners to enable them to learn the scent of the medical condition with which they will work. Successful Assistance dogs are paired with a client, and the charity continues to work with them in the client's homes. Assistance dogs pass when they are reliably alerting to a specific percentage of episodes and have accomplished public access requirements. They are then certified. However, if the relationship with the client does not work, the dog may be involved in a different client-dog partnership, trained for bio-detection, or withdraw from the system (Figure 3.1).

Table 3. 1 *Description of staff and volunteer roles from the charity based on the charity's website (www.medicaldetectiondogs.org.uk), trainer rating data and conversations with trainers.*

Staff member	Main role
Trainer	Handle and train the dog at different training stages; their activities vary with their training stage and task.
Co-trainer	Train the dog jointly or to a similar extent to another trainer.
Trainer assistant	Supports the main trainer in their work or trains the dog to a lesser extent.
Socialising Trainer	Monitor the young trainee dogs placed with volunteer puppy socialisers and provide puppy training lessons. They work mainly outside the charity's training centre.
Instructor	Train the dogs during specialised centre-based task training, including early scent detection training in prospective Bio-detection dogs, scent alert training, and client pairing in Assistance dogs.
Project trainer	Train and manage trained Bio-detection dogs in their regular detection projects, reinforce their performance in detection trials and record data for research and support dog foster.
Test assessor	A member of staff that evaluates the dogs' performance on behavioural tests or scent discrimination trials.
Puppy socialiser	A public volunteer that hosts, socialises, and practices training with young trainee dogs during the socialising stage with a Socialising trainer's guidance.
Dog fosterer	Volunteers from the public that permanently or temporarily keep trained Bio-detection dogs in their homes following the charity's standards and bring them regularly to the charity's training centre to work in their scent detection projects.
Client	Patients with a health condition (i.e. diabetes, Addison's disease, PoTS, severe allergies and other illnesses) are paired with an assistance dog that alerts them of an impending health crisis.

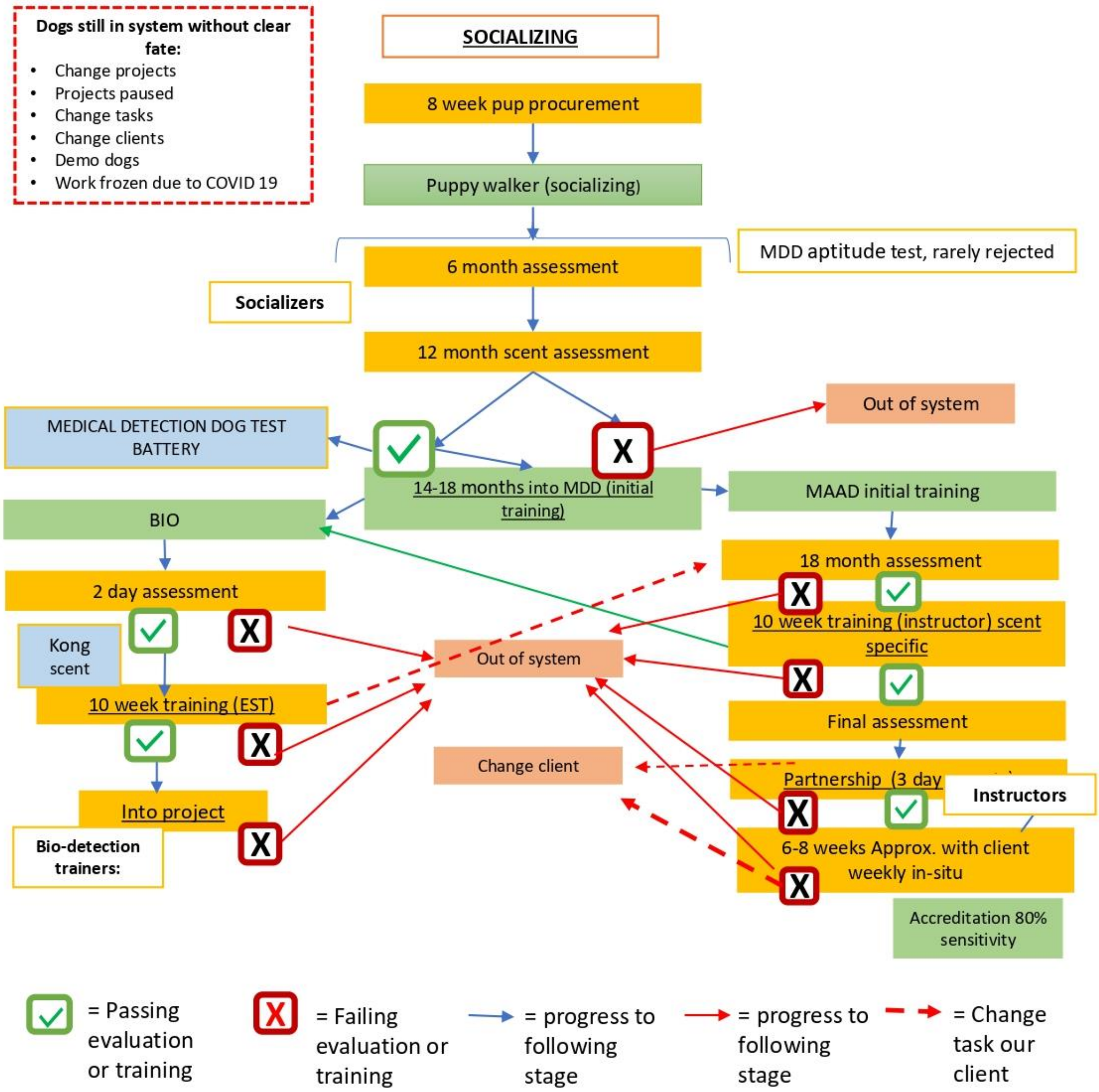


Figure 3. 1 Medical detection dogs® charity's training system. Dogs initially in socializing are canalised to different tasks and evaluated, where they may be filtered out from the system or changed to a different task.

3.1.5 Investigating the value of MDD success measures

In this chapter, I aimed to derive the most effective success measures to evaluate how different dog traits are associated with performance on medical detection tasks. I explored different parameters to derive meaningful measures of ability. It is critical to assess MDD performance when in training and when operational to understand how behavioural variation may be associated with their task functionality long-term as dogs' behaviour may vary with time (Chapter 1.2.5). Therefore I aimed to obtain specific measures for trainee and trained dogs.

The trainers' opinion on the dogs' performance can significantly impact trainee dogs' progress. I explored how trainers rate their perceived performance during training and working trials when evaluated internally. This enables the most relevant and reliable measures of success to be taken forward for the project's subsequent phases. The chapter is divided into two parts to facilitate navigation:

Study A. I focused on deriving performance measures by exploring the different aspects of how trainers rate the dogs' performance. I looked at:

a. The frequency of dogs' training outcomes (retain or withdrawn) and the main reasons for failing training to identify what is behind the outcome.

b. How the trainers used rating scales when scoring dog traits and their OA across training stages and disciplines, based on Rooney et al. (2007b).

c. How consistent were trainers rating their dogs at different training time-points for each discipline, and how this is associated with training outcome.

d. Which trait ratings are associated with OA.

e. The level of discrepancy from ideal levels from the medical detection dog survey (2.4.4) to derive a composite ability measure based on McConnell et al. (2022).

f. I investigated how the trait ratings associate with each other and cluster into components. Towards subsequently study how the trainers' opinion on dogs' behaviour in their training links with the dogs' performance in the test battery in Chapter 4.

g. Trained dogs' scent sensitivity and specificity levels and how these varied across scent projects and related to ratings of OA in their scent trials, aiming to derive objective parameters of scent performance.

Study B. The charity currently uses different assessment methods to select dogs that will ultimately work. I reviewed an in-house aptitude test to investigate whether the assessors could reliably rate dogs' performances with established rating scales when observing them on a single occasion. Specifically, I assessed the following:

a. How much variability do they show when rating each trait.

b. The strength of agreement between assessors when evaluating the dogs in a one-time task.

c. If assessors can differentiate each trait.

d. If the in-house test is predictive of the dogs' training outcome.

At the end of this chapter, I aimed to establish meaningful success measures for subsequent chapters and gain information on how the charity's trainers perceive their dogs' performance and how this may associate with dogs' successes and recommendations to enhance training and selection.

Study A: Determining medical detection dog's ability: value and reliability of performance measures

3.2 Methods

3.2.1 Ethical statement

The current research had ethical approval from the University of Bristol (Ref UB/19/05) and accorded to the charity Medical Detection Dogs standards. This covers this study and all subsequent studies in the dissertation.

3.2.2 Experimental Sample

The sample of dogs described here (N=58) is the same one used in subsequent chapters. It included all dogs working on bio-detection dog projects (N=19) at the test battery time-point and all prospective MDD over 12 months old that were in training from 09/09/2019 to 10/03/2020 (N=39).

Half of the dogs were female (N=29), 86.2% of males were neutered (N=25), and 82.8% of females were spayed (N=24). When undergoing the test battery (Chapters 4 and 5), the dog's mean age was 32.3 months (min=11, max=132, SD=26.38). Breeds included 31 Labrador Retrievers, 10 Labrador/Golden retriever cross, four Golden Retrievers, eight Cocker Spaniels, one Springer Spaniel, two Labradoodles, one Border Collie and one Hungarian Wire-haired Vizsla. Thirty-two dogs came from breeders procured as puppies, 16 from Guide Dogs of the Blind, seven from Dogs for Good, and three from dog rehoming charities.

Of the 39 trainee dogs at the test battery-point, 28 were still with their socialiser, and 11 were in initial bio-detection or alert assistance training (two dogs were already rejected at this stage). Twenty-two trainee dogs were intended for bio-detection tasks and 17 for assistance roles. Most dogs' disciplines were already known by then. However, five dogs (N=5) were rejected during the socialisation period before assigning them a role. These were assessed as being suitable for assistance tasks since most dogs are first considered for this discipline, and if not suited, they are trailed for bio-detection. The remaining sample was 19 trained dogs performing different bio-detection tasks. I had planned a larger dog sample, but due to the COVID-19 pandemic, all experimental work stopped in March 2020, thus limiting the number of dogs to those already tested (MDD sample in Appendix 4).

3.2.3 Trainers' rating forms

I designed a rating form to quantify the dogs' ability. Trainers were asked to fill this out on different occasions (specific to the research question as detailed below). The forms included the 27 most relevant traits (2.4.7) deduced from the survey based on the criteria established in Chapter 2 (2.3.1). I reviewed constructs from the charity internal rating sheets and substituted them with terminology standardised with the survey to make sure that the wording on the forms was understandable to the trainers and encompassed all the traits they regularly rate. For instance, 'Environmental behaviour' was replaced by 'Confidence in new environments' and 'Inappropriate high energy levels' with 'General excitability'. The sheets were piloted by three staff and modified based on their feedback (Appendix 5).

The trainers were asked to rate each trait based on their general impression of the dog's performance, using a rating scale from 1 to 6 for each attribute, where 1= Very low, 2=Low, 3=Intermediate, 4=High, 5=Very high and 6= Extremely high. In addition, the trainers rated the dog's OA from 1 (one of the worst dogs I have ever seen) to 10 (one of the best dogs I have ever seen). They also completed a withdrawn form for each dog rejected from training and briefly described the main reasons for rejection (Appendix 6). Assessment collection was ongoing from 13/05/2019 until 02/06/2021.

Initially, the dog trainers completed the paper forms collected at the charity. However, from March 2020, due to pandemic restrictions, I continued gathering the same data online with Jisc® surveys. We aimed for the trainers to complete the ratings at the end of the designated week. Still, on occasion, this was impossible, and trainers were asked to recollect ratings. The collection of those assessments completed retrospectively happened on average 89.7 days after the last designated data collection week (Min= 17, Max=496, SD=140.6). This was highly variable since the training of some dogs was interrupted due to the pandemic and resumed after conditions allowed. Some dogs transited through more than one programme (e.g., Bio-detection, Assistance or Covid training), and different trainers completed their assessments.

3.2.4 Binary outcome: Pass and Fail as a measure of success

To investigate what influences whether the dogs succeed in training or not, I classified from the withdrawn forms (3.2.3) whether each trainee dog remained in the system at

02/06/2021 data cut-point (14 weeks after the last dog progressed into initial training) and the main reasons for the dog being rejected.

Dogs that remained in the system did not necessarily complete training effectively since some dogs' training was suspended due to the pandemic, and others were between detection projects in the process of client pairing or changing clients at the data cut-point. However, due to the lack of standardisation, it was considered the fairest outcome measure for this study.

3.2.5 Dogs' training ability: How do dog trainers perceive their dogs' performance? And how are dog trait ratings associated with their OA?

3.2.5.1 Rating variability

Each trainer rated their dog using the form described in Appendix 5 (N=56; two trained dogs that trainers' did not rate were excluded). I considered the spread of the scale of trainers' ratings for each trait and their OA during working trials. The data was taken from the final assessment for each dog. Trainee dogs' ratings (N=39) were assessed separately for Bio-detection (N=22) and Assistance dogs (N=17) since the training methods differed with the task. I also evaluated how the project trainers of trained Bio-detection dogs (N=17) rated their dog's performance in their bio-detection projects at the test battery time-point.

3.2.5.2 Trainers' consistency when rating their dogs over training

I examined intra-rater consistency when testing a subject on repeated occasions (Taylor and Mills, 2006) for assistance and bio-detection tasks separately to determine if there was variation between them. The dogs were assessed at weeks 1, 4, 7 and 10 (approximately) of training.

However, the number of rating forms varied for each dog as not all the trainers completed the information when required. In several cases, the dogs were withdrawn before completing training, e.g. after week four or week seven, reducing the number of assessments for some time points, and 14 dogs were assessed retrospectively. Consequently, for analysis, I compressed the time points in three-time intervals for Bio-detection dogs and two for Assistance dogs, based on the data available for each task. I analysed only dogs rated on multiple occasions. Therefore, the sample for this analysis was 13 Bio-detection dogs and six Assistance dogs.

3.2.5.3 Associations between individual behavioural traits and ratings of OA

I assessed the association of trait ratings with OA with multiple regression analysis for all dogs and each discipline to identify trait differences across tasks. For Bio-detection dogs (n=39), I included trainee and trained Bio-detection dogs to improve the sample size.

3.2.5.4 Dogs' discrepancies from ideal levels

The discrepancy from ideal is a parameter developed by Rooney et al. (2004). I aimed to assess the deviation of each trait on the current dog sample from desirable levels from the Survey (2.4.4). Subsequently, I derived an overall discrepancy score (ODS) to measure the dogs' general deviation from the ideal levels of bio-detection and Assistance dogs from the survey.

3.2.5.5 Total ability composite score

I calculated the CTAS, which integrates the dogs' OA ratings with the dogs' ODS in a single combined measure (McConnell et al., 2022). The equation to calculate CTAS is shown in the data analysis section (3.3.2.7). Subsequently, I assessed its association with training outcomes.

3.2.5.6 Clustering of dog traits into components

To investigate how the trainers' opinions on different aspects of dogs' performance associate with the dogs' performance in the test battery (4.4.5), I clustered trait ratings into components to analyse how these components associated with the variables derived from testing in Chapter 4 (4.4.1)

3.2.6 Trained dog's sensitivity and specificity

I sourced sensitivity and specificity data from 27 dogs working on different detection projects at the charity. These were from the trained dogs' sample (N=19) at the test battery time-point (3.2.2). Additionally, eight dogs that were trainees at the test battery-point reached a stage of training by the end of data collection, where they participated in detection trials and generated trial data. Hence scent discrimination data for these dogs were gathered.

The charity provided the partially summarised data collected during the dogs' trials as sensitivity and specificity percentages from their databases, according to the equations presented earlier (3.1.3). The bio-detection trainers trained the dogs (N=27) to discriminate a target scent sample from control or blanks within an odour line-up on an octagonal carrousel over trial runs. The trainers rewarded the dogs when they correctly indicated the target scent presence (1.2.1.3). The trial duration, procedures and materials, and the samples' nature varied as the dogs were trained to work on a range of seven diseases. Twelve dogs were still in training, and their data came from unblinded (29.6% of the total data), single-blinded (3.7%) or double-blinded (11.1%) training sessions. Fifteen dogs had progressed to testing trials; hence, all their data was double-blinded (55.6%).

I selected the most appropriate data for each dog. Several dogs had data from multiple projects or trials of different complexity, e.g. from training and testing trials. I requested from the charity the data from the dogs' most recent trial, regardless of the disease they worked with, considered to provide more current information. I prioritised that from the most advanced testing session available; ideally, data from double-blind trials to standardise the trial difficulty. If the dog had data from both training and testing trials, the latter was selected. I investigated if sensitivity and specificity were associated with one another and if they were linked with the dog's CTAS.

3.3 Data analysis

3.3.1 Binary outcome: Pass and Fail as a measure of success

I conducted all the statistical analysis in IBM SPSS software®. I assessed dogs' training outcome frequencies and demography variations with descriptive statistics. To summarise the primary causes of rejection, I clustered those with similar terminology to the Medical detection dogs survey (2.2.2.2). I counted the number of times trainers referred to each of the main withdrawal reasons, given that dogs could fail for more than one cause. I performed Chi-Square tests to assess if there were significant differences in training success or failure according to dogs' sex or forthcoming task and Mann–Whitney U test to investigate age differences related to outcome.

3.3.2 How do the trainers perceive their dogs' performance, and how does this associate with OA?

3.3.2.1 Traits rating variability

Using descriptive statistics and plots, I assessed the rating distribution for all dogs and how these varied for each trait. I first evaluated the variation in ratings for each trait (N=56) for all dogs. I established a criterion where for traits showing wider variation, most ratings (excluding outliers) extended to at least three points of the scale, and at least 20% of the ratings differed from the Mode. And traits with ratings mainly ranging within one or two points were considered to have low variability and those which less than 20% of the ratings were different from the mode.

3.3.2.2 Differences in trainers' ratings across tasks

I explored if the trainers' traits ratings differed and whether their dogs were being trained for bio-detection or assistance tasks using the last ratings for each dog. I performed a Mann-Whitney U test to assess significant differences between each discipline's mean traits and OA ratings.

3.3.2.3 Differences in trainers' ratings across training stages

I conducted Mann–Whitney U tests to assess significant differences in trait ratings and OA between trainee and trained Bio-detection dogs since the trained dogs were also all bio-detection. All their trainers belonged to the same team and used standardised training methods.

3.3.2.4 Trainers' consistency when rating their dogs over training

I accounted the number of assessments for each discipline. I compressed them to three-time intervals for Bio-detection dogs and two for Assistance dogs since there were insufficient assessments for each time-point. I assessed frequencies of OA ratings and visually inspected individual dog progress with line graphs. For the 13 Bio-detection dogs, I initially performed a chi-square test of homogeneity between training progress tendencies (whether dogs' OA ratings 1. improved, 2. remained the same or 3. decreased) and training outcome. However, all expected cell count less than five, indicating that the sample size assumption was violated. Then I performed a Fisher's exact test instead, as suggested by Laerd Statistics (2017). Initially, Friedman tests were considered to assess differences in ratings across time intervals. However, since there was data of three time intervals for four dogs only, I conducted Wilcoxon signed-rank test using the last time intervals recorded for all dogs.

I performed a Krippendorf's Alpha analysis. This analysis measures inter-rater agreement and can also be used to evaluate intra-rater consistency over time. Initially, ICC was considered for this analysis for compatibility and interpretation against other papers in the field (Rooney and Clark, 2021). However, the number of dog assessments varied for each time interval. Krippendorf's Alpha shows advantages against ICC since it examines consistency with any number of time-points and sample sizes and allows for missing data (Hayes and Krippendorff, 2007), making it more appropriate for the current data set. Krippendorff's alpha values (α)=1 indicate perfect reliability, and 0= No reliability. For interpretation, $\alpha > 0.8$ means high reliability, and $\alpha < 0.8$ show low reliability, although this may vary with the context, and >0.70 may be acceptable when reliability requirements are not very strict. Hence, for this study, $\alpha > 0.70$ to ≤ 0.80 is considered to deliver acceptable reliability. $\alpha < 0.67$ shows very low reliability (Hayes and Krippendorff, 2007).

For each discipline, I performed these analyses for all time intervals and for each pair to identify time-points where OA ratings differed the most. I also carried out a Krippendorf's Alpha for each trait.

3.3.2.5 Associations between individual behavioural traits and ratings of OA

I investigated if trainers' trait ratings are associated with OA with a multiple regression model for the whole sample (N=56) and then for Bio-detection dogs (N=39) and trainee Assistance dogs (N=17) separately.

I conducted Spearman's Rank correlations to reduce the number of independent variables in the model. Excluding those traits highly associated with coefficients over 0.70, keeping the most biologically meaningful member of the pair, thus eliminating eight traits (Table 3.2). I introduced the variables to the model in repeated trials with stepwise methods and manually, considering traits' main effects significantly related to the dependent variable, increasing the overall significance of the model and the adjusted coefficient of determination (R2) (Laerd Statistics 2015). I performed a similar procedure for the Bio-detection dogs and Assistance dogs models.

Table 3. 2 *Traits highly correlated excluded for multiple regression analysis (N=56).*

	Correlated with	
Trait eliminated	traits included in multiple linear analysis	Other traits excluded
Tendency to search by smell alone	Acuity of sense of smell (R=0.712), Ability to solve problems (R=0.711).	Ability to concentrate (R=0.707)
Ability to concentrate	Ability to solve problems (R=0.711).	Tendency to search by smell alone (R=0.707), Tendency to remain specific to the target odour (R=0.763), Independence (R=0.736).
Tendency to remain specific to the target odour	Ability to solve problems (R=0.812).	Ability to concentrate (R=0.763).
Persistence when alerting	Ability to solve problems (R=0.744).	Independence (R=0.704).
Independence	Ability to solve problems (R=0.743)	Ability to concentrate (R=0.736), Persistence when alerting (R=0.704).

Ability to learn from being rewarded	Ability to solve problems (R=0.735)	Motivation to obtain food (R=0.724)
Motivation to obtain food	Ability to learn from being rewarded (R=0.724)	

3.3.2.6 Dogs' discrepancies from ideal trait levels

Based on (Rooney et al., 2007b), I calculated the dogs' absolute discrepancy from ideal for each trait by assessing the deviation between each trait's mean ideal score from the Survey for the relevant discipline. For each trait, I subtracted their ideal level from the dogs' actual trait levels. The resulting figures were squared and square-rooted to have only positive numbers (regardless of whether the trait was too high or too low). Subsequently, I added together the traits discrepancies to obtain an overall discrepancy score (ODS) for each dog. This was a general indicator of how far an individual is from ideal.

$$\textit{Discrepancy}=\sqrt{(Ideal\ \ dog\ score(survey)\text{-}\textit{individual dog score})^2}$$

3.3.2.7 Composite Total Ability score

To derive the CTAS, firstly, I assessed if the dogs' OA ratings were associated with their ODS Pearson correlations, which revealed a significant negative link ($R=-0.808$, $p=<0.001$). This justified combining both measures into a composite measure that indicates the dog’s ability but also its deficiencies in their performance based on McConnell et al. (2022).

I used the trainers' scores for OA to create a rank order. I ranked the trainee and trained dogs independently according to their OA from the highest ratings to the lowest with the SPSS IBM Software® Rank cases function, computing new ranked variables. Similarly, I ranked the overall ODS for each dog from the lowest to the highest. Then, the ranked scores were inverted (making the lowest rank the highest and vice versa) and added together so that the best-performing dog would have the highest CTAS. Therefore, a high-scoring dog would have a high ranking on OA and a low ranking on deviation away from ideal).

CTAS = $\begin{pmatrix} Overall\ discrepancy\ scores \\ ranked\ and\ inverted \end{pmatrix} + \begin{pmatrix} Trainer\text{'s}\ overall\ ability \\ ratings\ ranked\ \ and\ \ inverted \end{pmatrix}$

3.3.2.8 CTAS association with training outcome

I investigated if the CTAS varied dependent on whether they remained in the programme or failed training with a Mann-Whitney U test.

3.3.2.9 Clustering of dog traits into components

To investigate how different traits are linked with each other, I performed a PCA with Varimax rotation to condense these traits into components. I visually inspected the correlation matrix produced in the output and excluded seven traits highly correlated with others (R>0.7). The most biologically meaningful of the pair or the ones with more correlations were kept. I assessed Kaiser-Meyer-Olkin Measure (KMO) for overall sampling adequacy and each variable. KMO values may range from 0 to 1, although Kaiser (1974) considered measures > 0.5 the minimum acceptable for data reduction using PCA (KMO ≥ 0.9 as "Marvellous"; 0.8 ≤ KMO < 0.9 = as "Meritorious"; 0.7 ≤ KMO < 0.8 as "Middling"; 0.5 ≤ KMO < 0.6 as Miserable and KMO < 0.5 as "Unacceptable"). I removed three traits with individual KMO values under 0.5. The remaining variables produced a KMO of 0.746, and Bartlett's test of sphericity was significant (χ 2 (0.746) = 373.683, p< 0.001). I held components with eigenvalues over 1

3.3.3 Trained dogs' sensitivity and specificity scores

For each dog, I produced a single scent sensitivity score –the mean percentage of scent presentations that were correctly indicated (true positives) out of the total number of encounters with the target odour – and a specificity score – the number of control samples that were correctly ignored (true negatives) out of the total number of exposures to control samples– these across several search trials for the same dog. I assessed the frequencies of dogs working on each condition, the trial sort, the number of exposures and samples for each, and sample materials (when available). I calculated the mean sensitivity and specificity scores for all dogs and their ranges. I investigated whether these measures correlated between them and with CTAS with Pearson's correlations.

3.4 Results

3.4.1 Binary outcome: Pass and Fail as a measure of success

Of the trainee dogs (N=23), 59% remained in the system until the end of data collection, whilst 41% (N=16) were rejected. On average, the dogs that remained in the system were 19.3 months old at the test battery time-point (min=11, max=31, SD=6.28). Approximately half of them were males (52.2%). When data collection finished, 73.9% of the dogs had progressed to be operational, and 26.1% were still in specialised training (The training of these dogs was temporarily paused or slowed down due to lockdown, but they were still in the programme). 56.6% of the retained dogs were intended for bio-detection tasks (N=13) and 43.5% for assistance (N=10).

The rejected dogs had a mean age of 18.5 months (min=13, max=31, SD=5.23), and 10 were males (62.5%). Eleven dogs were still with their socialiser when rejected, while five were in specialised training. 56.3% of the dogs (N=9) were Bio-detection dogs, and 43.7% (N=7) were Assistance dogs. Three dogs were withdrawn from the system when they were already operational. There were no significant differences between dogs of different sexes, tasks or ages in whether they were retained or failed.

The trainers mentioned between one to eight rejection reasons for each dog being rejected. These were classified into 13 categories, and each was mentioned between one to 11 times (Table 3.3). The three most common rejection causes were related to deficient searching ability, followed by anxiety or stress-related and poor progress. Others were associated with a lack of motivation, adaptability, initiative, problem-solving skills, high aggressive reactivity, impulsivity, and poor frustration tolerance. Two dogs were rejected due to health issues besides behavioural concerns.

The proportion of dogs reported to be withdrawn for each reason generally appeared similar for bio-detection and assistance tasks. However, there were some differences. For instance, there were more Bio-detection dogs rejected for Lack of searching abilities and Anxiety/stress-related behaviours than Assistance dogs. Some Assistance dogs were withdrawn for Possessiveness/ scavenging behaviours, but no Bio-detection dogs failed for this reason (Table 3.3).

Table 3. 3 *Main withdrawn reasons mentioned by the trainers of rejected dogs (N=16) (often more than one for each dog).*

Withdrawn reason	Percentage of trainers that mentioned withdrawn reason (%)	percentage of trainers Bio-detection dogs (%)	percentage of trainers Assistance dogs (%)
Lack of search ability	68.7	77.9	57.1
Anxiety/stress-related behaviours	56.3	66.7	42.9
Lack of Motivation	18.8	22.2	14.3
Lack of adaptability/ Resilience	31.3	33.3	28.6
High frustration/ impulsivity levels	18.8	22.2	14.3
Low problem-solving skills	25.0	33.3	14.3
Low decision-making ability	12.5	22.2	0.0
Possessiveness/ scavenging	12.5	0.0	28.6
Lack of initiative	12.5	11.1	14.3
Aggressive reactivity	18.8	11.1	28.6
Poor progress	37.5	44.4	28.6
Health issues	12.5	11.1	14.3
Fostering issues	6.3	11.1	0.0

3.4.2 How do the trainers perceive the dogs' behaviour during training, and how does this associate with overall ability?

3.4.2.1 Trait rating variability

All traits showed acceptable variability, with >20% of dogs being rated differently to the mode. Nine traits extended wider within the scale, including 'Tendency to vocalise', 'Frustration during training', 'Adaptation to crate or kennel', 'Attachment to human partner', and 'Ability to solve problems'. The rest were mainly concentrated between Intermediate and High.

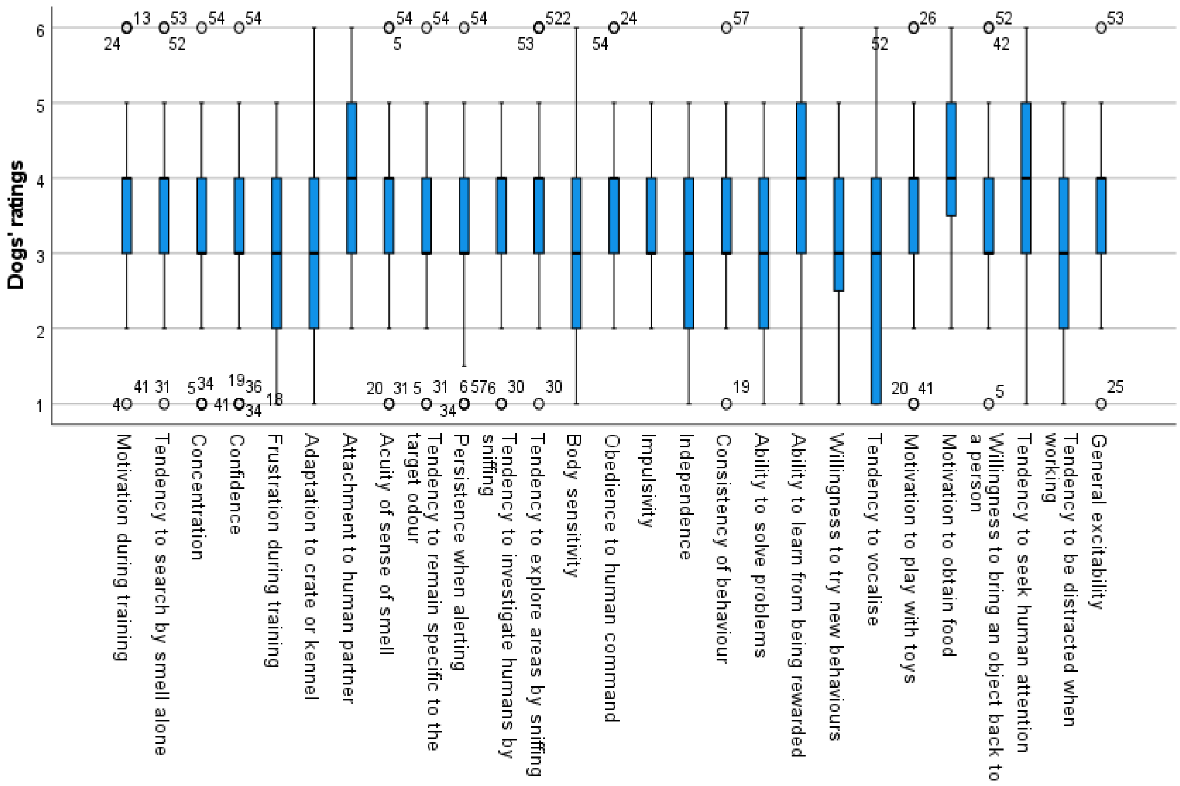


***Figure 3. 2** Distribution of trainers' ratings for different MDD traits over performance (N=56). Trait ratings: 1=Very low, 2=Low, 3=Intermediate, 4=High, 5=Very high, 6=Extremely high. Boxes show the median (bar within the box), the 25th interquartile (lower box border) and the 75th interquartile (upper box border). The whiskers indicate the minimum and maximum mean latencies. The circles and stars represent outliers.*

3.4.2.2 Differences in trainers' ratings across tasks

The trainers' ratings differed between dogs' disciplines for nine traits. For all traits except 'Tendency to Vocalise', Assistance dogs had higher ratings than Bio-detection dogs, including 'Tendency to search by smell alone', 'Ability to solve problems', 'Ability to learn from being rewarded', 'Ability to concentrate', 'Tendency to explore areas by sniffing', ''Confidence', 'Tendency to remain specific to the target odour' and 'Willingness to try new behaviours' (Table 3.4).

Table 3. 4 *Significant differences in trainers' dog ratings between prospective bio-detection (N=22) and assistance tasks (N=17). P=<0.05. *P <.05; **P <=.01; ***P <=.001*

Dog trait	Task Mean ± SD		Mann-Whitney U	P-value
	Bio-detection	Assistance		
Tendency to search by smell alone	2.98 ±1.12	4 ±1.06	271.5	0.015*
Ability to concentrate	2.59 ±1.04	3.65 ±0.79	281	0.007**
Confidence	2.73 ±1.03	3.59 ±1.12	269	0.020*
Tendency to remain specific to the target odour	2.93 ±0.98	3.65 ±0.79	259.5	0.039*
Tendency to explore areas by sniffing	3.20 ±1.15	4.47 ±1.12	294	0.002**
Ability to solve problems	2.43 ±0.93	3.59 ±0.71	307.5	<0.001***
Ability to learn from being rewarded	2.91 ±0.97	4.35 ±0.86	320	<0.001***
Willingness to try new behaviours	2.64 ±1	3.53 ±1.01	270	0.018*
Tendency to vocalise	3.14 ±1.49	2 ±1.17	97	0.016*

3.4.2.3 Differences in trainers' ratings across training stages

Ten traits were rated significantly higher in trained dogs than in trainee Bio-detection dogs: 'Tendency to search by smell alone', 'Concentration', 'Confidence', 'Ability to solve problems', 'Persistence when alerting', 'Independence', 'Willingness to try new behaviours', 'Motivation to obtain food' and 'Ability to learn from being rewarded'. Only 'Tendency to be distracted when working' was higher in trainee dogs (Table 3.5). Ratings on OA were significantly higher in trained dogs (Mdn =7) than in trainee dogs (Mdn=3.75: U=296.5, p= 0.001).

***Table 3. 5** Significant differences in trainers' dog ratings between bio-detection trainee (N=22) and trained dogs (N=17). P=<0.05. *P <.05; **P <=.01; ***P <=.001.*

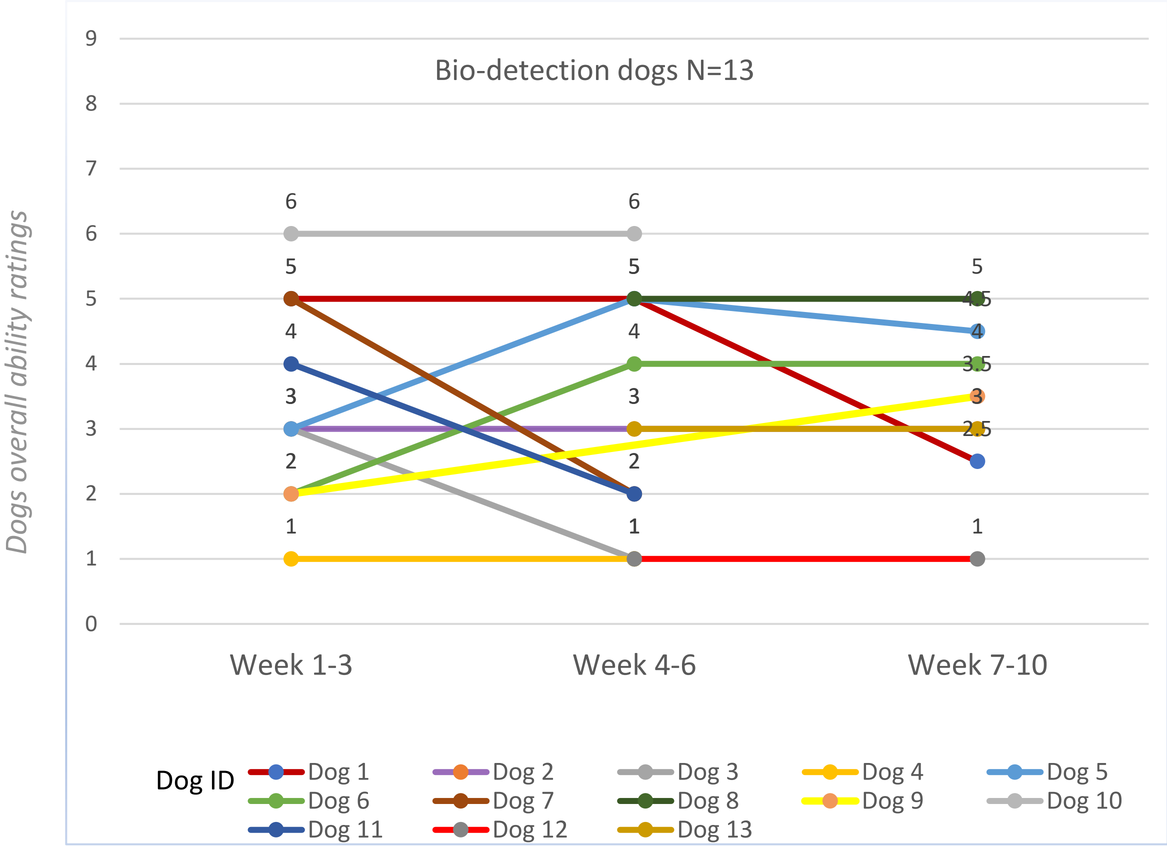
Dog trait	Training stage Mean ± SD		Mann-Whitney U	P-value
	Trainee	Trained		
Tendency to search by smell alone	2.98 ±1.12	3.76 ±1.09	257.5	0.038*
Ability to concentrate	2.59 ±1.04	3.76 ±1.03	284.5	0.004**
Confidence	2.73 ±1.03	3.69 ±1.25	253	0.022*
Persistence when alerting	2.82 ±1.06	3.82 ±0.95	276	0.011*
Independence	2.48 ±1.01	3.59 ±1	288	0.008**
Ability to solve problems	2.43 ±0.93	3.53 ±1.12	289	0.003**
Ability to learn from being rewarded	2.91 ±0.97	4.24 ±1.03	306	<0.001***
Willingness to try new behaviours	2.64 ±1	3.41 ±1.08	260	0.033*
Motivation to obtain food	3.84 ±0.76	4.65 ±1.1	261	0.036*
Tendency to be distracted when working	3.43 ±1.07	2.47 ±1.01	98.5	0.008**

3.4.2.4 Trainers' consistency when rating their dogs over training

For Bio-detection, there were ten dogs rated at Week 1-3 (OA Mdn=3, Min=1, Max=6, SD= 1.57), 12 dogs at Week 4-6 (OA Mdn=3, Min=1, Max=6, SD= 1.80) and eight dogs at Week 7-10 (Mdn=3.25, OA Min=1, Max=5, SD= 1.25). For assistance, six dogs for Week 1-4 (OA Mdn=7.5, Min=4, Max=8.5, SD= 1.62) and for week 7-10 (OA Mdn=6.5, Min=5, Max=8, SD= 1.21). Of the 13 Bio-detection dogs, OA increased over training for two dogs, six dogs remained the same, and four dogs decreased (Figure 3.3). Eight dogs were successful (61.5%). Three improved their OA, three maintained the same ratings, and two decreased their performance. However, five dogs were rejected without completing training (38.5%), and their ratings diminished (N=2) or remained equal over time (N=3), although with poor scores.

For Assistance dogs (N=6), one dog improved their OA ratings across training, two remained steady, and three showed a decrease. Of the dogs that failed (33.3%), one increased OA, and another remained stable. The rest remained in the programme (66.7%) (Figure 3.3).

Neither Fisher's exact test nor Wilcoxon signed-rank test showed significant variation over training OA tendencies or progress for either group.



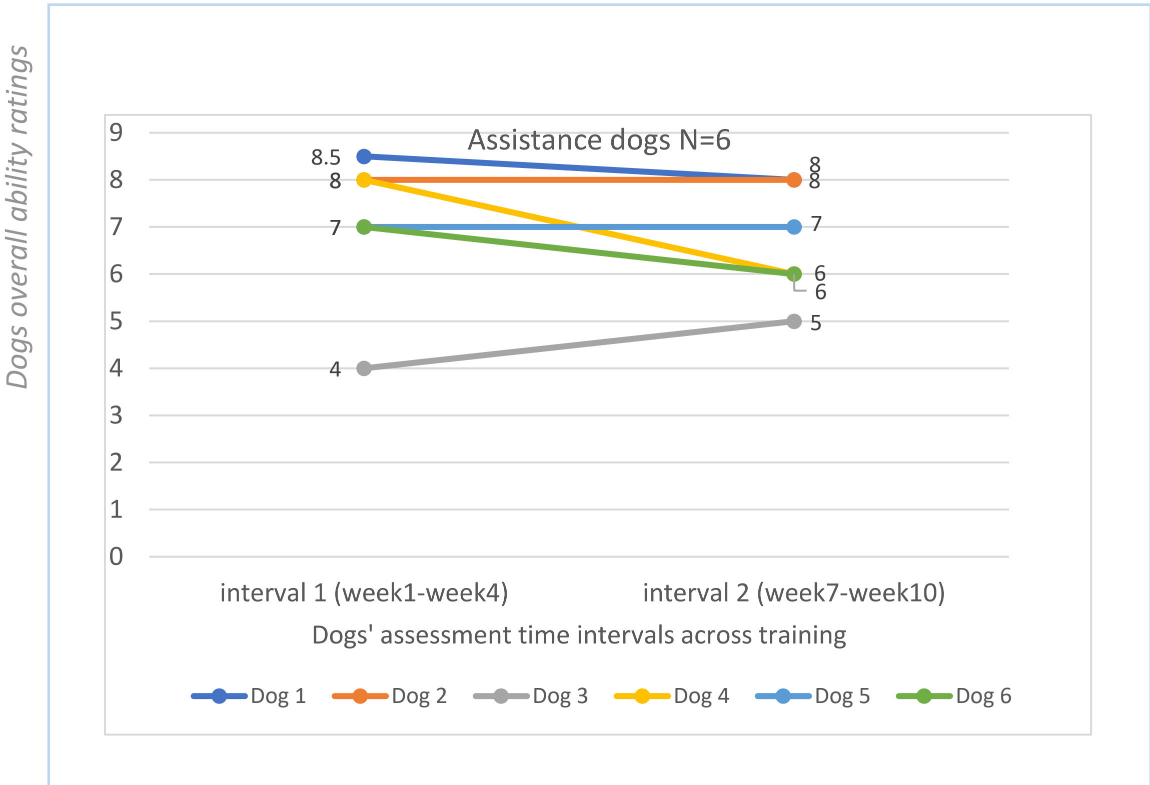


Figure 3. 3 *Dogs' training progress over time. Three time intervals for Bio-detection dogs'(N=13) and two for Assistance dogs (N=6). Cases with multiple assessments.*

For Bio-detection dogs, Krippendorf's Alpha analysis showed very low consistency of OA over time (α= 0.5). However, consistency slightly improved between weeks 4-6 and 7-10 (α = 0.58) than weeks 1-3 and 4-6 (α= 0.49). OA showed a relatively higher consistency for Assistance dogs than for Bio-detection dogs between time intervals (α = 0.67).

Four traits were relatively consistent for Bio-detection dogs: 'Impulsivity' and 'Obedience', 'Willingness to bring an object back to a person', and 'Motivation to play with toys'. Consistency was low for 'Tendency to investigate humans by sniffing' and very low for the remaining traits. For Assistance dogs, four traits had high consistency: 'Obedience', 'Adaptation to crate or kennel', 'Ability to learn from being rewarded', and 'Tendency to vocalise'. 'General excitability' and 'Confidence' had acceptable consistency, and the rest very low (Table 3.6).

Table 3.6 *Consistency of trainers' ratings for each trait and OA over training for Bio-detection dogs (N=13; three assessment time intervals) and Assistance dogs N=6 (two assessment time intervals).*

Krippendorf's Alpha analysis		
Trait	Bio-detection dogs	Assistance dogs
Motivation during training	0.31	0.52
Tendency to search by smell alone	0.38	0.63
Concentration	0.49	0.28
Confidence	0.48	0.80
Frustration during training	0.32	-0.22
Adaptation to crate or kennel	0.22	0.89
Attachment to human partner	0.17	0.57
Acuity of sense of smell	-0.12	0.37
Tendency to remain specific to the target odour	0.26	0.45
Persistence when alerting	0.45	0.19
Tendency to investigate humans by sniffing	0.70	-0.35
Tendency to explore areas by sniffing	0.01	0.42
Body sensitivity	0.37	-0.49
Obedience	0.80	1.00
Impulsivity	0.76	0.62
Independence	0.40	0.30
Consistency of Behaviour	0.25	0.18
Ability to solve problems	0.11	0.36
Ability to learn from being rewarded	0.49	0.82

Willingness to try new behaviours	0.22	-0.29
Tendency to vocalise	0.50	1
Motivation to play with toys	0.76	-0.24
Motivation to obtain food	0.49	0.51
Willingness to bring an object back	0.80	-0.19
Tendency to seek human attention	0.65	0.53
Tendency to be distracted	0.01	0.53
General excitability	0.50	0.77
OA	0.48	0.66

High reliability (>0.8) Acceptable reliability (>0.7- ≤0.8) Low reliability (≥0.67 ≤0.7)

Very low reliability (<0.67)

3.4.2.5 Associations between individual behavioural traits and ratings of OA

The model with the trainers' trait ratings as independent variables and their ratings of OA as the dependent variable was statistically significant ($F_{(5, 50)} = 50.26$, $p < 0.001$, adj. R2 = 81.7%). Higher ratings on 'Ability to solve problems', 'Willingness to try new behaviours', and 'Tendency to investigate humans by sniffing'; and lower ratings on 'Body sensitivity' and 'Adaptation to crate or kennel' were significantly associated with higher OA (Table 3.7).

Table 3. 7 *Multiple regression results with trainers' trait ratings as dependent variables and OA as independent variables (N=56).*

Independent variables	B	95.0% CI for B		SE B	*P*	β	R2	ΔR^2
		LL	UL					
Model							0.83	0.81***
Constant	-1.01	-2.47	0.46	0.73	0.175			
Ability to solve problems	1.36	1.02	1.70	0.17	<0.001	0.62		
Tendency to investigate humans by sniffing	0.57	0.27	0.88	0.15	<0.001	0.24		

Willingness to try new behaviours	0.54	0.22	0.87	0.16	0.002	0.25		
Body sensitivity	-0.21	-0.39	-0.02	0.09	0.028	-0.13		
Adaptation to crate or kennel	-0.26	-0.49	-0.02	0.12	0.035	-0.13		

Note. B= unstandardised regression coefficient; CI=confidence interval; LL= lower limit; UL=upper limit; R2=Coefficient of determination; SE B= standard error of the coefficient; β=Standardised coefficient; R^2= Coefficient of determination; ΔR^2= Adjusted R2, ***P <=0.001

The model for Bio-detection dogs (N=39) was statistically significant ($F_{(3, 35)} = 64.43$, $p < 0.001$, adj. R2 = 83.4%). Ratings on 'Ability to solve problems, 'Tendency to investigate humans by sniffing', and 'Willingness to try new behaviours' were significantly associated with OA (Table 3.8).

Table 3. 8 *Multiple regression model for Bio-detection dogs (N=39) with trainers' trait ratings as independent variables and OA as dependent variable (abbreviations meaning in Table 3.7) ***P <=0.001.*

Independent variables	B	95.0% CI for B		SE B	*P*	β	R2	ΔR^2
		LL	UL					
Model							0.85	0.83***
Constant	-3.23	-4.63	-1.82	0.69	<0.001			
Ability to solve problems	1.36	0.99	1.74	0.18	<0.001	0.63		
Tendency to investigate humans by sniffing	0.77	0.43	1.12	0.17	<0.001	0.31		
Willingness to try new behaviours	0.58	0.19	0.96	0.19	0.005	0.25		

For Assistance dogs (N=17), the model was statistically significant ($F_{(2, 14)} = 22.74$, $p < 0.001$, adj. R2 = 73.1%). Higher ratings on 'Willingness to bring an object back to a person' and 'Tendency to investigate humans by sniffing' were significantly associated with OA (Table 3.9).

Table 3. 9 *Multiple regression model with assistance dog trainers' (N=17) trait ratings as fixed factors and overall ability as outcome (abbreviations meaning in Table 3.7) ***P <=0.001.*

Independent variables	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.76	0.73***
Constant	-1.01	-3.45	1.43	1.14	0.39			
Willingness to bring an object back to a person	1.19	0.67	1.71	0.24	<0.001	0.66		
Tendency to investigate humans by sniffing	0.83	0.27	1.39	0.26	0.007	0.43		

3.4.2.6 Dog trait discrepancies from ideal levels

The discrepancy for each trait varied between disciplines. The traits with the highest ODS for Bio-detection dogs were 'Ability to solve problems' (Too low), 'Tendency to be distracted when working' (too high), 'Concentration' (Too low), 'Independence' (Too low), and 'Body sensitivity' (Too high). For Assistance dogs, the most discrepant traits were: 'Body sensitivity' (too high), 'Tendency to seek human attention' (too high), 'Tendency to be distracted when working' (too high), 'Impulsivity' (too high), and 'General excitability' (too high) (Table 3.10).

Some traits were near to their ideal levels, such as ’Motivation to obtain food’, ‘Motivation to play with toys’ and ‘Ability to learn from being rewarded’ for Assistance dogs, and 'Willingness to bring an object back to a person' and 'Motivation to and 'Obedience to human command' for Bio-detection dogs.

Table 3. 10 *Ten traits with the highest discrepancies from ideal for Bio-detection (N=39) and Assistance (N=17) tasks (listed in descending order).*

	Bio-detection	Discrepancy	Assistance	Discrepancy
1	Ability to solve problems	↓ 2.26	Body sensitivity	↑ 2.02
2	Tendency to be distracted when working	↑ 2.2	Tendency to seek human attention	↑ 1.95
3	Concentration	↓ 2.06	Tendency to be distracted when working	↑ 1.92
4	Independence	↓ 2.02	Impulsivity	↑ 1.91
5	Body sensitivity	↑ 1.85	General excitability	↑ 1.61
6	Tendency to search by smell alone	↓ 1.78	Adaptation to crate or kennel	↓ 1.44
7	Persistence when alerting	↓ 1.77	Persistence when alerting	↓ 1.32
8	Acuity of sense of smell	↓ 1.73	Frustration during training	↑ 1.32
9	Tendency to vocalise	↑ 1.72	Independence	↓ 1.31
10	Impulsivity	↑ 1.69	Concentration	↓ 1.19

Note. The arrow indicates the direction of the deviation. ↓ = Lower than ideal ↑ = Higher than ideal

3.4.2.7 CTAS association with training outcome

Dogs that remained in the system tended to have significantly higher CTAS over training (median=59.5) than those rejected (median=29.7: U=306, p <0.001; Figure 3.4).

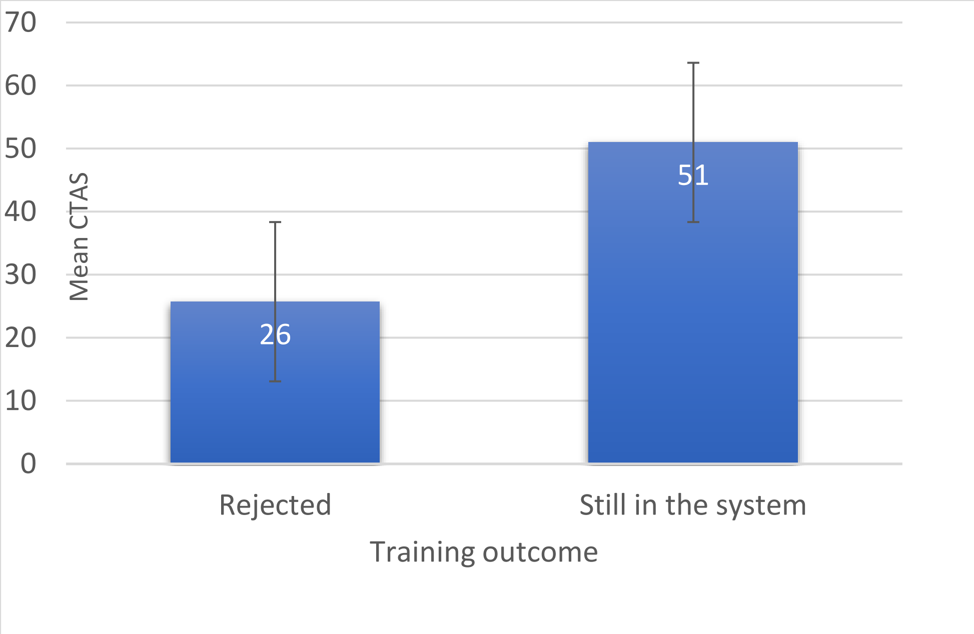


Figure 3.4 *Differences in mean CTAS related to training outcome (N=39) Bars shows standard error.*

3.4.2.8 Clustering of dog traits into components

The PCA produced five components which explained 68.8% of the total variance. I interpreted the components derived from orthogonal Varimax with loadings ⩾ 0.4. The components were labelled according to their major loadings as follows: Component 1: 'Responsiveness to training', Component 2: 'Keenness to please', Component 3: 'Human orientation', Component 4: 'Tendency to be impulsive', Component 5: 'Vocal excitability' (Table 3.11).

Table 3. 11 *Structure matrix of PCA with Varimax rotation, main loadings contributing to each component ≥0.4 bolded. Eigen values >1 (N=56).*

Variable	*Component*					
	1. Responsiveness to training	2. Keenness to please	3. Human orientation	4. Tendency to be impulsive	5. Vocal excitability	Communalities
Motivation during training	**0.664**	0.224	0.263	0.253	-0.010	0.624
Confidence	**0.622**	0.383	-0.103	0.204	-0.288	0.668
Attachment to human partner	0.137	0.240	**0.752**	0.146	-0.109	0.675
Acuity of sense of smell	**0.818**	-0.096	0.23	-0.145	0.179	0.784

Tendency to investigate humans by sniffing	0.397	0.053	**0.575**	-0.009	-0.060	0.495
Obedience to human command	0.046	**0.703**	0.385	-0.043	-0.030	0.648
Impulsivity	0.023	-0.067	0.071	**0.914**	0.062	0.849
Consistency of behaviour	0.153	**0.815**	-0.171	-0.240	0.019	0.774
Ability to solve problems	**0.718**	0.237	0.277	0.135	-0.128	0.683
Willingness to try new behaviours	**0.442**	**0.542**	0.075	**0.477**	-0.154	0.747
Tendency to vocalise	-0.219	-0.135	-0.055	0.001	**0.728**	0.600
Willingness to bring an object back to a person	0.325	**0.561**	0.323	0.290	0.026	0.609
Motivation to obtain food	**0.820**	0.157	0.162	0.067	-0.020	0.729
General excitability	0.097	-0.058	0.307	**0.597**	**0.431**	0.649
Tendency to be distracted when working	**-0.556**	-0.120	0.032	0.087	**0.493**	0.575
Motivation to play with toys	0.300	0.363	0.007	0.230	**0.691**	0.752
Tendency to seek human attention	0.130	-0.051	**0.873**	0.138	0.160	0.827
% Variance explained	32.41	13.51	8.76	7.48	6.59	*Total variance* 68.75%

3.4.3 *Trained dogs' scent sensitivity and specificity scores*

The mean scent sensitivity score was 79%, varying from 50% to 100% between conditions (SD= 0.14), and the mean specificity was 80%, ranging from 70% to 100% (Table 3.12). Pearson correlation showed that the two values were not significantly correlated. Therefore, I utilised them independently as outcome measures in the following phases of the research project.

The dogs' projects included COVID-19 (N=13), Canine Bladder Cancer (N=2), Malaria (N=1), Parkinson's disease (N=2), Pseudomonas Aeruginosa (Pa) (N=3), Prostate Cancer (N=4) and Canine Olfactory Thresholds' project (COT) (N=3). The number of scent exposures for each dog ranged from 14 to 3125 occasions (M= 1130.78, SD= 1007). The number of samples varied from 5 to 113. The scent samples' materials differed between targets and included socks, tennis balls and biological substances, such as urine and Pa in mixed culture (Table 3.13). Few dogs were trained for each disease, and the complexity of the procedures and 'trials' differed notably by condition (Table 3.13). Hence it was not possible to split the sample by target disease.

Concerning trial complexity, For 15 dogs, sensitivity and specificity were calculated from testing trials across conditions (except Malaria, where testing data was unavailable). For Prostate Cancer, Pa and Parkinson's disease, data was from testing trials only (Table 3.13). Some dogs trained on COVID-19, Canine Bladder Cancer, Malaria and COT did not yet have testing trials or were withdrawn before testing. Therefore, I used the most recent training data with the highest complexity. The COT project data was from an experiment investigating the dogs' olfactory threshold for bio-detection purposes with amyl acetate, a compound frequently utilised for olfactory studies (Concha et al., 2014). The Pseudomonas project has data from four testing stages (Davies et al., 2019). I used the data from the last stage (Detecting Pa in mixed culture) as it was the most recent and complex.

Table 3. 12 *Ranges of sensitivity and specificity scores for dogs trained on each target scent (N=27).*

Target scent	N	Sensitivity				Specificity			
		Mean	Min	Max	SD	Mean	Min	Max	SD
Covid-19	13	0.80	0.51	1.00	0.12	0.89	0.70	1.00	0.11
Pseudomonas	3	0.67	0.50	0.90	0.21	0.92	0.88	0.94	0.034
Parkinson's	2	0.85	0.70	1.00	0.21	0.89	0.88	0.90	0.014
Malaria	1	0.80	0.80	0.80	0.00	0.92	0.92	0.92	0.00
Prostate Cancer	4	0.80	0.71	0.92	0.11	0.82	0.70	0.98	0.12
Canine Bladder Cancer	2	0.85	0.84	0.86	0.14	0.89	0.84	0.95	0.08

Amyl Acetate (canine olfactory threshold trial)	2	0.78	0.57	0.99	0.30	0.91	0.90	0.91	0.01
All	27	0.79	0.50	1.00	0.14	0.88	0.70	1.00	0.09

Table 3. 13 *Numbers of exposures and presented samples, target source, whether these were training or testing trials, for each target scent (Information incompletely provided for some conditions).*

Target scent	N	Exposures nº				Samples nº				Material/ scent	Trial (N)		
											Training		Testing
		Mean	Min	Max	SD	Mean	Min	Max	SD		Unblinded	blinded	
Covid-19	13	1069.92	14	2811	1033.22	201.77	5	425	196.89	Sock-Tennis Ball	6	2	5
Pseudomonas	3	-	-	-	-	104	-	-	-	Pa in mixed culture with other organisms	0	0	3
Parkinson's	2	-	-	-	-	-	-	-	-	-	0	0	2
Malaria	1	3125	3125	3125	3125	340	340	340	340	Sock	0	1	0
Prostate Cancer	4	-	-	-	-	-	-	-	-	Urine	2	0	2
Canine Bladder Cancer	2	790	761	819	41.01	212	211	214	2.12	Urine	0	0	2
Amyl acetate	2	870	741	999	182.43	29	21	37	11.31	Amyl Acetate solution	0	1	1
All	27	1130.78	14	3125	1007	186.84	5	425	174.75	-	8	4	15
										Total of each trial sort %	29.6%	14.8%	55.6%

3.5 Discussion

This chapter aimed to explore how medical detection dogs are evaluated and selected to ultimately produce standardised success measures as performance parameters for subsequent chapters, deriving three success measures

a) Training outcome, a potential functional binary parameter of success

b) CTAS, a subjective measure that allows assessing dogs' broader ability range and their lacking in ability.

c) Scent detection sensitivity and specificity scores, an objective detection measure.

Study A explored how trainers rated their dogs' with established rating scales on different occasions, finding variations for each trait across disciplines and training stages and identifying traits significantly associated with OA. The ODS showed some attributes highly discrepant from the ideal levels reported in the survey (2.4.4). The CTAS integrates OA with ODS in a composite measure that indicates the dogs' aptitude when training or working but considers their discrepancy from ideal. The study also explored and derived scent sensitivity and specificity scores for each dog.

3.5.1 *Binary outcome: Pass and Fail as a measure of training success*

There were 39 trainee dogs at the test battery; slightly more than half remained in the system (53.1%) at the end of data collection. This figure corresponds with previous studies on working dogs, with an approximate success rate of 50% (Wilsson and Sundgren, 1997a; Slabbert and Odendaal, 1999; Batt et al., 2008; Sinn et al., 2010). However, nearly half of the dogs procured to be MDD (41%) were withdrawn. This rate is still high, leaving space for improvement since rejection due to behavioural issues causes critical economic losses (Cobb et al., 2015).

Most dogs (N=11) were rejected during the socialisation stage, and five were in specialised training. Suggesting that the internal methods of the charity identify dogs unsuitable for the task early and exclude them before investing more resources in them. Only three dogs were rejected after ending specialised training. Possibly, dogs that passed training had a greater success probability when operational. However, whether these dogs will work efficiently in the long-term is unknown since they were only followed until data collection ended. In longer-duration studies, several operational dogs were dismissed for behavioural reasons (e.g. Caron-Lormier et al., 2016). Longitudinal studies assessing dogs over time may aid in identifying factors associated with long-term success.

Among dogs’ withdrawal reasons were deficient search ability, lack of problem-solving skills, decision-making and tendency to take the initiative. Also, undesirable traits like anxiety, frustration, impulsivity and aggressiveness. Some of these traits, considered highly important for dog selection in the survey (2.4.5), were rated as far from ideal in the current sample.

Training outcomes may indicate a dog's potential for future working performance. However, its binary nature, and lack of standardisation, limit our understanding of which aspects of performance lead to failure. Hence, exploring how the trainers assess the dogs and which traits are most relevant for evaluating the dogs' performance may help develop more accurate performance parameters.

3.5.2 How do the trainers perceive the dogs' behaviour in training, and how does this affect their OA ratings

The trainers utilised most of the rating scale for nine traits, such as 'Tendency to vocalise', 'Frustration during training', and 'Ability to solve problems'. Some others showed less variability (but still ⩾20% were different to the mode). Some traits, such as 'Acuity of sense of smell' or 'Confidence', may be homogeneous due to genetics, standardised breeding, selection and training methods, or trainers may not distinguish them effectively.

The trainers' interpretation of the traits' implications and the rating scales may have influenced rating differences. I researched the terminology MDD professionals used to derive standardised and clear trait descriptions and benchmarked the rating scales (e.g. 1= Very low 6=Extremely high). The development of the scales was based on a systematic approach validated for other specialised detection tasks (Rooney et al., 2004; Rooney and Clark, 2021). Previous research observed that assigning labels for each scale point increased rating effectiveness (Clark and Rooney, 2021) compared to just a numerical range.

Human social studies indicated that even minor scale alterations might change ratings (Conti and Pudney, 2011). Each trainer may perceive the dog trait names and scales differently from their perspective. Labelling each scale point with a specific description for each trait may increase clarity. However, it could be impractical to elaborate and arduous for the raters when rating many behavioural aspects.

3.5.3 Differences in ratings across disciplines

Ratings of Assistance dogs' OA and most traits surpassed those of Bio-detection dogs. For instance, 'Confidence', 'Ability to solve problems' and 'Ability to learn from being rewarded' were significantly higher in Assistance dogs than in Bio-detection dogs. Differences across disciplines may relate to the trainers' awareness of the traits needed for each task, as dog traits differ depending on the nature of the role (Bray et al., 2021b).

Assistance dogs may have been preselected using stricter criteria, as these dogs should be calmer, cope with public environments and work continuously. Hence their training performance may be better. While for other detection roles, Bio-detection dogs tend to have a more active profile and work for limited periods without being constantly exposed to the external environment. Hence selection standards may be more lenient.

Some traits may be easier to observe in certain circumstances. For Bio-detection dogs, 'Tendency to vocalise' was rated higher than in Assistance dogs. Bio-detection dogs remain near their trainers in holding pens between training sessions, and the trainers would notice if they vocalised. However, Assistance dogs spend longer with their trainers, placing them in pens when needed in a different facility. Thus, they may not hear them vocalise as much as Bio-detection dog trainers.

Trainers may be biased when assessing the dogs. Raters tend to be lenient when scoring performance that reflects their work (Pronin et al., 2002; Clark et al., 2020b). The trainers' background, experience and interaction with their dogs may affect their judgment (Clark et al., 2020b), but this may differ between disciplines. For instance, assistance dog trainers may be more optimistic, while bio-trainers perhaps are more critical as their role imply specific research skills to collect and quantify detection data. Each training team may have a consensual opinion on what to observe and the individual dog's performance, developing a group bias in 'competition' with others (Brewer, 1979).

3.5.4 Differences in ratings across training stages

Trained dogs were rated higher than trainee dogs for OA and several desirable traits. Suggesting those who succeed in training tend to perform better than those still learning, as expected. Some behavioural attributes may have improved with maturity (e.g. Bray et al., 2021a), such as 'Concentration' and 'Ability to learn from being rewarded'. However, 'Motivation when training' was relatively similar for both groups. Trainers may have a higher opinion on dogs that succeeded in training, reflecting their achievement or their group's work (Brewer, 1979; Clark et al., 2020b). They may be more attached to the dog as they spent longer together, and some trained dogs were owned by the trainers or lived temporarily with them.

However, most undesirable traits did not differ between training stages, except for 'Tendency to be distracted when working', which was higher in trainee dogs, as younger dogs may struggle more to concentrate during training.

3.5.5 Dog trainers' ratings' consistency over time.

OA ratings of some Bio-detection dogs increased or decreased, but most stayed the same. Most Bio-detection dogs with rising scores or remaining stable were kept in the system (61.5%), while dogs with declining ratings tended to fail (38.5%). Assistance OA ratings were high on the scale and were similar for both time points, dogs whose scores increased over time generally passed.

Trait consistency varied between disciplines. For Assistance dogs, traits highly consistent were. 'Obedience’ and 'Ability to learn from being rewarded' 'Adaptation to crate or kennel' and 'Tendency to vocalise'. 'Confidence' and 'General excitability' had acceptable consistency. For Bio-detection dogs, no traits were highly consistent; only some characteristics had acceptable consistency, i.e., 'Obedience', 'Impulsivity', 'Motivation to play with toys', and 'Willingness to bring an object back'.

Personality is characterised by being consistent throughout life (Jones and Gosling, 2005). Traits may be stable over time due to their inherent nature and preselection. Dog traits have shown varying consistency when tested on different occasions (Fratkin et al., 2013). Some traits were highly consistent, such as 'Impulsivity' (Riemer et al., 2014a) and 'Exploratory activity' (Riemer et al., 2014b). Others were less consistent, such as 'Trainability' and 'Sociability' (Fratkin et al., 2013; Harvey et al., 2016). Although these studies differ in methodology, populations assessed, number of assessments and intervals between them. Performance is expected to improve with training and development (Fratkin et al., 2013; Bray et al., 2021a). Hence ratings may change for some traits with training progress.

Trainers' perceptions of dog performance may be similar over time. Trainers may have noticed certain traits early and rated them similarly over training, influenced by their experience or preconceived notions about the dogs' behaviour (Clark et al., 2020b). Negative traits, such as 'Tendency to vocalise' in Assistance dogs, and 'Impulsivity' in Bio-detection dogs, may be rated more consistently due to the dog instructors paying particular care to characteristics potentially disruptive to performance.

Most traits showed very low consistency. This may be due to: a) the trait's nature, b) the trainers not being reliable enough, and c) the assessment method may have failed to capture traits consistency. However, this is unlikely since ratings have shown more consistency over time than codings (McGarrity et al., 2016). It has been suggested that using the same evaluation methods with short intervals between assessments, as done here, increases consistency (Fratkin et al., 2013). Similar methods were used successfully to assess detection dogs on repeated occasions (e.g. Rooney et al., 2007a; Rooney and Clark, 2021). In contrast, intra-rater reliability in these studies was generally good, while here was weaker. However, these studies had larger, more homogeneous samples and applied different analyses than here.

The disparity of evaluations for each discipline did not allow for comparing progress across them. Ideally, larger samples and all trainers completing the assessments for all time points may have improved this analysis. However, a systematic follow-up of dogs' training progress may determine aspects needing continuous attention.

3.5.6 Associations between individual behavioural traits and ratings of OA'

Higher scores of 'Ability to solve problems', 'Willingness to try new behaviours' and 'Tendency to investigate humans by sniffing', but lower levels of 'Body sensitivity' and 'Adaptation to crate or kennel' were significantly associated with increased OA ratings. When assessing these relationships for each task, Bio-detection dogs were similar to the whole sample, probably because they were the majority. For Assistance dogs, the most relevant traits were 'Willingness to bring an object back' and 'Tendency to investigate humans by sniffing'. I will discuss results for the overall population and mention significant differences between the two disciplines.

'Ability to solve problems' was highly associated with OA for all dogs. Problem-solving skills have been linked to better performance in different detection and service roles (e.g. MacLean and Hare, 2018; Lazarowski et al., 2019b). This trait was essential for bio-detection tasks in the survey (2.4.5), but it was rated as below ideal in this study. Suggesting a highly relevant trait that may need attention.

Dogs with a higher 'Willingness to try new behaviours' were more likely to receive a higher OA rating. This trait has been associated with better alerting in Assistance dogs, according to trainers' reports (Rooney et al., 2019) and was also among the most important traits for assistance tasks in the survey (2.4.5).

'Tendency to investigate humans by sniffing' was expected to be more relevant for assistance tasks, and it did contribute significantly to their OA. These dogs must screen humans continuously to identify odour changes. This attribute was related to OA for all dogs', showed high variation and was highly consistent over time for the bio-detection and trainee population. Dogs investigating humans may reflect their inherent curiosity about the environment and relate to a higher tendency to cooperate with their handler regardless of the task (1.2.6.2).

'Body sensitivity' was negatively associated with OA and was higher than ideal for all dogs. Body sensitivity may be from, e.g. early experience, behavioural disorders or health conditions (e.g. Wiseman (Wiseman-Orr et al., 2001)-Orr et al., 2001). High body sensitivity may elicit dogs' over-reactivity in different situations, affecting their ability to focus on their task, and it may impair Assistance dogs' wearing the required service coat.

Surprisingly 'Adaptation to crate or kennel' was negatively linked with OA. In several studies, dogs more relaxed when kennelled are preferred for better performance and welfare (Rooney et al., 2003a; Rooney et al., 2007a). Bio-detection dogs may be more energetic and potentially more reactive when confined. This trait in Assistance dogs was much lower than ideal. Dogs suited for this role are preselected to bond strongly with the clients, and some may be restless when apart. Their inability to relax in isolation may be aversive to performance (Batt et al., 2009).

For Assistance dogs, 'Willingness to bring an object back' was positively associated with OA. This trait was also rated significantly higher for Assistance dogs than Bio-detection dogs, as they are trained to bring a medical kit to clients with a health crisis.

3.5.7 Discrepancy scores

Assessing discrepancies for each trait allowed us to identify particular behavioural aspects below or above the rater’s perceived ideal levels, similar to Serpell (1996) and Rooney et al. (2004). These divergences may be problematic during training and ultimate task operation. Hence these traits require higher focus when training and selecting the dogs to prevent them from being a cause of withdrawal. The ODS added each trait discrepancy as an integrative measure of dogs' performance shortcomings.

Overall, trained dogs tended to have lower discrepancy scores than trainee dogs, and Assistance dogs' trait ratings were nearer to ideal levels than Bio-detection dogs. These tendencies also correspond with the differences in ratings across disciplines (3.4.2.3) and training stages (3.4.2.4) discussed before. For Bio-detection dogs, 'Ability to solve problems' was farthest from ideal. Low levels of this relevant trait were related to dogs' training deficiencies. However, it was not among the most discrepant characteristics in the survey

(2.4.6.2), suggesting it maybe not be a generic issue for detection dogs but may be specific for this organisation.

'Tendency to search by smell alone' was highly discrepant for bio-detection. Lack of searching skills is a frequent withdrawal reason in this population (3.4.1). For Assistance tasks, traits potentially related to human attachment, i.e. 'Tendency to seek human attention' and 'Adaptation to crate or kennel', were far off from ideal, like for the Assistance dogs from the survey (2.4.6.2). Negative traits, like 'Body sensitivity', 'Tendency to be distracted when working', 'Impulsivity' and 'Tendency to vocalise', were among the ten most discrepant traits for one or both tasks in both the current study and the survey outcomes.

Traits like 'Ability to learn from being rewarded' for Assistance dogs and 'Obedience' for Bio-detection dogs may be closer to ideal since attention to these may be emphasised in selection and training.

3.5.8 Deriving the CTAS

The CTAS provides a complete overview of the dog's ability and general deficiencies. Similar composite scores in explosives detection (Rooney et al., 2007b) and stuck-guarding dogs (McConnell et al., 2022) have shown strong associations with objective performance measures.

In McConnell et al. (2022), the same dog handler rated and ranked the livestock guarding dogs for most attributes. However, in this study, different trainers rated different dogs. They may have different interpretations of the ability rating scale. In the charity, the same person cannot examine all dogs. Staff training may aid in homogenising the evaluation of dogs' ability to address this limitation.

The CTAS varied significantly with training outcome. Suggesting that the measure effectively reflects the dog's success level. However, the training outcome is a valid measure, but its binary character does not give the same gradation as the CTAS. This supports an argument for using both of these measures independently in subsequent chapters as they provide different performance information.

3.5.9 Dog traits' overall relevance

All the traits assessed here are relevant at different levels for early assessment. However, several characteristics stood out for their variability level, association with OA, and their difference with ideal levels and importance in the survey. Also, specific traits differed across disciplines and training stages.

Some traits that repeatedly emerged for both disciplines were 'Ability to solve problems', 'Willingness to try new behaviours', 'Ability to learn from being rewarded', and 'Body sensitivity'. These were highly linked with OA ratings and were among the most important traits in the survey. Focussing on these traits during socialisation, training, and evaluation may improve ultimate performance. Some traits far from ideal or related to rejection, like 'Ability to solve problems' or ‘Body sensitivity', require special attention to approach expected levels.

3.5.10 Clustering of dog traits into components

The components summarise trait ratings into main dog behavioural aspects over training. The first two were ostensibly positive for performance: 'Responsiveness to training' involved traits related to general motivation, confidence, problem-solving and search ability, and 'Keenness to please' included traits related to trainability and consistency. 'Human orientation' was linked to attachment and human attention demand (1.2.6.2). 'Tendency to be impulsive' and 'Vocal excitability' involved traits desired at very low levels in the survey (2.4.4). These components were effective in assessing their relationship with the dogs' performance in the test battery (4.4.5).

3.5.11 Trained dogs' sensitivity and specificity scores

Sensitivity and specificity scores have been useful in indicating individual detection effectiveness in previous studies (e.g. Lazarowski et al., 2015; Concha et al., 2019). However, these can vary significantly between working disciplines. Here, the scent samples for each health condition were small, and each condition required different sample handling procedures and training protocols, thus not allowing comparisons between target scents. Larger dog samples for each scent project would have allowed greater insight into variations within and between illnesses (Jezierski et al., 2015).

Finding large MDD populations is challenging. One way to increase the number of dogs would be to accumulate the data over a long period. Future research could focus on individual diseases to obtain specialised information or comparative studies with larger samples.

The data was collected in training or testing trials with varying complexity. Determining which to use for each dog was challenging. However, the criteria to use data from the most complex and recent trial aided in increasing data standardisation.

I evaluated scent sensitivity and specificity in subsequent chapters (Chapters 5-6) to analyse the association of dogs' behavioural factors with their decision-making propensities in search scenarios.

Study B: How predictive and reliable is an in-house test for dogs’ overall ability?

3.6 Methods

Thirteen prospective medical detection dogs (N=13) in socialisation, aged 9 to 14 months, were assessed in a one-off aptitude test used in-house by the MDD charity. Between two to five assessors (varied across tests) evaluated each test. The assessors were all dog trainers with differing backgrounds and experiences. The leading assessor conducted the trial and manipulated the materials, the trainer handled the dog, and the observer monitored the trial and assisted the leading assessor.

3.6.1 Aptitude test procedure

During the trial, the dogs were encouraged to find a tennis ball hidden under one of three metallic cans (similar to the S-12 Ball-searching subtest from the test battery; 4.2.5.2.4). Initially, the dog entered the room with their trainer, explored the space and was permitted to greet the people inside. The observers were seated, standing or assisting the assessor.

The lead assessor threw a tennis ball several times at the dog. Subsequently, the trainer held the dog by the collar facing the lead assessor while they aligned three metallic cans upside down on the floor and hid the tennis ball under one of them. The assessor then slid the cans with crossing movements to mix them, so the dog would lose sight of the target can and need to locate the ball using their sense of smell. When the assessor stopped, the dog was encouraged to sniff them to indicate the can hiding the ball. If the dog found the ball, the assessor praised them verbally, threw it, and encouraged the dog to retrieve it; if it did not find the ball, the assessor repeated the mixing. After several repetitions, the assessor substituted the tennis ball with smaller ball fragments (there was no standardised number), reducing their size progressively and throwing them a tennis ball if they found the smaller ball pieces (Figure 3.5).

Figure 3. 5 *Dog tested with aptitude test.*

3.6.2 Assessors' ratings on the aptitude test

I asked each test assessor to rate the dogs following the test using the same rating scales from Study A (3.3.3). They rated only 21 traits since I excluded those not potentially observable in this test, e.g. 'Confidence in different environments', 'Motivation to obtain food', and 'Level of attachment' (Appendix 7). The assessors completed the rating sheets with the dog's information (name, age, sex and sexual status), the assessor's name and their role in the test. They indicated any previous assessment dates. The assessors were advised to rate the dogs individually without discussing the trial's performance. I explored how much variation the assessors showed in their ratings, the level of agreement between the test assessors, and how predictive their ratings of training outcomes were.

3.7 Data analysis

3.7.1 Assessors' ratings variability

I explored the assessors' ratings with descriptive statistics and graphic aids. I examined the spread of each trait and their OA ratings to explore how they used the scale throughout. Assessors differed between dogs, with each evaluating a different number of dogs (Assessor one- 12 dogs, Assessor two- 8 dogs, Assessor three- 8 dogs, Assessor four-7 dogs, Assessor five-7 dogs, assessor six- 2 dogs, assessor seven- 1 dog). Since Assessors six and seven rated one and two dogs, they were excluded from further analysis.

3.7.2 Agreement between assessors

To assess the strength of agreement between the five assessors, I carried out a Krippendorf's Alpha analysis for each trait(3.3.2.4). I also performed this analysis for each pair of raters for each trait to explore whether all assessors agreed when rating the dog traits or if there was an outlier. Most raters tended to agree with their peers strongly; however, Assessor three had a very low agreement with three assessors out of four (α

<0.67; Table 3.15). I repeated the analysis for each trait, excluding Assessor three's ratings, to examine changes in agreement.

3.7.3 Independence of traits ratings

I performed Spearman's Rank correlations between ratings for the different traits. I extracted and organised the correlation coefficients by attribute. Then, I calculated the mean and transformed negative into positive numbers to get positive figures and produce the mean coefficient for assessors' correlations for each pair of traits. Similar to Rooney and Clark (2021), I created a matrix indicating the number of assessors with correlation coefficients ≥0.70 for each trait and the maximum number of assessors that rated that trait (not all rated all the attributes, e.g. some were rated by four or five assessors). Pairs of traits that were highly correlated by four or five assessors (≥80% of the raters) were considered highly correlated; by three assessors, acceptably correlated, and by less than three assessors, weekly correlated. However, when only three raters assessed a trait, it was not considered in the final analysis as there were few raters. Highly correlated traits were removed from further analysis by keeping the attribute with more correlations and those deemed more biologically relevant and observable during the test.

3.8 Results

3.8.1 How predictive and reliable is an in-house test for medical detection dogs?

The assessors generally showed a broad use of the scale with OA ranging from two to nine (Mdn=7, Min=2, Max=9, SD= 1.25). Several traits showed high variability, like 'Motivation during training', 'Confidence', and 'Willingness to try new behaviours'. Five traits had a limited spread: 'Frustration during training', 'Tendency to be distracted when working', 'Tendency to explore areas by sniffing' and 'Tendency to seek human attention'. The assessors rated 'Tendency to explore areas by sniffing', 'Tendency to seek human attention' and 'Tendency to investigate humans by sniffing' in less than 70% of the dogs, contributing to the low variability. I excluded these traits from further analysis (Figure 3.6).

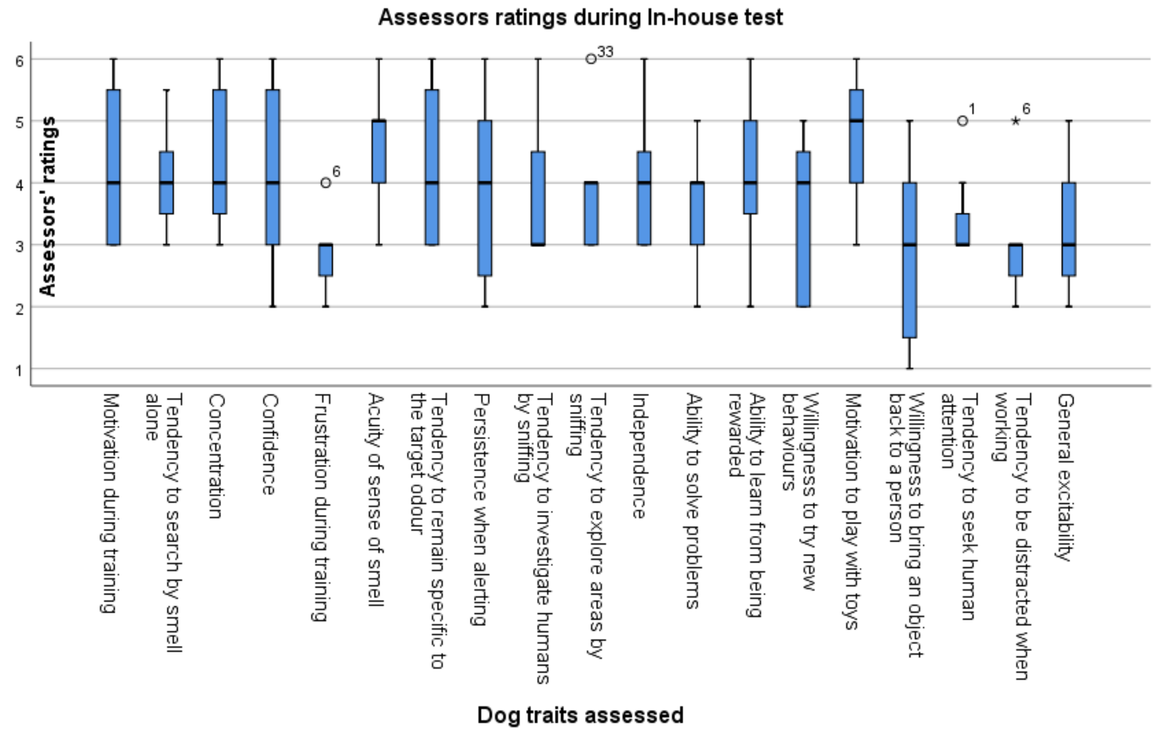


Figure 3. 6 *Assessors' rating distribution for different dog traits. Dogs N=13.Trait ratings: 1=Very low, 2=Low, 3=Intermediate, 4=High, 5=Very high, 6=Extremely high (plots meaning in Table 3.2).*

3.8.2 Agreement between assessors

'Level of Motivation during training' showed high agreement between assessors, and 'Tendency to search by smell alone' showed acceptable agreement. 'Willingness to try new behaviours' showed low agreement. However, most traits had a very low agreement. Excluding Assessor three increased Krippendorf's Alpha coefficient for six traits (Table 3.14; 3.15). The increase was greatest for 'Motivation during training', 'Tendency to search by smell alone and 'Concentration'. Several other characteristics slightly decreased, but the agreement for several was still very low (Table 3.15).

***Table 3. 14** Krippendorf's Alpha coefficients for each pair of assessors (N=5) rating trainee dog performance (N=13) in an in-house aptitude test. High agreement (>0.8), Acceptable agreement (<0.7 ≤0.8), Low agreement (≥0.67-≤0.7), Very low agreement (<0.67)*

	Assessor One	Assessor two	Assessor three	Assessor four	Assessor five
Assessor One	1.00	0.83	0.30	0.51	0.76
Assessor Two	0.83	1.00	0.48	0.70	0.70
Assessor Three	0.30	0.48	1.00	0.88	0.55
Assessor Four	0.51	0.70	0.88	1.00	0.79
Assessor Five	0.76	0.70	0.55	0.79	1.00

***Table 3. 15** Krippendorf's Alpha coefficients for each trait with five assessors and how these change when taking out a* potential outlier assessor three

Trait rated	Krippendorf's Alpha with five assessors	Krippendorf's Alpha without Assessor three
Motivation during training	0.81	↑ 0.84
Tendency to search by smell alone	0.71	↑ 0.78
Concentration	0.58	↑ 0.69
Confidence	0.41	↓ 0.38
Acuity of sense of smell	0.49	↓ 0.43
Tendency to remain specific	0.54	= 0.54
Persistence when alerting	0.56	↓ 0.48
Independence	0.37	↑ 0.52
Ability to learn from being rewarded	0.65	↓ 0.64
Willingness to try new behaviours	0.69	↓ 0.68
Motivation to play with toys	0.28	↑ 0.41
Willingness to bring an object back	0.21	↑ 0.52
General excitability	0.56	↓ 0.50

High agreement (>0.8) Acceptable agreement (<0.7 ≤0.8) Low agreement (0.5-0.59)

=

Very Low agreement (≥0.67-≤0.7)

↑ = Increase ↓ = Decrease = Remain the same after excluding Assessor three

3.8.3 Independence of trait ratings

'Concentration', 'Tendency to remain specific to the target odour' and 'Independence' were highly correlated with six other traits and 'Motivation during training' with five. 'Motivation during training' was kept over the others since it showed high variability and agreement between assessors and correlated with several characteristics. Some traits showed little correlation to others, like 'Willingness to bring an object back' and 'General excitability', where only two assessors or fewer had high correlations. (Table 3.16)

.

Table 3. 16 *Number of assessors (N=5) with correlation coefficients >0.70 for each trait pair in an aptitude test (less than five raters rated several traits).*

Correlations	Motivation during training	Search by smell alone	Concentration	Confidence	Acuity of sense of smell	Remain specific to odour	Persistence	Independence	Ability to solve problems	Ability to learn	Try new behaviours	Toys motivation	Bring an object back	General excitability
Motivation during training	5	4	5	3	3	3	3	5	3	4	3	3	2	0
Search by smell alone	4	5	4	3	3	4	1	2	1	2	2	2	2	0
Concentration	5	4	5	3	4	3	4	5	3	5	3	3	2	0
Confidence	3	3	3	3	2	3	2	3	1	2	3	2	1	1
Acuity of sense of smell	3	3	4	2	5	3	2	4	3	4	3	2	2	0
Remain specific to odour	3	4	3	3	3	5	4	4	4	5	3	3	1	0
Persistence when alerting	3	1	4	2	2	4	5	3	4	4	3	1	2	0
Independence	5	2	5	3	4	4	3	5	4	4	3	3	0	0
Ability to solve problems	3	1	3	1	3	4	4	4	4	3	3	1	1	0
Ability to learn	4	2	5	2	4	5	4	4	3	5	2	2	2	0
Try new behaviours	3	2	3	3	3	3	3	3	3	2	5	1	0	1
Toys motivation	3	2	3	2	2	3	1	3	1	2	1	5	0	2
Bring an object back	2	2	2	1	2	1	2	0	1	2	0	0	5	0
General excitability	0	0	0	1	0	0	0	0	0	0	1	1	0	5

Number of assessors with correlations ≥0.70 =Five =Four =Three =Less than Three =Correlated with themselves (maximum number of raters rating that trait since not all 5 raters rate each trait)

3.8.4 Association with training outcome

The dogs' OA in the test was higher in successful dogs (7.9) than in those who failed training (6), though this difference was not significant.

3.9 Discussion

Study B assessed the reliability and predictive value of a single one-off in-house aptitude test. Traits showed different levels of variation. There was low agreement between assessors, and they perceived several characteristics to be similar. These results led to prioritising the most valuable attributes to develop assessment methods that are effective, practical and clear, to include essential and easy-to-understand constructs, distinguishable by the raters at an individual and group level (Rooney and Clark, 2021). For this sample, the test was not predictive of training outcome.

3.9.1 The in-house test procedures

The in-house test is relatively short and easy to implement, does not require expensive materials, and is not potentially aversive to the dogs. It allows observing aspects related to motivation, stamina, obedience and olfactory abilities.

However, the aptitude test lacks standardisation, the number of assessors varies, and it was not well-established what elements they evaluated. The assessors tend to discuss the dog's performance while testing the dogs, influencing the assessor's individual opinions. The observations on general performance may be absorbed in a group consensus (Rooney and Clark, 2021), possibly led by the members with more experience, which affects the variability and reliability of the observations. Yet, assessors showed low agreement when rating the individual traits, which may be that they get to a consensus when discussing the dogs' overall aptitude, but not when evaluating different behavioural aspects with systematic rating scales they were naïve to.

For several traits, ratings had a wide distribution on the scale, such as 'Motivation during training', 'Confidence' and 'Willingness to try new behaviours', which were highly important in the survey. Variability was low for four traits. As in Study A, some traits may show low variability due to preselection or because the assessors overlooked them. However, some traits may be difficult to infer from a single observation, like 'Frustration during training'. Whilst others may not be evident during testing like 'Tendency to investigate humans by sniffing' or 'Tendency to seek human attention'. These traits may not be particularly useful when evaluating the dogs, hence may be removed in further assessments. The test could increase its effectiveness by including relevant items and being easy to distinguish in a single assessment, adding descriptions, and training the assessors on what to observe during the test (Clark and Rooney, 2021).

3.9.2 Assessors' agreement and independence of traits

Most traits showed low agreement as assessors strongly converged only for 'Motivation during training' and 'Tendency to search by smell alone'. This suggests that the assessors' interpretations differed notably for most traits and may indicate a lack of shared understanding of the construct (Rooney and Clark, 2021). Removing assessor three slightly improved agreement but did not substantially impact Krippendorf's Alpha levels.

Thus, it is worth considering if the test is testing what it is supposed to assess, if people agree with what they see and if the assessors understand what the trait refers to. Accordingly, the test items may be reduced to those directly observable and those for which the trainers' have a standardised interpretation.

Among the study's limitations was the reduced data available. Since the number of assessors differed for each dog and each assessor examined a different number of dogs. Some ratings were incomplete. This may have been due to a lack of understanding or because it was not apparent from the test. Also, the assessors have different backgrounds limiting standardisation. Assessors' agreement could improve with a larger sample and a standardised number of assessors for each dog completing the sheets in full.

Eight traits had several correlations with others (R≥ 0.7) for most assessors (4-5). Most of these were relevant in the survey, such as 'Motivation during training', 'Concentration' and 'Independence'. These traits may be biologically linked and are all part of a multifaceted construct. However, the assessors may perceive these traits as similar, a halo effect of one trait entailing the others, and then rate them all the same (Rooney and Clark, 2021). Raters' training on the constructs' implication and the use of scales may aid in increasing their understanding and increase agreement (Fratkin et al., 2015; Clark and Rooney, 2021) since the assessors were naïve to the assessment and may have interpreted the traits descriptions differently when observing the dogs.

3.9.3 Association with training outcome

The in-house test was not significantly associated with success in training. Assessors' agreement was generally low. However, the sample is too small to reach accurate conclusions.

Dogs are mainly tested when they are young. Several may progress to specialised training and then show inappropriate behaviour, eventually leading them to fail. The test may be insufficient to decide the dogs' fate since it allows only a snapshot of the dogs' general behaviour.

Initially, I filmed dogs during the test to collect objective data. However, assessing this was not possible due to time limitations. Including objective measures to quantify the dogs' performance, e.g. measuring frequencies and durations when finding and retrieving the ball, would increase the test robustness. Future research could replicate these methods and include these. The current test could be combined with cognitive and temperament tests to integrate the measurement of essential traits, such as problem-solving skills, initiative-taking, and cooperation with the trainer (further assessed in the test battery). Setting test-retest consistency from multiple evaluations would be valuable to determine the dogs' progress over time.

3.10 Conclusion

The current chapter thoroughly explored how trainers evaluate dogs during training and in operational trials. Study A derived three robust measures of medical detection dogs' success: training outcome, CTAS and scent sensitivity and specificity scores. It assessed how different behavioural aspects varied to produce these measures and prioritised traits that may be more relevant when assessing dogs' performance. Study B investigated how assessors used traits rating scales in their internal testing methods and to what extent they agreed with their ratings.

Validated and standardised evaluation instruments for dogs' performance can significantly improve the consistency in dogs' training and save money and time. The decision of which traits to include could vary for the organisation's general needs, the dogs' discipline or training stage. Hence, specialised evaluation scales can target dogs for particular tasks.

Overall, prioritising more relevant traits that are easy to understand and distinguish from the raters is recommended. This is an initial exploration of the effectiveness of dog assessments and performance measures. Further validation is needed. The replication of these procedures with preliminary raters trained on the implication of each construct and how to use the rating scales, e.g., regular workshops and training sessions for the staff, would help potentially use these findings for practical purposes to improve the consistency in dogs' training and produce more successful animals.

Subsequently, Chapters 5 and 6 explored the link between the dogs' behaviour in a test battery and a cognitive bias test with the dog's success measures from this study.

Chapter 4. Development of a test to measure individual attributes in MDD dogs:

Which variables can be measured and how do they associate with each other?

Abstract

This study explored the development of a test battery to identify how relevant attributes vary in a sample of medical detection dogs (MDD). The test included 18 temperament and cognitive subtests to assess the most important MDD traits from the survey (2.4.7). The test battery's measures derived 11 behavioural components. Several had significant associations with dogs' demography: 'Food orientation' was higher in females (p=0.05), 'Playfulness' was higher in younger dogs (p=0.005) and trainee dogs (p=0.027), 'Obedience' increased with age (p=0.030) and training stage (p=0.015) and 'Confidence' scores were higher in younger dogs (p=0.031). 'Playfulness' (p=0.016) and Persistence'' (p=0.031) were significantly associated with the dogs' impulsivity scores. Four components from trainers' trait ratings (3.4.2.8) were significantly correlated with test components: *'Responsiveness to training'* with 'Obedience' (p=0.016) and 'Self-control' (p=0.033), *'Keenness to please' with* 'Interest in exploring environment' (p=0.017) and 'Success in search' (p=0.012). *'Human orientation'* was negatively associated with 'Confidence' (p=0.020), and *'Vocal excitability'* was negatively correlated with 'Success at problem-solving' (p=0.009). These results indicate that a revised version of the test could eventually be used for MDD assessment and selection.

4.1. Introduction

The test battery is frequently used to evaluate working dog performance (1.2.4.1). However, there is no specific test for MDD. This study assessed the development of a test battery to evaluate the presence of traits considered relevant for medical detection (2.4.7) in an actual MDD sample. It examined how the measures from the test clustered into behavioural components and how these are linked with demographic factors. I also assessed the association between the dogs' test performance and the trainers' opinions on their performance in their training or working trials (3.4.2.8). Impulsivity is a trait with important implications for working dogs since it has been associated with the dogs' situation responses (1.2.6.10). Therefore, I investigated the association between dogs' impulsivity measures from the DIAS questionnaire (Wright et al., 2011) and the dogs' performance in the test.

4.1.1. The test battery

Test batteries can be objective and reliable tools for quantifying behavioural aspects in a dog population (Diederich and Giffroy, 2006; Taylor and Mills, 2006; Fratkin et al., 2013). They are used in research to understand dogs' behaviour better and, in practice, to select animals with individual characteristics desired for a specific purpose or to rule out individuals regarded as unsuitable for it (Serpell and Hsu, 2001; Bray et al., 2021b).

Test batteries include subtests measuring behaviour using quantitative or qualitative methods or a combination of both (Taylor and Mills, 2006); these often involve coding or rating responses to particular stimuli (Serpell and Hsu, 2001; Jones and Gosling, 2005; Sinn et al., 2010) (1.2.4.1). The resulting variables are usually submitted to statistical data reduction to derive components involving relevant behavioural aspects (Goodloe and Borchelt, 1998; Sinn et al., 2010)

Test batteries are applied to assess rescued dogs' suitability for adoption (e.g. Hennessy et al., 2001; De Palma et al., 2005) and the evaluation of behavioural problems in companion dogs (e.g. Netto and Planta, 1997). In working dog populations, test batteries aid the selection of individuals most suitable for the task in question, frequently assessing how the dogs' performance associates with the training outcome (e.g. Svartberg, 2002; Batt et al., 2009; Bray et al., 2017b).

4.1.2. What do test batteries measure?

Test batteries examine variation in training performance regarding dogs' behaviour or features such as sex, breed or age (e.g. Wilsson and Sundgren, 1997a; Svartberg, 2002; Batt et al., 2008). Historically, test batteries have measured dogs' responses in a novel room or towards an unfamiliar person; their reaction to a specific stimulus; their quickness to obey trained cues or their involvement in games such as tug of war (e.g. Wilsson and Sundgren, 1997b; Slabbert and Odendaal, 1999). Some measures have been associated with successful training outcomes (e.g.Batt et al., 2008). Still, some are impractical and costly and use potential fear-eliciting procedures (e.g. startling stimuli) that risk sensitisation (King et al., 2003; Taylor and Mills, 2006). However, there is growing evidence that cognitive measures may predict dogs' training success (1.2.6.10). Recent test batteries' have addressed cognitive functions such as problem-solving, self-inhibition, persistence and social referencing (e.g. MacLean and Hare, 2018; Tiira et al., 2020; Bray et al., 2021a).

4.1.3. Test battery's standardisation

Diederich and Giffroy (2006) stressed the importance of producing scientifically standardised dog testing methods by considering test validity, consistency and feasibility (1.2.3). The selection of working dogs has traditionally relied on the trainers' opinion (3.1.1), and reported test batteries are not always standardised (Taylor and Mills, 2006). Therefore, it is essential to produce more dog selection scientifically scrutinised testing methods.

Test batteries offer greater objectivity than subjective assessments. However, these tend to show context-specific behaviours (McGarrity et al., 2016) and may not fully reflect dogs' reactions to real-life situations (Wallis et al., 2020). It is also challenging to know if the testing method measures what it is supposed to. Frequently, test measures are related to the owner's or trainer's ratings (1.2.4.2) and combining objective and subjective measures may provide a broader picture of dogs' behaviour predictive of working performance (Diederich and Giffroy, 2006; Taylor and Mills, 2006; Rocznik et al., 2015).

4.1.4. The importance of testing MDD

Test batteries have assessed mainly police dogs, drugs, explosives detection and guide dogs (e.g. Maejima et al., 2007; Batt et al., 2009; MacLean and Hare, 2018). Medical detection tasks differ from others, as Bio-detection dogs mainly discriminate scent samples in a controlled environment (1 2.1.4), and Assistance dogs combine scent detection with assistance tasks and have high public access demands (1 2.1.5). However, to my knowledge, there is no published assessment specific to MDD. Hence, the need to explore approaches to assess how MDD behaviour reflects current performance to improve dog selection and fulfilling specific task needs.

This chapter explored the development of a test battery for MDD, involving temperament and cognitive subtests measuring the most important traits for these roles (2.4.7). It also assessed the association between the variables produced and how these varied with the dogs' demography.

Dogs are tested at different ages (Slabbert and Odendaal, 1999; Harvey et al., 2016) to identify whether their performance is consistent over time and potentially filter out dogs deemed unsuitable for the task at a younger age (Serpell and Hsu, 2001; Cobb et al., 2015). However, longitudinal studies have found limited dog test-retest reliability for several behavioural traits, suggesting that some aspects of personality may change with development (Riemer et al., 2014b). The assessment of dogs already operational is seldom reported (Bray et al., 2021b). Chapter 3 identified differences in trainer’s ratings between MDD trainee dogs and trained dogs (3.4.2.3) and that several traits are not rated consistently over time In the current dog sample (3.4.2.4). Hence this study assessed test performance regarding age and training stage.

4.1.5. Assessing impulsivity in MDD

Impulsivity refers to the absence of inhibition when responding to a stimulus (Fadel et al., 2016). There have been recognised different aspects of impulsivity which may be related to each other (Wright et al., 2012): cognitive impulsivity, linked with low tolerance of delayed rewards and motor impulsivity, which refers to the failure to inhibit a behavioural response

(Brunner and Hen, 1997; Krishnan-Sarin et al., 2007). There is emerging interest in this inherent trait (1.2.6.10). It has been linked to self-harmful behaviours in humans and psychiatric disorders (Martin et al., 1994; Brunner and Hen, 1997; Moeller et al., 2001), and it has been considered detrimental to dog behaviour (e.g. Fatjo et al., 2005). It is of particular interest here because evidence suggests that working dogs may have a greater tendency to be impulsive or over-active due to the genetic selection of working breed lines (Müller et al., 2016). The DIAS is a psychometric scale intended to assess dogs' impulsivity levels from dog owners' ratings (Wright et al., 2011), validated with delayed reward tasks and physiological parameters (Wright et al., 2012). The DIAS includes 18 items to evaluate dogs' underlying cognitive impulsivity in different situations. The questionnaire statements such as "My dog shows extreme physical signs when excited" and "My dog reacts very quickly" are rated on a five-point rating scale from "Strongly disagree" to "Strongly agree". DIAS assesses three main factors: 'Behavioural Regulation', 'Energy and interest' and 'Aggression and Response to Novelty' and generates an overall score. This psychometric tool potentially provides a broader picture of dogs' impulsivity than specific tasks (Wright et al., 2012) and has shown to be stable over time (Riemer et al., 2014a). Studies assessing impulsivity in pet dogs have shown higher DIAS scores linked with behavioural problems, lack of adaptability and trainability (Wright et al., 2012; Riemer et al., 2014a). However, in a police dog study, active dogs' DIAS scores in 'Aggression and Response to Novelty' and 'Energy and interest' were significantly higher than those withdrawn early (Brady et al., 2018). There is limited research on the association of dog DIAS scores with working performance, and there is none in MDD.

In this study, the dog handlers were asked to complete the DIAS to measure the association of dog impulsivity scores with their behaviour in the test, and in subsequent chapters, with their task performance.

4.1.6. Chapter aims

This Chapter aims to explore the following:

a. The development and implementation of a test battery to quantify dog attributes considered relevant for medical detection (2.3.1.1) in an MDD sample composed of prospective and operational dogs to assess behavioural variation from their training stage.

b. Which variables measured show variation and which are associated with each other.

c) How these varied with the dogs' demography: their sex, age and whether they were trainee or trained when tested.

d) Whether the dogs' behaviour when tested was associated with the dog trainers' opinion on the dogs' performance (3.4.1).

e) The association of the dogs' DIAS scores, completed by the trainers during the test, with the dogs' test performance.

4.2. Methods

4.2.1. Ethical statement

The current research had ethical approval from the University of Bristol (Ref UB/19/05) and accorded to the charity Medical Detection Dogs standards.

4.2.2. Test battery development

The test battery included 18 subtests involving temperament and cognitive measures since previous research on prospective guide dogs suggests that both are deeply interconnected (Bray et al., 2017b). Combining them may improve the test's effectiveness. The test battery design relied on an extensive literature review of working dogs' personality and evaluation methods (Chapter 1). I considered the 27 most essential traits derived from the survey (2.4.7), aiming to measure the level at which these varied within the dog sample.

After reviewing several validated dogs' tests (1.2.6), I selected and adapted those deemed more suitable and feasible to evaluate these traits. The temperament subtests were based on military detection dogs' studies by Svartberg (2002) and (Rooney et al., 2003a) and similar to those in guide dogs by Batt et al. (2008), Harvey et al. (2016) and Bray et al. (2017b). These assessed aspects of the dogs' reaction to different stimuli, sociability to unfamiliar people, attachment to the handler, responsiveness to obedience cues, body sensitivity, playfulness, reactivity in a holding pen and specific tasks assessing detection ability. Cognitive measures were mainly adapted from MacLean et al. (2017), Bray et al. (2014); Bray et al. (2017b) and Tiira et al. (2020). Some subtests were originally from Hare et al. (1998) and Miklösi et al. (1998). These tasks measured reward motivation, social referencing, problem-solving, persistence, independence and inhibitory control.

The dogs' test performance was measured with scoring scales of state and point events and behaviour ratings, as the combination of subjective and objective variables increases behavioural predictability (Sinn et al., 2010).

I piloted the preliminary test on ten teaching dogs from the University of Bristol staff, with volunteers handling the dogs and made feasibility adjustments. I further rehearsed the test battery in the charity with three staff-owned dogs, which were not to be part of the experimental sample. After final modifications from piloting, I implemented the test battery with the experimental sample.

4.2.3. Experimental Sample

I tested the same 58 dogs assessed in Chapter 3 (3.2.2; Appendix 4), including 39 trainee dogs, of which 22 were intended for bio-detection tasks and 17 for assistance roles. The remaining 19 were trained dogs performing different bio-detection tasks. I ensured that the dogs were tested when they were at least 12 months old (except for one dog that was 11 months).

4.2.4. Experimental setting

The test battery took place in an experimental room in the Bio-detection wing of the Medical Detection Dogs® charity training facility (Figure 4.1). The test was video recorded with three cameras and a Swann® CCTV system.

4.2.4.1. *Experimental room description*

The room measured 6.80m x 3.7m, was partially carpeted and had a smaller tiled section (2.3m²). The room entrance was in the middle; a table was at the tiled end. A 3.5m wooden screen with three panels was at the opposite end, later utilised for the CBT (6.2.1). There were two windows in the wall opposite the room entrance, and each window shelf had a closed plastic container with the materials for the different subtests to ease its access. A Canon® video camera (primary camera) recorded the main view, and a GoPro Hero 4 ® camera captured the lateral angle for some subtests and filmed those happening outside the experimental room. A second one filmed the back view. The overhead Swann® CCTV system had a camera mounted to the ceiling in the middle of the room for backup recording. Another one pointed to the cognitive bias screen to monitor the dogs during that test (6.2.1). The room temperature was set between 21-22 Cº to avoid discomfort to the dogs from weather fluctuations. A bowl with fresh water was available all the time, refilled during the break.

4.2.4.2. Room's configuration at the test onset

At the start of the test, there was a chair by the entrance and another on the opposite wall. The chair positions and the materials were changed for subsequent subtests.

There were marks drawn on the carpet to aid the start location and the correct object positioning for each subtest. On the 'Holding area' in the tiled section, we indicated a start line with duct tape, and for several subtests, the handler held the dog behind the mark before releasing them for the task. All subtests were conducted in this room, except the last two: the S17 'Slippery surface test' (4.2.5.2.9), performed in a hall adjacent to the experimental room (to facilitate logistics) and the S18 'Pen holding test' (4.2.5.2.10), in the dogs' resting area (which contained four dog floor pens) in a different building or in the place where the dog was most frequently rested.

All the subtest lengths were measured with a stopwatch. For most subtests requiring a reward, I utilised Royal Canin Snacks Educ®. However, for the S4 Reward preference (4.2.5.1.4), S7 Cylinder task (4.2.5.1.7) and S11 Unsolvable task (4.2.5.2.3), I used Royal Canin Energy® as these needed higher value treats. For the S9 Box search (4.2.5.2.1) and S10 Noise distraction (4.2.5.2.2), I used Fish4Dogs Sea Jerky Squares® as they are larger and easier to notice when the dogs found them in the boxes.

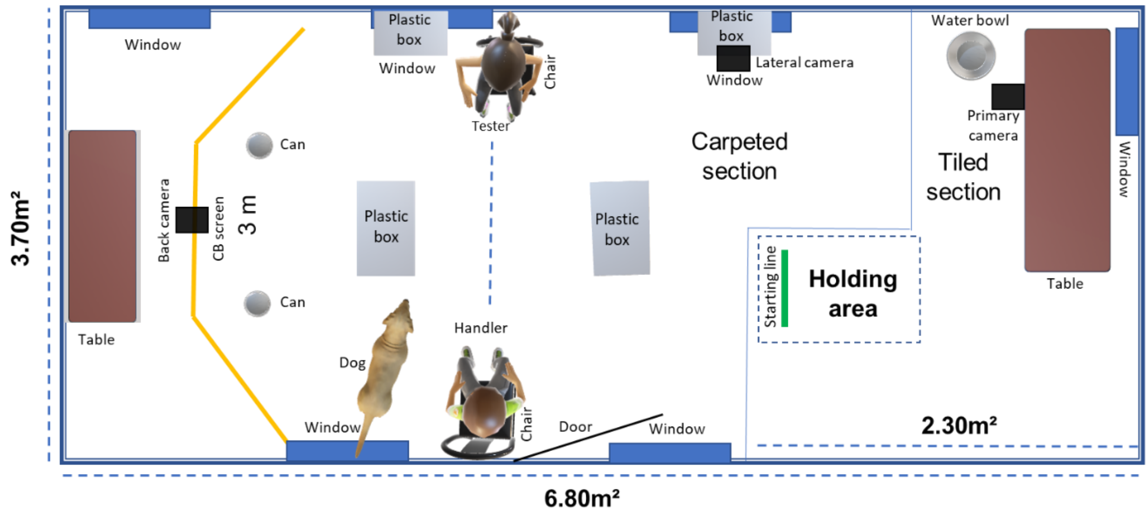


Figure 4. 1 *Diagram of experimental room with the test battery initial arrangement*

4.2.5. The test battery

I conducted the test battery between 09/09/2019 and 10/03/2020, one to two days a week. The test schedule was carefully coordinated considering the dogs' availability since they did not live on-site and the trainers' convenience to limit disruption of their daily work.

The 18 subtests were divided into two parts with a ten-minute break in between. The first involved the first eight subtests with an approximate duration of 30 minutes (4.2.5.1). The second included subtests 9-18 and lasted around 45 minutes (4.2.5.2; Figure 4.2).

I tested two dogs sequentially each morning, and they underwent the CBT (Chapter 6) in the afternoon. They were tested in the same task order ensuring a minimum of 120 minutes (Maximum =132, Mean=123 SD=1.6) between the end of the test battery and the start of the cognitive bias test. I established the test order considering the tests' content, logistics and feasibility. I attempted to leave the potentially most stress-eliciting subtests (i.e. 'Coat wearing', 'Vet check-up', 'Novel object', 'Slippery surface' and 'Holding pen' subtests) to the end to prevent dogs' cumulative -stress levels affecting their performance during multiple subtests. I conducted all tests (tester), and the dogs' handler (main trainer or assistant trainer) accompanied them throughout. I avoided interacting with any of the dogs before they were tested to ensure being unfamiliar at the outset of the test.

Before the test, I explained to the handler their role in the test battery with a diagram with instructions I had designed (Appendix 8). I also advised that if they noticed the dog responding negatively to any task, they could request a break or withdraw them (one dog was removed from the 'Unsolvable task' as it was scared when the container bounced; another from the 'Cylinder task' as it showed an adverse reaction to the device and the same dog did not perform the 'Novel object task' as the trainer thought the dog might react negatively).

The following section describes the 18 subtests (S=subtest). I used an ethogram based on the studies mentioned previously to measure the dogs' responses during the test (Appendix 9). The variables measured are detailed in Tables 4.1 and 4.2, and Figure 4.2 contains pictures of each subtest.

Table 4. 1 *Test battery Part 1: variables definition and rationale for inclusion in PCA (N=58): Variables excluded those where dogs had equal or less than 20% variation or those highly correlated (the most meaningful variable of the pair remained).*

Subtest, traits assessed and variables eliminated	*Variable*	*Description*	*Low variation ≥ 20%(LV) Min 20% or max 80% dogs*	*Correlated with another variable ≥ 0.70 (COR)*	*Included for data reduction (DR) Description of eliminated variables*

S1 Exploring 'Tendency to investigate humans by sniffing' 'Tendency to explore areas by sniffing' 'Friendliness to new people' 1 Eliminated	1. Duration exploring environment	Time investigating novel room within 1 min (Nose <=10 cm from object, floor or walls)	No	No	DR
	2. Duration investigating tester	Time sniffing experimenter within 1 min (Nose<= 10cm from experimenter)	No	Yes	DR COR with 4 More biologically meaningful
	3. Latency to approach tester	Time taken to approach experimenter within 1 min	No	No	DR
	4. Duration interacting with tester	Time making contact or < 50cm, from and paying attention to experimenter within 1 min.	No	Yes	Removed COR with 2 (R=0.993**)
	5. Duration interacting with handler	Time making contact or <50 cm, from and paying attention to trainer within 1 min.	No	No	DR
S2 Ignoring *'Attachment to human partner'* 1 Eliminated	1. Duration interacting with handler	Time making contact or <50cm, from and paying attention to trainer within 1 min.	No	No	DR
	2. Handler contact frequency	Number of times in contact with their trainer within 1 min	No	No	DR
	3. Duration interacting with tester	Time making contact or < 50cm, from and paying attention to experimenter within 1 min.	No	Yes	DR COR with 4 More biologically meaningful
	4. Tester contact frequency	Number of times in contact with experimenter within 1 min.	No	Yes	Removed COR with 3 (R=0.775*)
S3 Following *'Attachment to human partner'*	1. Duration following handler	Time dog spends making contact or < 50cm, from and paying attention to their trainer within 1 min.	No	No	DR
S4 Reward preference 'Motivation to play with toys' 'Motivation to obtain food' 2 Eliminated	1. Frequency choosing food	Number of trials in which the dog chooses and eats food ahead of toy over six trials.	No	Yes	DR COR with 2 mutually exclusive Food used more frequently as a reward in training
	2. Frequency choosing toy	Number of trials in which the dog chooses and contacts toy ahead of food. Amongst six trials.	No	Yes	Removed COR with 1 (R=-0.949**)
	3. Food motivation level	Subjective rating of enthusiasm to reach the chosen reward: 1. Very low to 5. Very high.	Yes 88% max. value	No	Removed LV >80%
	4. Toy motivation level	Subjective rating of enthusiasm to reach the chosen reward: 1. Very low to 5. Very high.	No	No	DR
S5 Arm pointing 'Ability to solve problems' 'Independence' (Cooperation) 1 Eliminated	1. Occasions selecting pointed can	Number of trials that the subject's first approach (to <10cm) contact the can or indicate (as in scent discrimination trials) that the experimenter pointed to within six trials	No	Yes	DR COR with 2 More biologically meaningful Mutually exclusive
	2. Occasions selecting opposite can	Number of trials that the subject's first approach (to <10cm) or contact the can that the experimenter did not point to within six trials	No	Yes	Removed COR with 1 (R=-0.911**)

S6 Obedience	1. Total latency to obey "Sit" and "Lie down"	Sum of the latencies to perform the "Sit" and "Lie down" tasks each twice	No	No	DR
'Obedience to human command'	2. Duration Stay	Mean latency before moving after the "Stay" cue is given maximum 30 sec.	No	No	DR
S7 Cylinder: Inhibitory control 'Impulsivity'	1. Immediate detour frequency	Number of times the dog performs a detour to take out the food from one of the cylinder's edges without previously nosing, pushing or pawing the transparent cylinder (out of max six trials)	No	Yes	DR COR with 4 More biologically meaningful Mutually exclusive
'Willingness to try new behaviours	2. Mean latency to obtain food	Average time to extract the food from one of the cylinder's edges over six trials 1 min max.	No	No	DR
	3. Different behaviours presented	Total number of different actions presented when trying to access food	No	Yes	DR
1 Eliminated	4. Total behaviours presented	Number of behaviours presentedwhen trying to access food	No	Yes	Removed COR with 1 (R=-0.720**) and 3 (R= 0.763**)
S8 Puzzle	1. Number of treats recovered	Number of food items located and consumed (out of 20) within 3 minutes	No	No	DR
'Ability to solve problems'	2. Duration on task	Time engaged in locating food over 3 minutes (sec)	No	Yes	DR COR with 3, 4 and 5. More biologically meaningful Mutually exclusive
'Independence'	3. Latency to gaze at human	Time before looking at experimenter or handler within 3 minutes (sec)	No	Yes	Removed COR with 2 (R=-0.784**), 4 (R=0 .901**) and 5 (R=0.918**)
'Persistence'	4. Duration gazing human	Total time (within 3 min) looking at experimenter or handler	Yes 88% min. value	Yes	Removed LV >80% COR with and 2 (R=-0.854**), 3 (R=0.901**) and 5 (R=0.902**)
3 Eliminated	5. Duration not on task	Total time (within 3 min) not focused on finding treats	No	Yes	Removed COR with 3 (R=0.839**) and 4 (R=0.902**) COR with 2 (R=-0.918**)

***Table 4. 2** Test battery Part 2: variables definition and rationale for inclusion in PCA (Explanation in Table 4.1).*

Subtest, traits assessed and variables eliminated	*Variable*	*Description*	*Low variation ≥ 20%(LV)* *Min 20% or max 80% dogs*	*Correlated with another variable ≥ 0.70 (COR)*	*Included for data reduction (DR)* *Description of eliminated variables*
S9 Boxes search	1. Latency to find treats	Mean latency to find food item over three trials (1 min max each)	No	No	DR
'Acuity of sense of smell'	2. Number of treats located	Total food items found and consumed over three trials (maximum 3)	No	No	DR

'Tendency to remain specific to the target odour' 1 Eliminated	3. Time not on task	Mean percentage of the total time not focused on finding treats over three trials (1 min max each)	No	No	Removed Almost significantly COR with 1 (R=0.690) methodological issues confounding effect of length on trial
S10 Noise distraction *'Tendency to be distracted when working'* 1 Eliminated	1. Time attending to noise	Duration oriented and apparently focussed on the sound (max 30 sec)	Yes 87.9% min. value	No	Removed LV >80% Described but rarely occurred
S11 Unsolvable task	1. Duration on task	Time attempting to access food items (1 min max)	No	Yes	DR COR with 2,3, 5, 7 and 8. More biologically meaningful
'Ability to solve problems'	2. Duration gazing at human	Total duration dog looked towards the experimenter or trainer over 1 min.	No	Yes	Removed COR with 1 (R=-0.820**), 3 (R=0.938**), 4 (0.895**), 5 (R=-0.855**), 7 (R=-0.853**) and 8 (R=0.952**)
'Independence'	3. Duration gazing at tester	Time dog looked towards the experimenter over 1 min.	No	Yes	Removed COR with 1 (R=-0.773**), 2, (R=-0.862**), 7 (R= 0.825**) and 8 (R= 0.895**)
'Persistence'	4. Duration gazing at handler	Time dog looked towards trainer over 1 min.	No	Yes	Removed COR with 2 COR (-) with 1 R=(-0.773**), 6 (R=-0.862**) and 7 (R=0.825**)
'Willingness to try new behaviours'	5. Latency to gaze at tester	Time before dog looks towards tester maximum 1 min.	No	Yes	Removed COR), 1 (R=0.816**) 2, 3, 7 (R=0.972** and 8 (R=-0.862**)
	6. Latency to gaze at handler	Time before dog looks towards handler within 1 min.	No	Yes	DR COR 1, 4 and 8 Included as not COR with 1 More biologically meaningful
	7. Latency to gaze at human	Latency to gaze human within 1 min.	No	Yes	Removed COR with 1 (R=-0.849**), 2,3,5 and 8 (R=-0.863**)
	8. Activity alternation frequency	Number of changes from task to other activity and vice-versa	No	Yes	Removed COR with 1 (R=-0.832**),2, 3 ,4, 5 and 6
6 Eliminated	9. Number of different behaviours presented	Different behaviours presented during task	No	No	DR

S12 Ball searching 'Acuity of sense of smell' 'Tendency to remain specific to the target odour'	1. Total correct choices	Number of choices of the baited pot (first container they contact with snout or paw or indicate with trained alert 0= fail, 1= found with assistance from tester 2= found (scoring 0-6)	No	No	DR
Willingness to bring an object back to a person	2. Ball retrieval score	Number of occasions when the ball is retrieved to the experimenter 0= not retrieved, 1= retrieved with intervention (speaking to dog or slightly puling ball from mouth), 2= retrieved to handler or experimenter (scoring 0-12)	No	No	DR
S13 Toy playing 'Motivation to play with toys' 'General excitability'	1. Duration playing	Time interacting with toy over 60 seconds	Yes 81% Max value	No	Removed LV >80%
	2. Activity alternation frequency	Number of changes from playing with toy to other activity and vice-versa	Yes 81% min value	No	Removed LV >80%
	3. Play intensity	Subjective rating for involvement with tug of war game on a scale of 1 - very low to 5 - very high	No	No	DR
	4. Variety of behaviours presented	Number of different behaviour types presented when playing	No	No	DR
	5. Latency to drop	Mean time to drop toy after being instructed by the handler maximum 1 min	No	No	DR
	6. Total latency to obey "Sit" and "Lie down"	Sum of the latencies to perform the "Sit" and "Lie down" tasks once	No	No	DR
2 Eliminated	7. Duration Stay	Latency before moving after the "Stay" cue is given maximum 30 sec.	No	No	DR
S14 Coat wearing 'Body sensitivity'	1. Level of restraint required	Subjective rating for the level of restraint required when putting coat on 1 very low-Staying still, relaxed o 5 very high-Does not tolerate	No	No	DR
	2. Biting coat frequency	Number of times dog places coat in their mouth within 30 sec	Yes 89.7% min. value	No	Removed LV >80%
1 Eliminated	3. Frequency of approaches to human	The dog makes contact or < 50cm from trainer or experimenter	No	No	DR
S15 Body condition check-up 'Body sensitivity' 'Confidence'	1. Level of tolerance of check-up	Subjective rating for the level of dog's acceptance of a veterinary check-up with a scale of 1 to 5: 1 very low - Tries to escape to 5 very high - Stays still, relaxed	No	No	DR
S16 Novel object	1. Level of approach towards new object	Subjective ratings of dog's response to moving ball object: 1 Very low/retreat to 5 to Very high/Approaches to object	No	No	DR
'Confidence'	2. Number of different behaviours towards new object	Total variety of behaviours displayed when the dog is	No	No	DR

		presented with the moving object			
	3. Vocalisations frequency	Total vocalisations during test	Yes 96.5% Min value	No	Removed LV >80%: rarely occurred
	4. Frequency of approach to tester	Number of times the dog approaches experimenter, making contact or < 50cm	No	No	DR
2 Eliminated	5. Frequency of approach to handler	Number of times the dog approaches handler, making contact or < 50cm from them	Yes 91.1% Min value	No	Removed Low variation >80%: rarely occurred
S17 Slippery surface	1. Latency to step onto surface	Time taken to place all four paws on the slippery surface maximum 30 sec.	No	No	DR
	2. Duration on slippery surface	Total time with four paws n on novel surface maximum 30 sec.	No	No	DR
1 Eliminated	3. Readiness of approach to slippery surface	Subjective rating of dog's response to a novel surface. Scale: 1 Very low-does not step on surface to 5 - Readily -jumps straight onto surface	Yes 80.7% Max value	No	Removed LV >80%
S18 Holding pen	1. Latency to rest	Latency to lie down	No	No	DR
	2. Duration resting	Total time lying down	No	Yes	DR COR with 5 More biologically meaningful Mutually exclusive
	3. Duration moving	Total time walking in pen within 5 minutes	No	No	DR
	4.Duration exploring	Total time investigating environment within 1 min (nose <10cm from object, floor or walls/ fences)	No	No	DR
	5. Duration standing	Total time standing in pen within 5 minutes	No	Yes	Removed COR with 1 (R=-0.930**) Mutually exclusive
	6. Pen gate approach frequency	Number of times the dog approaches from within 20cm or less of the holding pen gate within 5 minutes	No	No	DR
	7. Pawing on pen or enclosure area	Number of times the dog places their paw on the gate or walls/ fences within 5 minutes	Yes 80.4% Min value	No	Removed LV, 20% showing the behaviour
	8. Jumping up frequency	Number of times when the dog jumps or climbs up supported on its back paws within 5 minutes	No	No	DR
	9. Duration vocalising	Total time vocalising whilst in holding in pen within 5 minutes	No	Yes	DR COR with 10 More biologically meaningful
3 Eliminated	10. Vocalisation frequency	Total number of vocalisations made whilst in holding pen within 5 minutes	No	Yes	Removed COR with 9 (R=0.906**)

4.2.5.1 Test battery subtests Part 1

4.2.5.1.1. S1. Exploring

Adapted from Rooney et al. (2003a), this subtest assessed the dog's tendency to investigate the environment as compared to an unfamiliar person and functioned to acclimatise the dog to the experimental room.

The room was prepared with objects new to the dog and used in further subtests: two large plastic containers L 58.2 x W 38 x H 18.8cm placed upside down perpendicularly across the central axis of the room with 75cm between them and two metallic cans H 15 cm x W 7.5cm with a 2cm hole on the top, placed parallel to the wooden screen 50 cm from it and with 180 cm between them.

The handler entered with the dog on a leash and sat in a chair adjacent to the door. The tester was seated in the further chair, 2m from the handler. The dog was unleashed and allowed to freely explore the room and objects for two minutes while the handler answered a brief oral questionnaire about the dog's general details (Appendix. 10). The tester did not encourage them to approach, but she would talk and pet them gently if they did.

I measured duration exploring environment, duration investigating the tester and the handler correspondingly and latency to approach to the tester following release.

4.2.5.1.2. S2. Ignoring

Based on Rooney et al. (2003a), this subtest assessed the dog’s propensity to demand the handler’s attention when ignored.

The handler and experimenter sat in the room as in the previous test, and the dog roamed freely. The handler was asked to complete the written DIAS questionnaire (Wright et al., 2011) (4.1.5) and given three minutes to do so (a printed copy of the questionnaire was inside a folder on the chair before the handler entered the room). Meanwhile, both the handler and the experimenter ignored the dog. The handler attended to and praised the dog when the subtest ended.

I measured the handler and tester contact frequency and the duration interacting with the handler and the tester.

4.2.5.1.3. S3. Following

Adapted from Rooney et al. (2003a), this subtest assessed the dog's predisposition to follow their handler when moving.

The handler was initially in their seat and was asked to walk normally around the room for one minute while ignoring the dog. Afterwards, the handler stopped moving and paid attention to the dog.

I measured the dog's duration following their handler.

4.2.5.1.4. S4. Reward preference

Based on MacLean et al. (2017), this subtest investigated the preferred reward for each dog between food and toys and their level of motivation to obtain each.

The handler sat in the holding area while keeping the dog behind the start line. There were two familiarisation trials. The experimenter approached the dog holding a 40cm stuffed fox toy (Spot Skinneeez extreme®), calling their name and saying "Look", allowing the dog to smell it. Then the tester took six steps backwards and placed the toy on the floor 3.5m from the dog in a central position.

When the tester said "Ready", the handler released the dog with a "Go" cue to get the toy. In a second familiarisation trial, this was repeated with a treat in a small bowl. In the testing phase, the tester showed the dog a treat and a toy and placed them on the floor 180cm apart. The dog was released to choose a reward (eat the treat or make contact with the toy) within six pseudorandomised trials counterbalancing the locations of the items. Each reward was placed three times on each side in the same order for all the dogs to overcome any side bias. (The treat was twice on the right side, twice on the left, then one to the right and one to the left). The toy and type of treat were the same for all trials. If the dog chose food, it was allowed to eat it, and when selecting the toy, it could play with it briefly before the handler removed the toy.

I measured the frequency of choosing food and the frequency of choosing the toy over six trials. I also assessed the motivation to approach each reward on a subjective scale from 1 to 5.

4.2.5.1.5. S5. Arm pointing

Modified from Hare et al. (1998), Miklösi et al. (1998) and (MacLean et al., 2017). In this subtest, the dog was encouraged to find a reward hidden in a can, based on the tester's arm pointing.

The two metallic cans used in the 'Exploring subtest' (S1) were placed upside down, 3.5m away for the dog and separated by 1.80m. Initially, for familiarisation, the tester baited one can in a central position. The handler released the dog from the holding area, and when they made contact with the can, it was raised, and the dog was allowed to take the treat. In a second familiarisation trial, this was repeated but by baiting two cans while the dog watched.

For the test trials, the tester baited the containers with a treat while saying the dog's name and "Look", covering the cans with a cardboard view blocker (37 x 2cm) when baiting them (so the dog couldn't see which was baited). The view blocker was then removed, and the tester extended her arm, pointing to the baited container with the index finger. The dog was released to choose a can, and the tester raised the chosen can, allowing the dog to take the treat. This was repeated six times with a pseudorandomised balance (three times for each side when pointing to the cans as in the reward preference test).

I recorded the number of trials when the dog selected the pointed can.

4.2.5.1.6. S6. Obedience

Adapted from Rooney et al. (2003a), this subtest assessed the dog's ability to complete three previously learned obedience tasks on cue.

The handler stood in the centre of the room facing the dog and was asked to consecutively instruct the dog to "Sit", "Lie down", and "Stay". For the "Stay" task, the handler walked three steps backwards, facing the dog and waited for up to 30 seconds (timed by the tester with the stopwatch) before calling the dog. The handler repeated this series of exercises and rewarded the dog with a treat after the second time.

I measured the total latency to obey "Sit" and "Lie down" tasks after the first command, and the mean duration holding the "Stay" task without moving summed across both commands and both trials.

4.2.5.1.7. S7. Cylinder: Inhibitory control

Based on Bray et al. (2014), MacLean et al. (2017), and Tiira et al. (2020), this subtest assessed the dog's ability to inhibit a motor response when facing a treat in a transparent cylindrical container and perform a detour to obtain it.

The tester placed a transparent hollow acrylic cylinder 20cm in diameter and 28 cm long on a wooden base in the centre of the room. Initially, the training phase consisted of six trials where the cylinder was covered with black fabric secured with four Velcro® stickers. The tester faced the dog held behind the start line, called the dog’s name, said "Look," and placed a treat in the cylinder. The dog was released and allowed to retrieve the treat out of the cylinder for up to one minute.

For the testing phase, the cover was removed from the cylinder, and the test was repeated six times. The side where the treat was inserted was counterbalanced for each trial.

I measured the number of times the dog performed a detour to take out the food from one of the cylinder's edges without previously nosing, pushing or pawing the transparent cylinder out of a maximum of six trials; the mean latency to obtain food from one of the cylinder's edges across trials and the total number of different behaviours presented when trying to access food.

4.2.5.1.8. S8. Puzzle

Based on Bray et al. (2017b), this subtest assessed the dog's ability to solve a three-step puzzle and the tendency to persist independently or rely on human aid to solve it.

The tester baited a dog puzzle (Outward Hound OttonsonPuzzle Brick®) with 20 treats. The puzzle was composed of four rows with three sorts of compartments. Each row had two compartments on each side with flip lids that could be opened to retrieve hidden treats and two more where the lid had to be slid to uncover an additional hiding spot under each. A middle compartment was covered by a brick that needed to be removed to retrieve the treat inside and reveal the lateral compartments. A treat was introduced in each compartment with a total of 20, and the dog was allowed to extract them. After calling the dog's name and saying "Look", the tester left the puzzle in the centre of the room and returned to her seat. Then the dog was released and allowed to get the food out of the puzzle for three minutes.

I measured the number of treats recovered, the dog's duration on task and the latency and duration within three minutes gazing at the human.

4.2.5.2. Test battery subtests Part 2

4.2.5.2.1. S9. Boxes searching

Modified from Svartberg (2002) and Rooney et al. (2003a); This subtest explored the dogs' ability to find a hidden food item with their sense of smell.

Four transparent plastic storage containers (like those used in S1) filled with craft paper shreds were placed around the central part of the room in a semi-circular array 25cm from each other. The tester faced the dog, which was held by the handler behind the start line and baited one of the boxes with a Fish4Dogs Sea Jerky Square® and sham baited the rest of the boxes (so the dog was not able to see which box was baited) while saying "Where did it go?" to gain the dog's attention.

The tester returned to her seat, and the dog was released to seek the food treat for one minute. If the dog found it sooner, the trainer praised them and the trial was ended; this was repeated three times.

We measured the mean latency to find each treat over three trials, the duration of the non-searching activity and the total number of treats located (maximum three).

4.2.5.2.2. S10. Noise distraction

This subtest was based on previous assessments where dogs were submitted to a distracting or startling noise, assessing their reaction and latency to recover (e.g. Svartberg, 2002; Batt et al., 2008). This task assessed the dogs' ability to concentrate on a searching task in the presence of distracting noise.

The previous task (S9) was repeated. This time, the tester played a recording of a standardised unfamiliar sound (Geiger Counter for radiation) at 75dB while the dog was seeking the treat. I chose what was likely to be an unfamiliar sound to minimise the likelihood of dogs' prior associations and potentially not fearful. The noise sounded after 10 seconds from the trial's onset so the dog could start searching and lasted five seconds. It was played during the one minute of searching three subsequent times with randomised intervals. I measured the total time attending to the noise.

4.2.5.2.3. S11. Unsolvable task

Adapted from Miklosi et al. (2003) and MacLean et al. (2017), this subtest aimed to evaluate the dog's tendency to independently persist in obtaining a food item that remains out of reach and to look for aid from a human. First, for familiarisation, the tester baited a transparent plastic container (25.2 X 17.6 X 9 cm) with a Royal Canin Energy® higher value treat.

While the handler kept the dog in the holding area, the tester approached the dog, called the dog's name and said, "Look", showing the treat inside the container and allowing the dog to smell it. Then the tester placed the container in the centre of the room with the lid slightly on but not fully closed so the dog could easily access the treat. After being released, the dog had one minute to retrieve the treat

The procedure was repeated, but with the container closed entirely. The dog could attempt to get the food out of the container for one minute, although this was impossible.

I measured the duration attempting to access the food, the latency and duration gazing at the handler or tester, the number of changes from task to other activities and the number of different behaviours presented.

4.2.5.2.4. S12. Ball searching

This was an abbreviated version of an in-house dog test from the Medical Detection Dogs® charity (3.6.1), it assessed the dog's ability to locate a hidden ball by its scent.

In the central area of the room, the tester lined up three metallic cans upside down (the same used in S5) with 25cm between them; the cans had plastic ashtrays under them to avoid scent contamination on the floor and were placed on a transparent plastic mat 76.2 x 121cm. Each can was discretely labelled with a ½ cm white sticker on the top marked with a number in black pen to enable the tester to distinguish the can where the ball was hidden (target can), which was numbered 1, and the other two, 2 and 3, so the difference was deemed unnoticed to the dog. The tester manipulated the target can and the ball with a disposable plastic glove to avoid contaminating the rest of the cans with the ball scent.

The handler sat on the floor, holding the dog approximately 50cm from the cans. Initially, the tester threw a tennis ball three times, and the dog was encouraged to retrieve it. For the testing trials, the tester hid the ball under the target can and slid the three cans crossing and mixing them, following a set pattern for each of the six trials, aiming for the dog to lose sight of the target can and rely on their sense of smell to find the ball. The plastic mat and ashtrays allowed the cans to slide smoothly.

The tester said, "Where is it?" to gain the dog's attention. After the tester moved the cans, the dog was released with the “Ready” cue and encouraged to find the ball. If the dog located the ball, the tester verbally praised them and threw it for retrieval. If not, she said, "Wrong," and the handler encouraged the dog to keep searching until they found the ball.

I measured the number of total correct choices the first time within six trials and scored the dog's willingness to retrieve and release the ball without resistance.

4.2.5.2.5. S13. Toy playing

Based on Rooney et al. (2003a), this subtest assessed the dog's tendency to get involved in interspecific social play with a toy.

The handler encouraged the dog to play tug of war with a stuffed fox toy (used in S4) for one minute in the centre of the room while the tester timed from her chair. After the time was up, the handler asked the dog to leave the toy and then asked them to sit, lie down and stay still for 30 seconds as in S6, but only once.

I measured play duration, the intensity of involvement with the game on a subjective scale, latency to drop the toy and the total time spent performing the "Sit" and "Lie down" tasks and the latency to move after the "Stay" command. I recorded six behaviours dogs frequently perform when playing, selected from a previous study listed in Table 4.1.

4.2.5.2.6. S14. Coat wearing

This subtest measured the dog's tolerance for wearing a service coat. The tester gave the handler a service fabric coat provided by the charity (of an appropriate size for the individual dog). The handler put the coat on the dog in the centre of the room while the tester was seated. Then the handler went back to their seat behind the start line. The dog was allowed to roam freely in the room for 30 seconds, timed by the tester, and the handler was asked to interact with the dog normally if it approached but not to excite it deliberately

I measured the level of restraint required when putting the coat on, the duration wearing the coat without struggling and the number of times biting the coat. I also recorded the frequency of ear height, tail height score, tail wagging and the presentation of any of 15 stress behaviours (4.5.4) whilst the coat was put on and five seconds after its placement.

4.2.5.2.7. S15. Body condition check-up

Modified from Harvey et al. (2016), this subtest measured the dog's tolerance to a body condition check-up. The handler held the dog in the centre of the room while the tester, crouched on the floor, handled the dog around the head and body, similar to a vet's simple body examination while giving the dog treats. She also examined the dog with a stethoscope.

I measured the tolerance level on a subjective scale and the dogs' ear height, tail height and tail wagging five seconds after the start of the check-up and when the stethoscope was placed on the dog's chest.

4.2.5.2.8. S16. Novel object

Based on Rooney et al. (2003a), this subtest evaluated the dog's tendency to interact with or avoid an unfamiliar object in motion.

The tester stood in the centre of the room, took a motorised toy that jumps and makes noise (Bumble ball®) out from a fabric bag, activated it, and placed it 1.5m from the dog held in the holding area. The handler released the dog when the toy was moving on the floor to explore the toy freely but without encouragement. When the toy stopped moving after approximately 20 seconds, the tester retrieved it and returned it to the bag.

I measured the dog's tendency to approach the object. I also recorded ear height, tail height and the presence of tail wagging five seconds after placing the novel object on the floor

4.2.5.2.9. S17. Slippery surface

Adapted from Rooney et al. (2003a), this subtest assessed the dog's tendency to step onto an unfamiliar surface to access food rewards.

The slippery surface was provided by a sheet of polished wood 190 x 110cm and an adjacent smooth acrylic one 120 x 110cm. These were placed together to make a large surface for the dogs to walk over it and increase the novelty factor by using two different materials. The surface was located outside the experimental room. The tester exited the room, followed by the handler holding the dog on a leash. The handler had the dog 50cm from the surface while the tester scattered 15 treats over it. After the “Ready” signal, the dog was unleashed to explore the surface and get the treats whilst the tester filmed the auxiliary GoPro® camera and timed the dog.

I measured the dog’s latency to put all four paws onto the surface and the total duration with four paws on the surface over the next 30 seconds and rated the dog’s readiness to step onto the surface. We also recorded the ear height, tail height and tail wagging five seconds after release and the presence of any stress behaviours.

4.2.5.2.10. S18. Holding pen test

Modified from Rooney et al. (2003a); This subtest measured the dogs' ability to settle without human presence.

After the main test ended, the dog’s trainer placed them in a crate, pen, room or holding area where they regularly left them. The dog was filmed for five minutes without human company with the auxiliary GoPro camera on a clamp at an angle, allowing a full view of the dog's movement. After this period, some dogs were removed, while others remained as the handler decided. After a resting period of approximately 120 minutes, I proceeded with the CBT (Chapter 6).

I measured the dog’s latency to lie down and the duration of resting, moving, standing, exploring and vocalising over five minutes. I also measured vocalisation frequency, the number of occasions when the dog approached within 20cm of the pen door and the number of times the dog jumped up supported on its back feet within five minutes.

4.2.5.3. Stress behaviours

I quantified the frequency of each of 13 stress-related behaviours (Table 4.5) and their total occurrence over the test to get a better insight into the tests' welfare impact on the dogs, its association with performance in Chapter 5 and CBT outcomes in Chapter 6. Since the Holding pen test (S.18) was conducted separately, stress behaviours were not accounted for in this subtest.

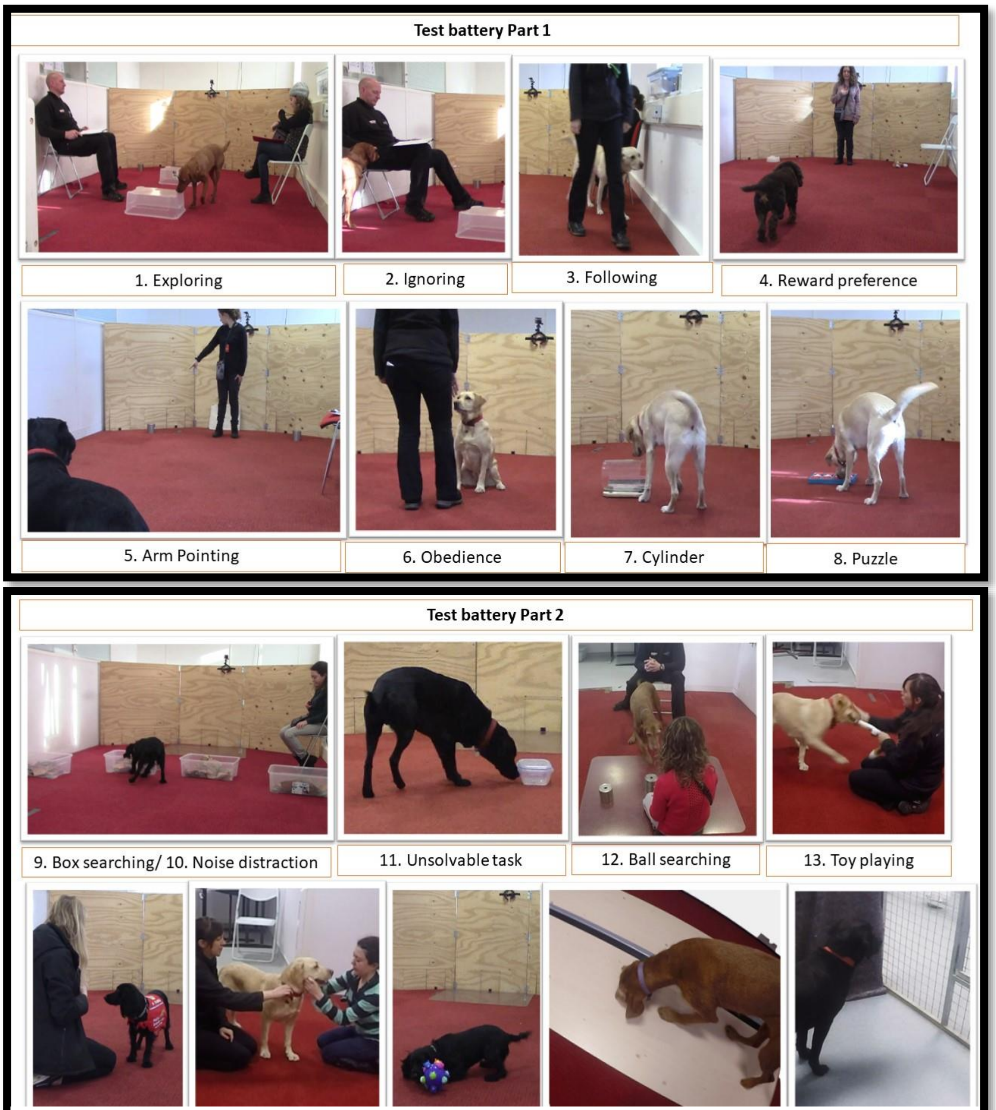


***Figure 4. 2** Subtests' pictures from the test battery in sequential order.*

4.3. Data analysis

4.3.1. Data reduction

4.3.1.1. Battery Test data extraction and initial data reduction

I decoded the video recordings from the battery test of the primary camera and the lateral Go Pro® with Boris® (Behavioural Observation Research Interactive Software) and extracted 170 variables. Variables were first cleaned to discard or condense anything redundant (e.g. durations that were added or averaged to produce total or mean), producing 96 variables for subsequent analysis with SPSS® software (Tables 4.1 and 4.2).

One variable, S14 'Restraint tolerance when putting the coat on', was reverse coded, so all subjective scores were positively oriented, e.g. less struggle, higher score (Field, 2018). We first inspected the variation of the data on histograms, and those variables where 80% or more of the dogs had the same score were rejected. This process led to exclude 12 variables, and the 'Noise distraction test' was eliminated. Then, I examined correlations between variables within each subtest. For all correlations $R>0.7$, one variable in the pair was excluded. I retained those more biologically relevant (i.e. the variable which best measured the construct), aiming for those showing more variation amongst dogs. I further eliminated 16 highly correlated variables (Tables 4.1 and 4.2).

I measured ear position, tail position, and tail wagging (Appendix 9) in four subtests, potentially assessing aspects associated with dogs' body sensitivity and confidence (S14-two time points, S15-two time points, S16-once, S17-once). I calculated averages for ear position, tail position and tail wagging to summarise them and the total number of stress behaviours exhibited across all subtests. This left a total of 49 variables for further data reduction with PCA.

4.3.1.2. Data reduction phase 2: Principal component analysis

We performed a PCA with Varimax rotation (based on the procedures suggested by Laerd Statistics, 2015, and Field, 2018) to summarise the remaining 49 variables into meaningful behavioural components and subsequently assessed their association with MDD performance measures.

The PCA initially included 49 variables from the initial data reduction. However, Kaiser-Meyer-Olkin Measure (KMO) for sampling adequacy showed low values (χ 2 (0.182) = 1321.3, $p<0.001$). KMO value ranges considered by Kaiser (1974), described in Chapter 3 (3.3.2.9), where values range from 0 to 1, but the minimum acceptable is ≤ 0.5. Here KMO was <0.5. Hence the number of variables was not suitable for the sample size and needed further reduction to increase the KMO value.

I excluded more variables to improve the data suitability for PCA, given the sample size, as suggested by Field (2013). Firstly, I manually introduced each of the 49 remaining variables to the PCA using a trial-and-error process. I entered each variable forwards and backwards, aiming to increase the general KMO and the individual KMO of each variable to over 0.5 (the minimum acceptable according to Kaiser (1974) with a statistically significant Bartlett's test of sphericity, and including the highest number of variables as possible). I retained at least one measure from each subtest (except ‘Body condition check-up’ since ‘Level of tolerance to body condition check-up’ decreased the KMO). There were variables potentially measuring similar constructs in different subtests (e.g. variables related to time near the handler were measured in three tests and obedience measures in two), where I kept only the one that had increased the KMO.

Ultimately, 22 variables were removed, leaving 27 variables which jointly achieved an acceptable KMO and a significant Bartlett's test of sphericity (χ 2 (0.513) = 424.4, $p = 0.004$). I retained the components with eigenvalues over one and extracted their scores for further analysis.

4.3.1.3. Stress behaviours

The total stress behaviours were not included in the final PCA since this variable did not contribute to improving the KMO. However, it was assessed separately for statistical analysis.

4.3.2. Relationship of dogs' demography with their behaviour in the test battery

I performed logistic models to investigate associations between dogs' measures from the test and demography to reduce multiple comparisons (Personal communication with Knowles, 2020) instead of correlations or models to study each component as a dependent variable. I focussed on the main effect of each independent variable on the dependent variable but not on interactions between independent variables. I considered associations, but not causality, between the independent and the dependent variable. The independent variables were introduced in the models with stepwise methods and then manually to include in the final model the most significant or close to significant independent variables adequate to sample size (as in Chapter 3.3.2.5). Only significant associations were discussed.

I performed binary logistic regressions with the 11 component scores and total stress as independent variables and sex and training stage as the dependent variable, respectively. I also assessed if the dogs' behaviour varied with their age when tested with multiple regression, including as independent variables the components and total stress and dogs' age as the dependent one. In addition, I conducted Mann-Whitney U tests to assess if the dogs' age significantly varied with the training stage.

4.3.3. Association between test battery's behavioural components and DIAS scores

The DIAS handwritten questionnaire's ratings were captured in Microsoft Excel® and computed scores for the three different factors mentioned in Sec. 4.1.5. (following the formulas from (Wright et al., 2011) and the overall score (a higher overall score indicates higher impulsivity). I decided to use only the DIAS Overall score as I considered it the most representative measure to assess if the dogs' impulsivity levels varied with the dogs' behaviour in the test battery.

I assessed the link between dogs' impulsivity scores and their behavioural attributes through multiple regression, including as independent variables: sex, age, training stage, the components and total stress behaviours from the test battery, and the dogs' DIAS scores as the dependent variable.

4.3.4. Association of the dogs' behaviour on the test battery with their trainers' trait ratings

To investigate if the trainers' opinion of the dog's performance concurred with the behavioural measures from the test, I performed a Pearson's correlation between the 11 PCA behavioural components from the test battery and the five PCA components clustering the trainers' trait ratings (3.4.2.8).

4.4. Results

4.4.1. Principal components from the test battery

The PCA with orthogonal Varimax rotation produced 11 components with eigenvalues over 1. The 11 components jointly accounted for 74.5% of the total variance. The rotated solution with the component loadings, the percentage of variance explained by each component and their commonalities are shown in Tables 4.3 and 4.4 (components divided into two tables).

I labelled each component according to the combination of variables clustered in each and their major loadings to facilitate future reference (Subtest numbers are in brackets).

Component 1 'Playfulness'. Dogs with higher scores showed higher toy motivation when choosing toys compared to food (S5), play intensity in the tug of war game (S13), mean tail positions and a greater variety of behaviours in the Unsolvable Task (S11).

Component 2 'Persistence'. Dogs that scored higher took longer to gaze at their handler during the Unsolvable task, spent more time engaged with it and attempted a greater variety of behaviours to take the treat out of the container (S11) and remove the treat from the cylinder faster (S8).

Component 3 'Reactivity in holding pen'. Dogs with higher scores spent more time vocalising and moving inside the pen; they rested less and more frequently approached the pen gate (S18). They also chose the correct signalled can less often than during the Arm pointing test (S6).

Component 4 'Food orientation'. Dogs scoring higher spent longer eating treats on the slippery surface (S17), chose food more frequently than toys in the reward preference test (S5), had a higher ear average position and retrieved the treat more quickly from the cylinder (S8).

Component 5 'Obedience'. Dogs with higher "Obedience" scores obeyed the "Sit" and "Down" commands faster and stayed lying longer in the "Stay" part of the Obedience test (S7).

Component 6 'Level of attention to handler'. Dogs that scored higher spent longer interacting with their handler in the Ignoring test (S2); they dedicated less time to investigating the tester in the Exploring test (S1) and had a higher average ear position.

Component 7 'Self-control'. Dogs with higher scores performed an immediate detour more frequently and were more likely to take the treat out of the cylinder faster in the Cylinder test (S8) and found the ball more frequently in the ball searching test (S12).

Component 8 'Confidence'. Dogs scoring higher displayed more behaviours towards the novel object (S16) and required less restraint when putting the coat on (S14).

Component 9 'Success at problem-solving'. Dogs with higher scores took out more treats from the puzzle (S9), rested less in the holding pen (S18) and displayed a greater variety of behaviours' to obtain the treat in the Unsolvable Task (S11).

Component 10 'Interest in exploring environment'. Dogs with higher scores spent longer exploring the experimental room in the Exploring test (S1).

Component 11 'Success in search'. Dogs that scored higher took less time to find the treats in the Boxes search task (S10) and spent less time vocalising and moving in the holding pen (S18).

Table 4. 3 *Components 1-6 and variables with loadings contributing to each component (N=58). Structure matrix of PCA with Varimax rotation, main loadings ≥ 0.4 bolded. Eigen values >1. subtest number on the left side of each variable. S= Subtest.*

Variable	*1.Playfulness*	*2.Persistence*	*3.Reactivity in holding pen*	*4. Food orientation*	*5.Obedience*	*6. Level of attention to handler*
S5 Toy motivation level	**0.839**	0.141	0.104	-0.017	-0.081	-0.219
S13 Play intensity	**0.792**	0.131	-0.015	0.005	0.033	0.126
Tail average position	**0.467**	0.234	-0.073	**-0.424**	-0.156	0.115
S16 Level of approach towards novel object	0.394	0.136	-0.097	-0.204	-0.181	0.313
S11 Latency to gaze at handler	0.143	**0.863**	-0.064	0.075	0.010	-0.009
S11 Time on task	0.106	**0.829**	0.039	-0.007	-0.109	-0.112
S11 Different behaviours presented	**0.422**	**0.623**	0.012	-0.028	-0.014	0.068
S18 Duration vocalising	0.105	0.071	**0.750**	0.161	-0.159	0.141
S18 Pen gate approach	0.300	-0.019	**0.675**	-0.329	0.164	0.037
S18 Duration moving	-0.041	0.065	**0.663**	0.023	0.160	-0.267
S6 Occasions selecting pointed can	0.133	0.064	**-0.660**	-0.091	-0.083	-0.069
Duration on slippery surface	0.016	0.085	0.195	**0.754**	-0.208	-0.089
S17 Frequency choosing food	-0.216	-0.040	-0.094	**0.659**	0.076	0.013
Ears average height	0.299	0.130	-0.108	**0.538**	0.154	**0.517**
S7 Latency to obey "Sit" and "Lie down"	0.009	0.118	-0.052	0.040	**-0.861**	-0.003
S7 Duration Stay	-0.115	0.039	0.028	-0.002	**0.623**	0.055
S3 Duration following handler	0.281	0.043	0.261	0.146	-0.394	0.285
S2 Duration interacting with handler	-0.079	-0.141	0.145	-0.065	0.004	**0.855**

S1 Duration investigating tester	0.075	-0.396	-0.025	0.125	0.252	**-0.434**
S8 Immediate detour frequency	0.025	0.057	-0.037	0.181	0.005	-0.076
S8 Mean latency to obtain food	0.271	**-0.371**	-0.234	0.096	0.169	-0.094
S12 Total correct choices	-0.252	0.030	-0.016	**-0.383**	0.095	0.009
S16 Number of different behaviours	0.133	0.120	-0.013	-0.252	-0.065	0.108
S14 Level of restraint required	0.001	0.190	-0.136	0.061	-0.157	0.020
S9 Number of treats recovered	0.173	0.017	-0.209	0.105	-0.017	-0.130
S18 Duration resting	0.155	0.267	**-0.507**	0.233	0.215	-0.160
S1 Duration investigating environment	0.150	0.056	0.087	0.101	0.005	-0.071
S9 Latency to find treats	-0.003	-0.218	0.011	-0.113	-0.075	0.056
Eigenvalues	2.507	2.478	2.460	2.051	1.701	1.647
% Variance explained	8.953	8.850	8.784	7.324	6.075	5.881

Table 4. 4 *Components 7-11, variables with loadings contributing to each component and communalities (See table 4.3 for PCA explanation).*

Variable	7. Self-control	8. Confidence	9. Success at problem-solving	10. Interest to explore environment	11. Success in search	Communalities
S4 Toy motivation level	-0.074	0.171	-0.071	-0.030	0.075	0.836
S13 Play intensity	-0.045	-0.011	0.233	0.158	-0.084	0.750
Tail average position	-0.008	0.044	-0.065	0.330	-0.106	0.621
S16 Level of approach towards novel object	0.020	-0.052	-0.386	0.213	0.269	0.626
S11 Latency to gaze at handler	-0.057	-0.073	-0.117	0.113	0.037	0.811
S11 Time on task	0.151	0.087	-0.050	-0.016	0.157	0.782
S11 Different behaviours presented	0.241	-0.094	**0.402**	-0.140	0.105	0.832
S18 Duration vocalising	-0.155	-0.098	-0.026	0.066	0.137	0.707
S18 Pen gate approach	0.065	0.066	-0.065	0.015	-0.386	0.844
S18 Duration moving	-0.113	0.220	0.012	-0.144	**-0.415**	0.797
S5 Occasions selecting pointed can	-0.261	-0.116	0.084	-0.124	-0.073	0.586
S17 Duration on slippery surface	0.081	-0.071	-0.177	0.100	0.071	0.723
S4 Frequency choosing food	0.141	-0.266	0.222	0.059	0.013	0.641
Ears average height	-0.059	0.029	0.133	-0.019	0.120	0.736

S6 Latency to obey "Sit" and "Lie down"	-0.058	-0.120	0.011	0.104	0.063	0.792
S6 Duration Stay	**-0.403**	-0.136	-0.093	0.264	0.309	0.762
S3 Duration following handler	-0.170	0.322	0.035	-0.292	-0.248	0.687
S2 Duration interacting with handler	-0.013	0.062	-0.133	-0.092	-0.058	0.817
S1 Duration investigating tester	0.302	-0.180	-0.172	-0.316	0.371	0.822
S7 Immediate detour frequency	**0.772**	-0.061	0.002	0.002	0.033	0.644
S7 Mean latency to obtain food	**-0.523**	0.148	0.006	-0.105	-0.029	0.620
S12 Total correct choices	**0.437**	0.241	-0.336	0.118	-0.072	0.602
S16 Number of different behaviours	-0.079	**0.772**	0.108	-0.119	0.099	0.749
S14 Level of restraint required	-0.020	**-0.693**	0.005	-0.347	0.023	0.685
S8 Number of treats recovered	-0.034	0.150	**0.719**	0.135	0.087	0.668
S18 Duration resting	0.044	0.089	**-0.586**	0.079	-0.087	0.846
S1 Duration investigating environment	0.044	0.080	0.052	**0.881**	-0.007	0.835
S18 Latency to find treats	-0.002	-0.105	-0.109	0.030	**-0.816**	0.758
Eigenvalues	1.589	1.588	1.558	1.501	1.498	
% Variance explained	5.677	5.673	5.565	5.362	5.350	*Cumulative % 73.495*

4.4.2. Stress behaviours

The dogs (N=58) presented a mean of 23.5 stress signals (Min=2, Max=52, SD=11). Lip licking was predominant, making up more than half of all stress signals (58.8%), followed by whining (16.95%). The rest occurred less frequently. The test that produced the highest number of stress signals was the Ignoring subtest (S2), followed by the Toy-playing (S13) and Coat-wearing (S14) subtests (Table 4.5).

Table 4. 5 Ethogram of stress behaviours recorded during Medical Detection Dogs Test Battery and percentage of stress behaviours for each subtest (S1-S17).

Stress signal	Description	Extracted from	Percentage of dogs showing the behaviour in any subtests (%)
Low body posture	Dog shows posture below horizontal	Rooney and Bradshaw (2002)	2.4

Ears back	Dog's ears are held flat to its head or pulled back laterally	Rooney and Bradshaw (2002)	1.3
Tail tucked	Dog tucks its tail between its legs	Rooney and Bradshaw (2002)	0.0
Mouth	Dog places their mouth around the handler's or experimenter's hand	Rooney and Bradshaw (2002)	1.5
Lip lick	Brief protrusion of tongue from mouth, potentially licking nose or side of mouth.	Polgar et al. (2019)	58.3
Yawn	Fully opening mouth wide with ears back while inhaling and then closing while exhaling.	Polgar et al. (2019)	3.5
Startle	Short startled movement of the whole body	Van den Berg et al. (2003)	8.6
Tremble	Fine muscle tremors across the whole body or legs	Polgar et al. (2019)	Not presented
Body shake	Rapidly rotating abdomen and head from side to side	Polgar et al. (2019)	0.4
Whine	Producing a high-pitched vocalisation from the throat	Polgar et al. (2019)	17.0
Bark	Producing a resonating vocalisation by opening and closing mouth	Polgar et al. (2019)	4.3
Growl	Producing low-pitched rumbling vocalisation from throat	Polgar et al. (2019)	2.9
Bare teeth	Dog protrudes their lips or corners of their mouths exposing teeth	Van den Berg et al. (2003)	Not presented

Percentage of stress signals presented in each subtest (%)																
S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*	*S11*	*S12*	*S13*	*S14*	*S15*	*S16*	*S17*
4.9	*17.2*	*4.5*	*7.0*	*6.7*	*6.4*	*6.3*	*2.0*	*7.7*	*1.7*	*0.7*	*6.9*	*10.4*	*9.1*	*8.3*	*0.1*	*4.9*

4.4.3. Association of demography with dogs' behaviour in the test battery

4.4.3.1. *Dogs' sex*

The model (female was the reference category) was statistically significant $\chi 2(2) = 7.711$, $p < 0.05$, explaining 16.6% of the variance (Nagelkerke R2). For females (N=29), 'Food orientation' scores were likely to exceed those of males (N=29; p=0.05) (Table 4.6 and Figure 4.3)

Table 4. 6 *Association of the dogs' behaviour in the test with dog's sex (Male N=29, Female N=29). Reference category: Female.*

Test components	*B*	*SE*	*Wald*	*df*	*p*	*Odds Ratio*	*95.0% CI for Odds Ratio*	
							LL	UL
Constant	-0.01	0.28	0.00	1	0.98			
Persistence	-0.57	0.32	3.13	1	0.08	0.57	0.30	1.06
Food orientation	0.64	0.33	3.77	1	0.05	1.89	0.99	3.60

Note. B= unstandardised regression coefficient; SE= standard error of the coefficient; Wald=Wald test; df= degrees of freedom; p=p value, CI=confidence interval; LL= lower limit; UL=upper limit.

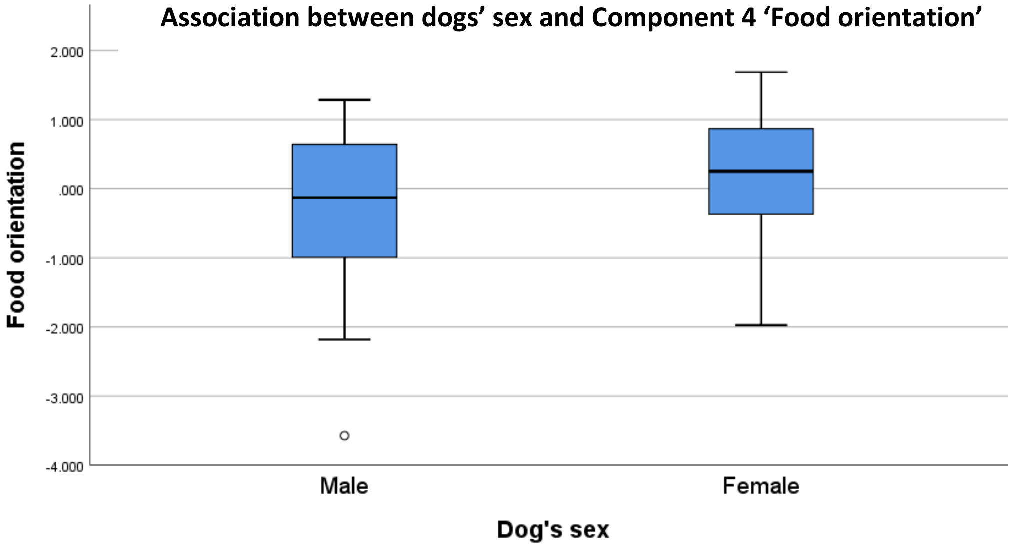


Figure 4. 3 *Logistic regression showed marginally significantly higher 'Food orientation' in females (N=29) than males (N=29). P=0.05. The 25th interquartile (lower box border) and the 75th interquartile (upper box border). The whiskers indicate the minimum and maximum mean latencies. The circles represent outliers.*

4.4.3.2. Dogs' training stage

The model was overall statistically significant (Trained dogs as the reference category) $\chi 2(2)= 13.13$, $p <.005$, explaining 46.6% of the variance (Nagelkerke R2). Trained dogs (N=19) were significantly more likely to have higher 'Obedience' than trainee dogs (N=39) (p=0.015), and 'Playfulness' was significantly higher in trainee dogs (p= 0.027) than in trained dogs. (Table 4.7 and Figure 4.4).

***Table 4. 7** Association of the dogs' behaviour in the test with training stage (Trainee N=39, trained N=19). Reference category: Trained dogs (Meaning of abbreviations in Table 4.6).*

Test components	*B*	*SE*	*Wald*	*df*	*p*	*Odds Ratio*	*95.0% CI for Odds Ratio*	
							LL	UL
Constant	-1.03	0.36	8.18	1	0.004	0.36		
Playfulness	-0.75	0.34	4.91	1	0.027	0.47	0.24	0.92
Obedience	1.44	0.59	5.89	1	0.015	4.22	1.32	13.48

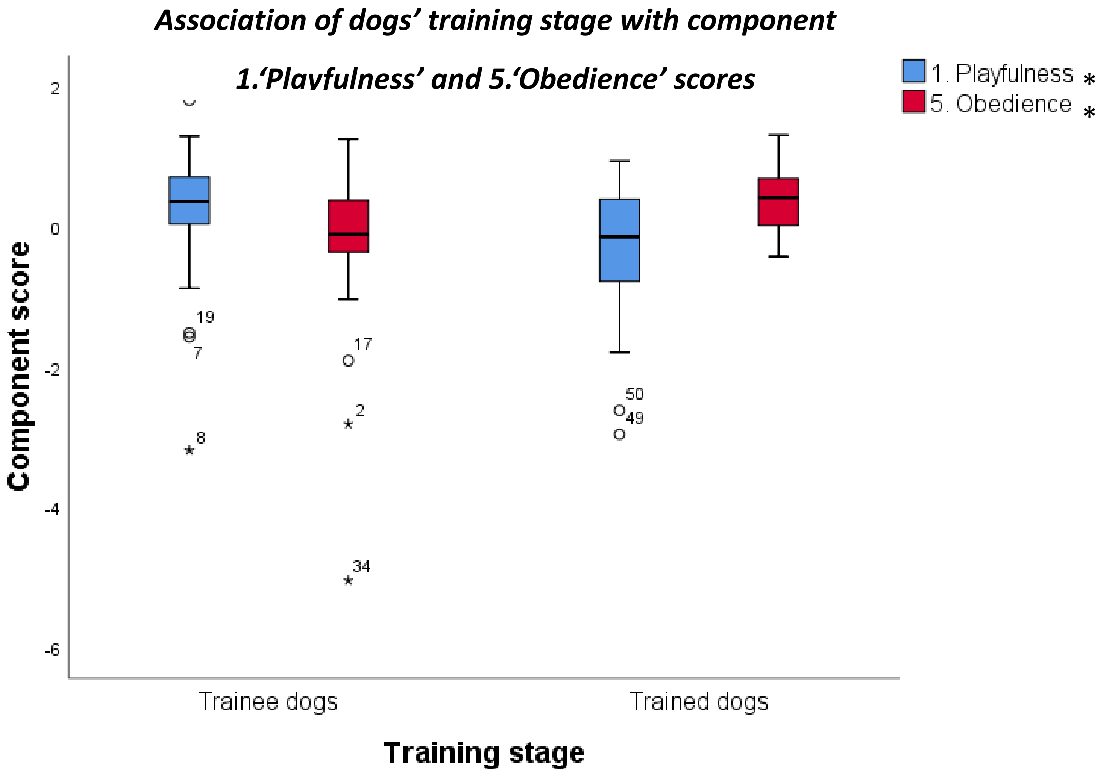


***Figure 4. 4** Significant variation of dogs' 'Playfulness' and 'Obedience' scores whether they were trainee (N=39) or trained (N=19). *P<0.05 (plots meaning in Table 4.3).*

4.4.3.3. *Dogs' age*

The model was overall statistically significant ($F_{(3, 54)} = 6.126$, $p < 0.005$, adj. $R^2 = 21.2\%$). Lower 'Playfulness' (p= 0.005) and higher 'Obedience' (p=0.030), and lower 'Confidence' (p=0.031) were significantly associated with greater age (Table 4.8).

Mann-Whitney U test showed that the median age of trained dogs was significantly higher (55.5) than trainee dogs (18) (U= 708.5, p= <0.001).

Table 4. 8 *Association of the dogs' demography and their behaviour in the test with dogs' DIAS scores (N=58).*

Test components	B	95.0% CI for B		SE B	P	β	R2	ΔR²
		LL	UL					
Model							0.25	0.21***
Constant	33.36	26.92	39.81	3.21	<0.001			
Playfulness	-9.47	-15.98	-2.97	3.24	0.005	-0.34		
Obedience	7.21	0.71	13.71	3.24	0.030	0.26		
Confidence	-7.18	-13.68	-0.68	3.24	0.031	-0.26		

Note. B=unstandardised regression coefficient; CI=confidence interval; LL= lower limit; UL=upper limit; R2=Coefficient of determination; SE B= standard error of the coefficient; β=Standardised coefficient; R^2= Coefficient of determination; ΔR^2= Adjusted R2, ***P <=0.001.

4.4.4. Association between test battery's behavioural components and DIAS scores

For all dogs (N=58), the mean Impulsivity overall score was 0.53 (± 0.09), ranging from 0.33 to 0.75

The model assessing the link between DIAS and the behaviour in the test was overall significant ($F_{(2, 55)}$ = 5.513, p < 0.01, adj. R2=13.7%). Higher ‘Playfulness’ (p=0.016) and higher ‘Persistence’ (p =0.031) were significantly linked with higher DIAS scores (Table 4.9).

Table 4. 9 *Association of the dogs' demography and behavioural factors with dogs' DIAS scores (N=58) (Meaning of abbreviations in Table 4.8). **P <=0.01.*

Dogs’ DIAS scores	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.17	0.14**
Constant	0.54	0.51	0.56	0.01	<0.001			
Playfulness	0.03	0.01	0.05	0.01	0.016	0.31		
Persistence	0.03	0.00	0.05	0.01	0.031	0.27		

4.4.5. Association of the dogs' behaviour on the test battery with their trainers' trait ratings.

Four components from the trainers' ratings from Chapter 3 (3.4.2.8) (In italics to differentiate them) showed significant correlations with behavioural test components. *'Responsiveness to training'* was correlated with 'Obedience' (R=0.315 p=0.016) and with 'Self-control' (R=0.280 p=0.033) and marginally with 'Success in search' (R=0.250 p=0.058). *'Keenness to please'* was associated with 'Interest in exploring environment' (R=0.313 p=0.017) and 'Success in search' (R=0.326 p=0.012). *'Human orientation'* was negatively correlated with 'Confidence' (R=-0.305 p=0.020), and *'Vocal excitability'* had a negative correlation with 'Success at problem-solving' (R=-0.341, p=0.009) and marginally significantly associated with 'Reactivity in holding pen' (R=0.249, p=0.06).

4.5. Discussion

This chapter aimed to develop a test battery potentially valuable for the selection of MDD. This is the first study (to my knowledge) exploring an assessment method specific to this dog population. The test included 18 temperament and cognitive subtests to assess the most important traits for MDD from the survey (2.4.7). The test battery's measures were clustered into 11 main behavioural components. Several significantly varied with dogs' demography, with the dogs' impulsivity scores, and some were correlated with the trainers' opinion on their dogs' performance from Chapter 3 (3.4.2.8). From these preliminary findings, the test could become a practical resource for MDD assessment and selection after confirmatory analysis and refinement. Next, I discuss the study's results, limitations and potential application.

4.5.1. The test battery development

The subtests within the test battery were based on those scientifically validated for dog detection and other working roles. I intended the test to be as specific as possible to MDD, measuring the behavioural traits considered most relevant for this role.

The test assessed both temperament and cognitive behavioural aspects as this has been suggested to increase test robustness (Bray et al., 2017b). Each subtest included several variables, often combining both objective and subjective measures (1.2.4.4). However, several variables were excluded upon data reduction to increase the tests' statistical strength. Yet, the components seem descriptively similar to the identified characteristics for MDD from the survey (2.4.7) and those that trainers prioritised when rating OA in Chapter 3 (3.4.2.5).

The extensive piloting process allowed me to refine the test and increase its feasibility by discerning which subtests to include in the test based on the dogs' responses and their feasibility and functionality, and adjusting their duration, so the dogs could go through the whole test without being fatigued or stressed. I also ensured that the equipment was economical and easy to move between subtests and locations.

I performed most piloting on companion dogs at the UOB. It would have been ideal to do more piloting with MDD and in the experimental facility since these dogs' rearing and training methods are standardised (3.1.4). In contrast, companion dogs tend to show greater response variability as they are not pre-selected and specifically trained (e.g. Hare et al., 2018). I aimed for the largest MDD sample available. Therefore, I could not spare them for piloting. However, I practised the final test version at the charity with three retired detection dogs that did not take part in the final test and adapted the subtests' into the charity's experimental setting. As this is an exploratory study, the test included numerous subtests assessing various dog traits. From this chapter's findings and the next, it may now be possible to develop a briefer version of the test to increase its feasibility for further confirmatory research.

I initially contemplated assessing Intra and inter-observer reliability and test-retest reliability. However, this was impossible for logistic reasons and time limitations, partly due to the disruption of the pandemic. It is also challenging to test a dog for a second time since they may habituate to stimuli which are no longer perceived to be novel. These would need to be replaced by new ones, while even items considered similar replacements might be perceived by the dogs differently (Taylor and Mills, 2006). However, assessing reliability in a shorter test version would be necessary if the test is to be applied for MDD selection.

4.5.2. Variables derived from the test battery

Similar to Rooney et al. (2003b) and Polgar et al. (2019), in the initial data reduction phase, I excluded 12 variables showing low variation attributed to the sample pre-selection and rearing process. Most of the dogs in the charity are procured from specific sources, socialised, and trained with standardised methods. Hence, they may adapt faster to different scenarios than other dog populations. Some stimuli may not be appropriate to trigger variation in dogs' responses due to their intensity or composition, even if the selected subtests attempted to target this population. For instance, the S10 'Noise distraction test' showed low variation (4.2.5.2.2) as most of the dogs did not react to the distracting noise, and the subtest was therefore excluded; and the S17 'Readiness of approach to the slippery surface' (4.2.5.2.2) where very few dogs hesitated to step in. I eliminated 16 variables for being highly correlated to others. Most were similar or were negatively correlated, being mutually exclusive.

The variables included in the PCA were reduced to 27 to reach the minimum criterion of sample adequacy from the KMO (Kaiser, 1974). I made systematic variable inclusion decisions (Field, 2013; Laerd Statistics, 2015). This still involved some subjectivity when deciding which variables to keep to increase the KMO >0.5 but attempted to ensure that the most relevant variables were represented within the components.

4.5.3. Behavioural components

The PCA produced 11 components which explained most of the total test variance (74.5%). I labelled them according to the variables' loadings on the behavioural components: 1.'Playfulness', 2.'Persistence', 3. 'Reactivity in holding pen', 4. 'Food orientation', 5. 'Obedience', 6.'Level of attention to handler', 7. 'Self-control', 8. 'Confidence', 9. 'Success at problem-solving', 7. 'Interest in exploring environment', 11.'Success in search'. For some components, it was straightforward to understand the relationship between variables. For instance, 'Interest in exploring environment' only included 'Duration exploring environment'

(S1). In ‘Obedience’, both variables were related to the dog attending to the handler cues, and ‘Reactivity in the holding pen’ mainly involved behaviours associated with restlessness in the ‘Holding pen test’ (S18).

Some associations between variables within the components were more complex: dogs with higher scores on ‘Confidence’ showed more approach behaviours towards the motorised ball (S16), which a bolder dog may display. Yet this may also relate to playful reactivity to the toy or high arousal due to fear. Also, dogs needed less restraint when putting the coat on. More confident dogs may be more tolerant of being manipulated and wearing a coat, hence being calmer. However, dogs could also freeze due to high stress or learned helplessness (Maier and Seligman, 1976) and then not moving when putting the coat on.

Dogs scoring higher for ‘Success at problem-solving’ retrieved more treats from the puzzle toy (S9), tried different behaviours to access the food in the Unsolvable Task (S11) and were more restless in the holding pen (S18). Counting the number of treats taken from the puzzle may partially reflect the dogs' involvement and ability to solve the task, as extracting the treats entails a series of steps. However, it did not measure whether the dogs were systematically solving the puzzle. Some dogs tended to paw, bite or hold the toy hastily. Others were slower to get the treats but worked more methodologically. Thus, dogs with higher scores were possibly more excitable, restless (also in the holding pen) and tended to try many ways to obtain the food quicker. The trainers described observing similar dogs’ working styles during scent discrimination trials (Personal communication with trainers of Medical Detection Dogs®). Future research could assess whether dogs being systematic when performing problem-solving tests correlates with their searching style in detection trials.

Dogs with higher scores for ‘Self-control’ extracted the treat faster from the cylinder (S8) and found the ball more often in the searching ball subtest (S12). They may tend to inhibit a spontaneous response and concentrate on finding the ball. Previous research found an association between higher motor inhibition in the cylinder test and better performance in an explosives detection task (Tiira et al., 2020). MDD dogs scoring higher in ‘Self-control’ may be more contained and able to focus when making decisions on searching tasks.

The body positions within some components may have different implications. Body language varies individually and should be interpreted carefully as it is affected by several factors. For instance, in 'Food orientation', the high ear positions may be associated with attention to the treats. Ear movement in animals has been linked with attention towards stimuli (Wathan and McComb, 2014; Descovich et al., 2017) and associated with positive states from obtaining a high-value reward (Panksepp, 1998; Caeiro et al., 2017; Bremhorst et al., 2019). Dogs with higher 'Playfulness' scores displayed higher tail positions, probably associated with higher arousal during playing or higher confidence. A study identified that dogs more involved in a tug of war game had higher scores of the component 'Confident Interactivity', involving higher average tail positions and less tucked tail frequency (Rooney and Bradshaw, 2002). The dogs' ear and tail positions give valuable information on potential motivation. Still, the dog's body language should be regarded as a whole to get a complete picture of their underlying states.

The 11 test components (and the variables interacting within each) represent behaviour aspects relevant to MDD. Hence, being helpful to examine their association with the dogs' performance measures in the subsequent chapters (Chapters 5 and 6).

4.5.4. Stress behaviours

The dogs displayed different stress signals during the test, predominantly lip licking, followed by whining. A high prevalence of stress-related behaviours may be as some dogs may have perceived some subtests as stressful or challenging. However, stress-related behaviours may rely on different underlying motivations that often overlap (Shiverdecker et al., 2013).

Past research has linked dog stress-related behaviours such as lip licking with physiological parameters indicative of high-stress levels (Beerda et al., 1998) and with less success in detection dogs (Rooney et al., 2003b). However, other studies have not found this relationship or suggested an alternative function as a coping mechanism. Lip licking has been considered a stress relieving system (Pastore et al., 2011; Shiverdecker et al., 2013) and a displacement behaviour in conflict or high arousal (Lund and Jørgensen, 1999). Whining has been associated with arousal with a positive motivation (e.g. excitement before accessing a reward) or a negative one (e.g. frustration when reward access is denied) (Jakovcevic et al., 2013).

Bremhorst et al. (2019) measured dogs' facial expressions with the Dog FACS (Facial Action Coding System) (Waller and Micheletta, 2013). They found lip licking and jaw-dropping more frequently in dogs submitted to an adverse condition (food access denial) compared to a positive one (access to food). However, Caeiro et al. (2017) found a higher presentation of lip-licking in positive anticipation contexts. In this study, the dogs tended to display these behaviours while held by their handler before performing the task, which frequently involved the obtention of food or toy rewards. Arousal potentially escalated from reward anticipation, not differing from the excitability build-up before a search training trial. Yet, in subtests involving more atypical situations, such as the Ignoring subtest (S2) or the Novel object subtest (S16), lip licking and other behaviours may be the product of negative affective states as lack of predictability has been linked with anxiety or frustration (Gray, 1987; Papini and Ludvigson, 1994). Also, the high number of subtests could have generated a 'trigger stacking' effect due to the exposure to multiple situations (Hargrave, 2015), potentially increasing stress signals seen in the later tests. Therefore, it is crucial to consider the impact of each subtest and its accumulative influence on dogs' welfare when refining the test battery.

4.5.5. Association of demography with dogs' behaviour in the test battery

The study showed a significant association of some components with sex, age and training status:

'Food orientation' was marginally significantly higher in female dogs than in males. Research in different species indicates that females had less energy expenditure and a naturally higher tendency to accumulate body fat than males, partly attributed to keeping reproductive activity deprivation times (Hoyenga and Hoyenga, 1982; Geary and Lovejoy, 2008). Females seem less food-selective than males (Houpt and Smith, 1981) and have a greater risk for obesity (Geary and Lovejoy, 2008). Most dogs in this study were spayed or neutered. Neutering modifies the influence of sex on individual traits due to alterations in hormone levels (Wallis et al., 2020). In rodents, females tend to gain weight after spaying due to the lack of the hormone estradiol, which regulates food intake (Eckel, 2011). Dog breed could have an effect as most of the dogs were Labrador Retrievers, who tend to be highly food-oriented (Gerencsér et al., 2018), but there were also Labrador retrievers of both sexes. Therefore, is unlikely that dog breed influenced differences between sexes.

Dog's age and training stage may have a confounding effect as trainee dogs were significantly younger than trained dogs, and the association of age and training stage with the behavioural components was similar. ‘Playfulness’ significantly decreased as dogs got older and in trained dogs. Although adult dogs continue displaying play-related behaviours, these tend to decline with social maturity.

‘Obedience’ scores were significantly higher with age and advanced training stage; this was likely as more mature dogs may develop higher self-control and focus on complying with their trainers' cues and have been exposed to more training. Past research found that older guide dogs tended to comply with obedience tasks more than younger dogs (Batt et al., 2008), and older cancer detection dogs were better than younger ones in increasingly complex training (Walczak et al., 2012).

‘Confidence’ scores significantly decreased with age (although they were not significantly linked with training stage). This is somehow surprising as it would be thought that older dogs are more confident than younger ones. The component ‘Confidence’ involved displaying more behaviours towards the novel object (S16) and struggling less when putting the coat on (S14). Younger dogs may be more playfully reactive towards the object and bolder towards a new situation. In comparison, older dogs may be more cautious when assessing new contexts without approaching immediately.

4.5.6. Association with impulsivity

Both 'Playfulness' and 'Persistence' were significantly associated with the dogs' DIAS scores. Highly playful dogs tend to be more energetic, excitable and quicker to respond to different situations. This sample was composed predominantly of younger dogs from working breeds which tend to be more playful and less inhibited (4.5.5).

For both components, 'Playfulness' and 'Persistence' dogs with higher scores displayed more behaviours during the Unsolvable Task (s11). More playful and impulsive dogs may persist in accessing the food in the container or consider chasing the container around the room as if it was a game.

Dogs with higher 'Persistence' were involved in the Unsolvable Task for longer and performed a greater variety of behaviours to take out the treat from the container (S11), but were quicker to take the treat out from the cylinder (S8), also an element from the 'Self-control' component. This was similar to Tiira et al. (2020), where explosives detection dogs that were quicker to extract the treat from the cylinder spent longer involved in the Unsolvable Task

Persistence is linked with self-inhibition, and both are inherently multifaceted (Tiira et al., 2020). Persistence may be observed as repeating a specific behaviour to achieve something or trying many behaviours to obtain a resource by trial and error. A highly reward-motivated dog may get frustrated when trying to access a reward unsuccessfully and react without inhibition (Bremhorst et al., 2019; McPeake et al., 2019). Companion dogs with higher DIAS scores performed more paw presses in a delayed reward apparatus during a no-reward period, and they were less able to wait for the reward (Wright et al., 2012), suggesting that more impulsive subjects may persevere in unproductive behaviours to achieve a goal.

Depending on how it manifests and the context, persistence may be a positive or a negative trait for working performance regarding how it mediates the dogs' decisions. For instance, a dog should persevere in learning a task or keep alerting a client of an impending crisis but not display several undesirable behaviours to achieve their goal, e.g. jumping and barking to get attention or falsely alerting a target odour to get rewarded.

4.5.7. Association between the dogs' behaviour on the test battery and their trainers' trait ratings.

I assessed if the behavioural components from the test battery were correlated with PCA components from dog trainers' ratings (In italics; full description in Chapter 3.4.2.8). *'Responsiveness to training'*, which included trait ratings related to task motivation, searching ability, obedience, Problem-solving skills, and willingness to try new behaviours and a lower tendency to be distracted, was significantly associated with higher 'Obedience', 'Self-control' and 'Success in search' in the test battery.

'Keenness to please', involving ratings on higher consistency, obedience and tendency to try new behaviours, was associated with 'Interest in exploring environment' and 'Search success' during the test. Similar findings in humans indicated that a greater tendency to take active initiative was linked with more successful behaviours (de Ridder et al., 2011).

Dogs rated higher on *'Human orientation'* tended to have significantly less 'Confidence' in the test battery. Dogs more attached to their handler may feel more reassured by remaining nearer to them instead of approaching the novel object (S16) or when wearing the coat (S14). Dogs with higher *'Vocal excitability'*, including traits related to higher arousal, had significantly lower 'Success at problem-solving' in the test battery and were marginally more 'Reactive in the holding pen', coinciding with the trainers' perception that these dogs were more excitable and distracted in their daily tasks.

These relationships suggest that the dogs' behaviour in the test battery reflects their trainers' opinion on aspects of their general performance to a certain degree showing acceptable criterion validity (1.2.3). This highlights the relevance of combining objective and subjective measures when assessing a dog for a task, as suggested in past studies (Rooney et al., 2007a; McGarrity et al., 2016). However, In Chapter 3, trainers failed to distinguish several traits and were inconsistent in their ratings over time. Also, these scales may not directly correspond with what the subtests measure. Therefore, the ratings' components require further study as criterion validity parameters. Future research may use revised scales that correspond directly to those measured in the test and assess the opinion of more than one person on the dog.

4.6. Conclusion

This study designed a test battery to quantify aspects of the behaviour of MDD performance. The test battery included temperament and cognitive subtests based on scientifically validated tasks. The test battery yielded 11 ostensibly meaningful behavioural components representing individual aspects relevant to MDD. Several components varied with dogs' demographic aspects and impulsivity scores and were significantly correlated with some components from trainers' ratings.

Future confirmatory research may aid in refining the test battery with larger MDD samples, assessing reliability and validity measures, to adapt the test for practical application in MDD training organisations.

The next chapter assessed the test battery's predictive value. It explored the associations between the test’s behavioural components, the dogs’ demography and the DIAS scores with the measures of success from Chapter 3 across training stages and prospective tasks.

Chapter 5. Associations between medical detection dogs' behaviour in the test battery and their performance:

How this varies for different training stages and tasks

Abstract

This study assessed how the medical detection dogs (MDD) performance in the test battery (Chapter 4) linked with the performance of their tasks and how this varied across disciplines and training stages. The dogs (N=58)were previously tested with a test battery, and the measurements extracted were clustered into 11 behavioural components (Chapter 4).

In the current chapter, I explore whether these components and demographic elements, including sex and age, were related to the success measures derived in Chapter 3 (3.5). For the trainee dogs (N=39), I assessed the associations between the test battery components, demographic elements and the dogs' training outcomes. I examined the relationship between test battery components, the dogs' Composite total ability score (CTAS) for trainee bio-detection (N=22) and Assistance dogs (N=17) separately to investigate whether the association between dogs' behaviour in the tests and their training performance varies with each prospective discipline.

To investigate how the trained dogs' behaviour in the test battery (N=19) was associated with their performance in their detection projects, I assessed the association between the behavioural components and demography and their CTAS. I also investigated whether these were linked with the dogs' scent sensitivity and specificity scores in all dogs enrolled in projects at the end of data collection (N=27).

For trainee dogs, those with higher ‘Interest in exploring environment’ (p= 0.031) and lower 'Playfulness' (p=0.02) and 'Level of attention to handler' (p=0.012) were significantly more likely to remain in the programme. In trainee Bio-detection dogs, higher 'Food orientation' (p= 0.047), 'Obedience' (p=0.005) and 'Self-control' (p=0.018), but lower ‘Interest in

exploring environment' (p=0.007) were significantly related to higher CTAS. In Assistance dogs, older age (p<0.001), higher 'Self-control' (p=0.03), 'Interest in exploring environment' (p=0.001) and 'Reactivity in holding pen' (p=0.018) but lower 'Playfulness' (p=0.003) and 'Persistence' (p=0.013) and the presentation of fewer stress behaviours during the test battery (p<0.001) were associated with higher CTAS.

In trained dogs, younger age (p<0.001), higher 'Food orientation' (p<0.001), 'Interest in exploring environment' (p=0.001) and 'Success in search' (p<0.001) but lower 'Confidence' (p<0.001) and 'Obedience' (p=0.004) and fewer stress behaviours (p=0.002) were significantly linked with higher CTAS. Of the 27 trained dogs at the data collection cut-off point, male dogs (p=0.008), older (p=0.028), with higher 'Confidence' (p=0.038) and 'Success in search' (p = 0.002) were significantly more scent sensitive in their projects. Scent specificity was significantly associated with younger age (p=0.003), 'Food orientation' (p<0.001) and 'Success at problem-solving' (p=0.005), but negatively linked with 'Level of attention to handler' (p<0.001), 'Self-control' (p=0.002), 'Interest in exploring environment' (p=0.01) and 'Success in search' (p=0.04).

Some test battery components were found to be related to task performance similarly across one or more datasets, whereas others were not clearly related or showed distinct relationship trends across analyses. These findings suggest that at least some test battery components may provide an effective predictive tool of task performance, aiding in identifying elements to focus on during dogs' training and selection, which may be relevant to long-term working ability. However, further refinement and confirmatory analysis are needed to improve the test battery's predictive value.

5.1 Introduction

Working dogs are frequently assessed with test batteries, allowing the objective evaluation of different individual characteristics to select the best-suited individuals for the task as early as possible (Jones and Gosling, 2005).

However, assessing any test battery's predictive value with appropriate measures of role success is crucial (1.2.5).

When examining the predictive validity of a test battery, training outcome (pass or fail the training programme) is often used as a success parameter (1.2.5). However, this measure can be argued to be subjective (3.1.1). Assessing tasks' success by integrating subjective (e.g. trainers' ratings) and objective parameters of ability (e.g. detection accuracy in search trials) may provide more information on a dog's future performance (1.2.5). However, few studies have relied on this approach (e.g. Rooney et al., 2007b; Tiira et al., 2020).

Dogs are frequently young when assessed with test batteries and probably have not fully developed (1.2.5). Some traits remain stable throughout animals' lives, e.g. distractibility (Harvey et al., 2016) and impulsivity (Riemer et al., 2014a). However, dog behaviour may change throughout maturity (Wallis et al., 2020), affecting their task performance during their working life. However, this has seldom been investigated (Brady et al., 2018). In analogy, students graduating from university with good grades are not always proficient in their professional careers. They may be more successful when performing simple tasks within a company but struggle when promoted to more complex projects (e.g. Bryła, 2015). Therefore, working dog assessments must be developed to predict and monitor long-term operational effectiveness.

The previous chapter explored the development of a test battery to investigate how important traits for MDD performance varied in a sample of dogs. The test battery included 18 subtests measuring relevant characteristics. In addition, it assessed the presentation of 13 stress-related behaviours based on an ethogram. Of the 96 variables extracted and submitted to data reduction, 11 components explaining the dogs' behaviour when tested were identified (4.4.1). It also assessed associations between the test battery components and dog demography elements (4.4.3), how these linked with the dogs' impulsivity scores (4.4.4), and with their trainers' ratings on different traits (4.4.5). This chapter investigates the predictive value of the test battery on MDD performance. Specifically, it assesses the association between the test battery components with success measures from chapter 3

(3.5). It also investigated whether total stress behaviours in the test battery were associated with task performance.

Chapters 2 and 3 found that the levels of several traits were similar, while others differed across bio-detection or assistance tasks (2.4.5.1; 3.4.2.2). In Chapter 3, I identified differences in how trainers rated trainee and trained dogs (3.4.2.3). In Chapter 4 certain components of the test battery differed according to the dogs' sex, age and training stage (4.4.3). Therefore, this study assesses the association between the dogs' behaviour in the test battery, demographic elements, and success measures across prospective disciplines and training stages.

More specifically, this chapter investigates the following:

a. The association between trainee dogs' behaviour in the test battery and demographic variables with training outcomes.
b. How the associations between dog behaviour in the test battery, and demography, vary with the ability of trainee dogs who were intended to perform bio-detection and assistance tasks.
c. How trained dogs' behaviour in the test battery and demography are linked with their ability at regular detection performance evaluations.
d. How trained dogs' behavioural components in the test and demography are associated with their scent sensitivity and specificity scores.

5.2 Methods and statistical analyses

5.2.1 Different samples: trainee dogs and trained dogs

As described in Chapter 3 (3.2.2; Appendix 4), the sample tested with the test battery (N=58) was composed of two groups: trainee dogs (N=39) and trained dogs (N=19). Additionally, at the end of data collection, data on sensitivity and specificity from 27 dogs enrolled in different detection projects were available (3.3.2). Since these dog samples differed in age, training and activity level, they were analysed independently to answer specific questions for each group.

5.2.2 Dogs' behaviour in the test battery and success measures

The 11 components extracted from the test battery were labelled according to the distribution of the variables with higher loadings: 1.'Playfulness', 2. 'Persistence', 3. 'Reactivity in holding pen', 4. 'Food orientation', 5. 'Obedience', 6. 'Level of attention to handler', 7. 'Self-control', 8. 'Confidence', 9.'Success at problem-solving', 10. 'Interest in exploring environment' and 11. 'Success in search' (4.4.1).

The 13 total stress behaviours were extracted independently from the test components (4.4.2). The most frequent were lip licking (58.33%) and whining (16.95%) (Table 4.5).

The demographic elements assessed, as in Chapter 4 (4.4.3), included dogs' sex and age.

The success measures used here are those derived from Chapter 3: for trainee dogs, these are training outcome (retained in the system or failed training: 3.2.1) and CTAS, based on the trainers' opinion on the dogs' training performance (3.2.2.7). For trained dogs, these were CTAS, based on the trainer's opinion on the dogs' ability in scent detection trials. Scent sensitivity (the proportion of targets that are correctly alerted; 1.2.1.2) and scent specificity scores (the dogs' propensity to avoid false positives when discriminating an odour; 1.2.1.2). These as objective parameters of the dogs' scent discrimination ability (3.3.2).

5.2.3 Statistical models

I performed a series of binary logistic regressions and multiple regression analyses to assess how the test battery behaviour and demographic variables were related to training success measures, dealing separately with trainee and trained dogs. I built models including as independent variables the 11 PCA component scores (4.4.1), total stress behaviours (4.4.2) and demographic variables. I used the different success measures as the dependent variables. The procedures were similar to Chapter 4 (4.3.2) to reduce multiple comparisons and focus on main effects.

5.2.4 Trainee dogs

5.2.4.1 Association between the dogs' behaviour in test battery and training outcome

I ran a binary logistic regression to assess whether any test components were related to whether the trainee dogs (N=39) were successful in the training programme or failed. The model included the test battery's behavioural components and demography as independent variables and the training outcome as the dependent variable (Remained in the system vs rejected).

5.2.4.2 Association between dogs' training outcomes and their CTAS

I investigated whether a dog's probability of remaining in the programme correlated to their CTAS since this measure provides broader information on dogs' ability than a binary success/failure. The model described previously (5.3.1.1) produced an estimated probability score of success in training, indicating that if the estimated probability of a dog succeeding was higher than or equal to 0.5 (better than chance), the dog was classified as 'Remained in the system'. If it was lower than 0.5, the dog was classified as 'Rejected' (Laerd statistics, 2015).

I extracted this score and correlated it with the dogs' CTAS with Pearson correlations to investigate if the probability of success in training was associated with the dogs' ability CTAS (Personal communication with Knowles, 2020).

5.2.4.3 Association between the dogs' behavioural attributes and their CTAS for each task

To investigate the association between the dogs' behaviour in the test battery and the ability of dogs intended for different tasks, I assessed prospective Bio-detection dogs (N=22) and dogs intended for assistance tasks (N=17) separately. I explored the association between the test components and demography for each task and the CTAS through multiple regression analyses since CTAS was significantly linked with the dogs' probability of training success (5.3.1.2) and may provide a broader parameter of a dog's ability than training outcome.

5.2.5 Trained Dogs

5.2.5.1 Association between dogs' behaviour in the test battery and their CTAS

To explore whether trained dogs' behaviour in the test battery was associated with their CTAS when performing regular search trials, I conducted a multiple regression analysis where the dogs' test components and demography were the independent variables, and their CTAS was the dependent variable.

5.2.5.2 Association between behavioural components and Scent Sensitivity and Specificity scores

I investigated whether the test battery components were related to the trained dog's sensitivity and specificity scores (3.2.3). Each dog worked to detect different conditions with different complexity. These measures were used to assess individual dog performance but not to make comparisons between conditions.

I conducted separate multiple regression analyses with the test components and demography as independent factors and dogs' sensitivity and specificity scores correspondingly as the outcome variables. For these analyses, I included the data of 27 dogs, including the 19 trained dogs at the test battery time point and eight dogs enrolled in detection projects up to the end of data collection (3.4.3), to assess the largest sample available.

5.3 Results

5.3.1 Trainee dogs

5.3.1.1 Association between the dogs' behaviour in test battery and training outcome

Overall the model was significant (χ2(4)= 19.33, p <.005) with 52.7% (Nagelkerke R2) of the variance explained, predicting dogs with favourable training outcomes in 91.3% of the cases and dogs failing in 68.8%, with an overall predictive value of 82.1%. The model showed that lower 'Playfulness' and 'Level of attention to handler'; and higher ‘Interest in exploring environment’ were significantly associated with a higher likelihood of remaining in the system (Table 5.1).

Table 5. 1 *Association between the 11 test behavioural components, dogs' demography, and training outcome (Remained in the system N=23, rejected N=16; reference category: Remained in the system).*

Test components	B	SE	Wald	*df*	*P*	Odds Ratio	95.0% CI for Odds Ratio	
							LL	UL
Constant	1.32	0.55	5.81	1	0.016	3.73		
Playfulness	-1.30	0.56	5.40	1	0.02	0.27	0.09	0.82
Obedience	1.14	0.65	3.07	1	0.08	3.11	0.87	11.10

Level of attention to handler	-1.35	0.53	6.37	1	0.012	0.26	0.09	0.74
Interest to explore environment	1.15	0.53	4.68	1	0.031	3.17	1.11	9.02

Note. B= unstandardised regression coefficient; SE= standard error of the coefficient; Wald=Wald test; df=degrees of freedom; p=p value, CI=confidence interval; LL= lower limit; UL=upper limit.

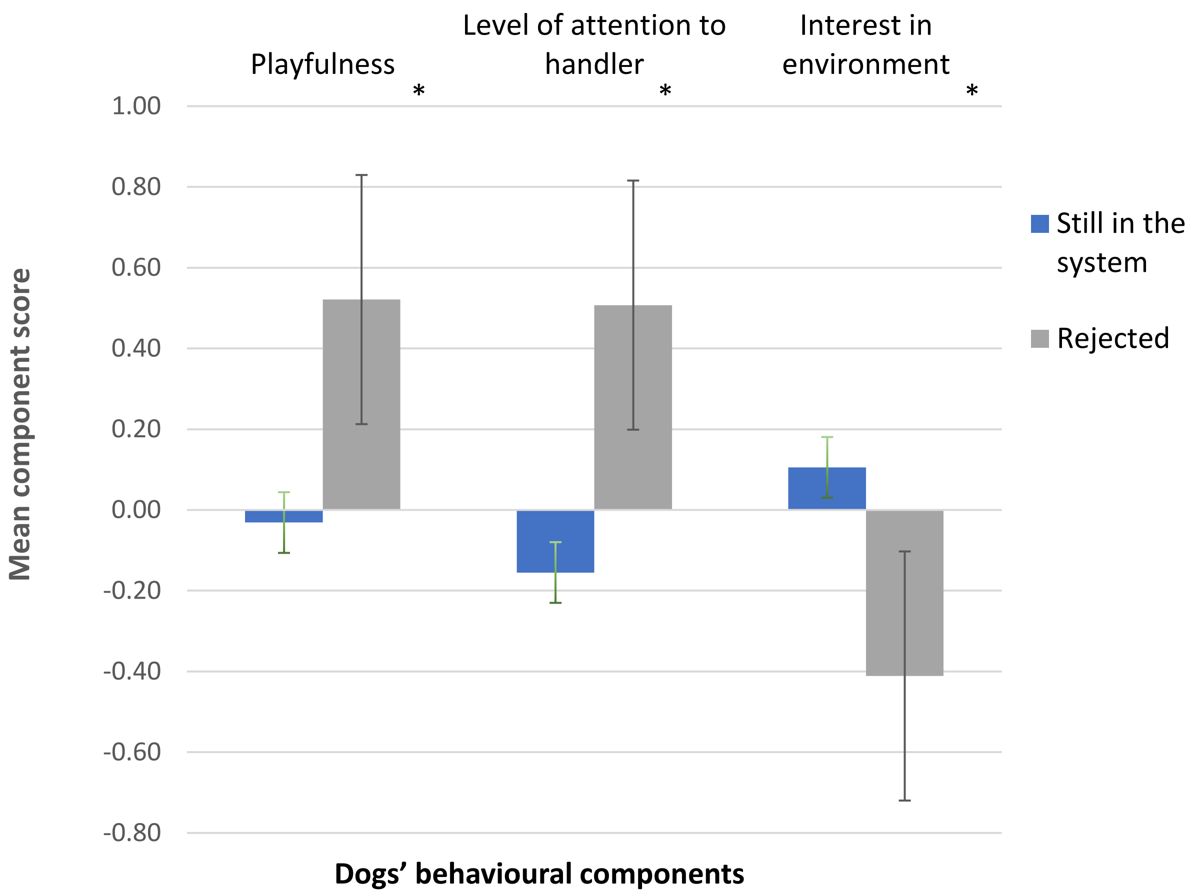


Figure 5. 1 *Significant difference in mean scores of 'Playfulness', 'Level of attention to handler' and 'Interest in exploring environment' related to training outcome (Remained in the system N=23, Rejected N=16). *P<.05.*

5.3.1.2 Association between dogs' training outcomes and their CTAS

Pearson correlation was significant (R=0.497 p=0.001), showing that the higher the dogs' CTAS, the higher the probability of them succeeding in training.

5.3.1.3 Association between the dogs' behaviour in the test battery and CTAS for each task

5.3.1.3.1 Bio-detection dogs

The overall model was significant ($F_{(4, 17)} = 5.159$, $p < 0.01$, adj. $R2 = 44.2\%$). Higher 'Food orientation', 'Obedience' and 'Self-control' scores but lower 'Interest in exploring environment' were significantly associated with higher CTAS (Table 5.2) .

Table 5. 2 *Association between Bio-detection dogs' (N=22) test behavioural components and demography and their CTAS*

Test components	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.42	0.35**
Constant	28.87	21.23	36.51	3.62	<0.001			
Food orientation	6.39	0.10	12.68	2.98	0.047	0.35		
Obedience	9.12	3.13	15.12	2.84	0.005	0.54		
Self-control	8.28	1.60	14.97	3.17	0.018	0.47		
Interest in exploring environment	-14.63	-24.73	-4.53	4.79	0.007	-0.53		

Note. B= unstandardised regression coefficient; CI=confidence interval; LL= lower limit; UL=upper limit; SE B= standard error of the coefficient; β=Standardised coefficient; R^2= Coefficient of determination; ΔR^2= Adjusted R2, **P <0.01.

5.3.1.3.2 Assistance dogs

The model was statistically significant ($F_{(9, 7)} = 23.88$, $p < 0.001$, adj. $R2 = 92.8\%$). Older age, higher 'Self-control', 'Interest in exploring environment' and 'Reactivity in holding pen', but lower 'Playfulness', 'Persistence' and fewer stress behaviours tended were significantly linked to higher CTAS (Table 5.3).

Table 5. 3 *Association between Assistance dogs' (N=17) test behavioural components and demography and their CTAS, ***P <0.001. (Abbreviations meaning in Table 5.2.)*

Test components	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.97	0.93***
Constant	64.77	49.90	79.64	6.29	<0.001			
Age in months	2.10	1.41	2.78	0.29	<0.001	0.78		
Total stress	-2.16	-2.82	-1.50	0.28	<0.001	-1.07		
Playfulness	- 10.78	-16.38	-5.18	2.37	0.003	-0.49		
Persistence	-4.83	-8.29	-1.37	1.46	0.013	-0.25		
Reactivity in holding pen	5.67	1.33	10.01	1.84	0.018	0.32		
Food orientation	-3.59	-7.89	0.71	1.82	0.089	-0.18		
Obedience	-5.36	-11.30	0.57	2.51	0.070	-0.28		
Self-control	6.19	0.78	11.60	2.29	0.030	0.33		
Interest in exploring environment	6.64	3.66	9.62	1.26	0.001	0.41		

5.3.2 Trained dogs

5.3.2.1 Association between trained dogs' behaviour in the test battery and their CTAS

The model was statistically significant ($F_{(7, 11)}$ = 9.530, p < .005, adj. R2 = 76.8%). Younger age, higher 'Food orientation', ‘Interest in exploring environment’ and 'Success in search' but lower 'Confidence’' and 'Obedience' and fewer stress behaviours were significantly associated with higher CTAS (Table 5.4).

Table 5. 4 *Associations between trained 'dogs' (N=19) test behavioural components and demography, and their CTAS **P <0.01 (Abbreviations meaning in Table 5.2).*

Test components	B	95.0% CI for B		SE B	*P*	B	R2	ΔR²
		LL	UL					
Model							0.32	0.77**
Constant	58.43	45.10	71.76	6.06	<0.001			
Age in months	-0.43	-0.58	-0.29	0.07	<0.001	-1.12		
Total stress	-0.65	-1.01	-0.29	0.16	0.002	-0.61		
Food orientation	9.16	5.03	13.29	1.88	<0.001	0.72		
Obedience	-11.25	-18.02	-4.49	3.07	0.004	-0.48		
Confidence	-18.40	-24.44	-12.35	2.75	<0.001	-1.05		
Interest in exploring environment	7.66	3.78	11.54	1.76	0.001	0.68		
Success in search	15.48	10.22	20.75	2.39	<0.001	1.07		

5.3.2.2 Association between behavioural components and Scent Sensitivity and Specificity scores

5.3.2.2.1 Dogs' scent sensitivity

The model was overall significant ($F_{(5, 21)}$ = 6.066, p < .005, adj. R2= 49.3%). Male dogs, and d older dogs, and those with higher 'Confidence' and those with higher 'Success in search' when tested, were significantly more odour sensitive during search trials (Table 5.5).

Table 5. 5 *Association of test components for dogs enrolled in detection projects at data collection cut-point (N=27) and their scent sensitivity scores **P <0.01. (Abbreviations' meaning in Table 5.2)*

Test components	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.59	0.49**
Constant	0.8	0.72	0.89	0.04	<0.001			
Dog's sex	-0.12	-0.20	-0.04	0.04	0.008	-0.42		
Age in months	0.002	0.00	0.00	0.00	0.028	0.37		
Playfulness	0.03	0.00	0.07	0.02	0.060	0.30		
Confidence	0.06	0.00	0.13	0.03	0.038	0.33		
Success in search	0.07	0.03	0.11	0.02	0.002	0.49		

5.3.2.2.2 Dogs' scent specificity

The model was significant ($F_{(8, 18)}$ = 8.938, p < 0.001, adj. R2 = 71%). Younger age, higher 'Food orientation', and 'Success at problem-solving', but lower 'Level of attention to handler', 'Self-control', ‘Interest in exploring environment’ and 'Success in were associated with higher scent specificity when discriminating odour samples (Table 5.6).

Table 5. 6 *Association between test components and demography and scent specificity scores from all dogs enrolled in detection projects at data collection cut-point (N=27) ***P <0.001. (Abbreviations' meaning in Table 5.2)*

Test components	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.80	0.71***
Constant	0.99	0.95	1.04	0.02	<0.001			
Dog's sex	-0.04	-0.08	0.01	0.02	0.082	-0.21		
Age in months	-0.001	-0.002	0.00	0.00	0.003	-0.40		
Food orientation	0.06	0.04	0.08	0.01	<0.001	0.73		
Level of attention to handler	-0.05	-0.07	-0.03	0.01	<0.001	-0.58		
Self-control	-0.05	-0.07	-0.02	0.01	0.002	-0.47		
Success at problem-solving	0.03	0.01	0.05	0.01	0.005	0.36		
Interest in exploring environment	-0.04	-0.07	-0.01	0.01	0.010	-0.39		
Success in search	-0.02	-0.05	0.00	0.01	0.040	-0.26		

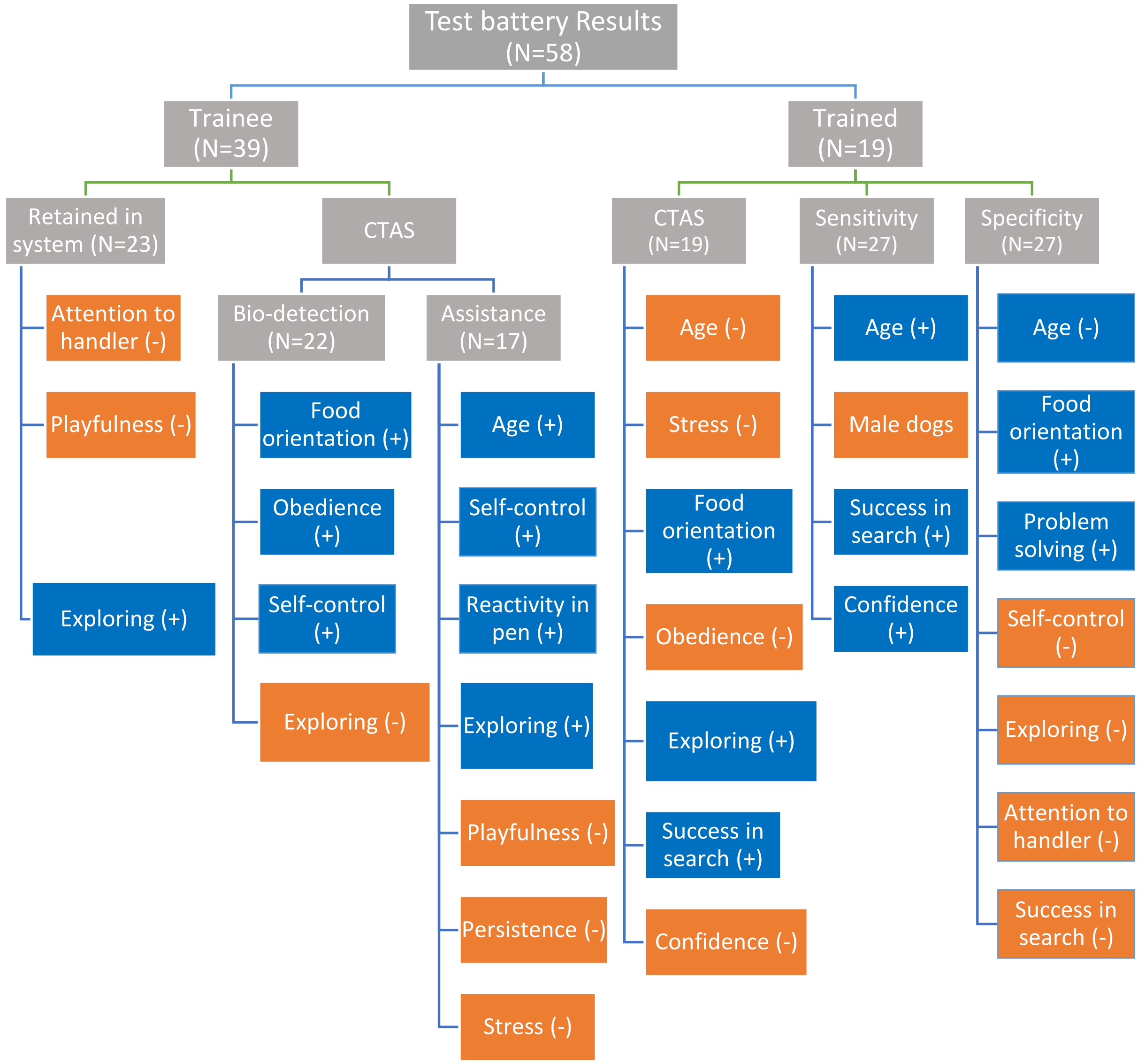


***Figure 5. 2** Overall results from test battery for each group. Significant relationships between Dogs' (N=58) test components and demographic factors with success measures.* The direction of the variable relationship with each measure is indicated. (+) = positive (-) = negative. 'Exploring'= 'Interest in exploring environment'.

5.4 Discussion

The current study aimed to investigate the association between dogs' behaviour during the test battery and MDD success in their roles. It assessed how this occurred for dogs at different training stages. In trainee dogs, it evaluated how dogs' behavioural characteristics correlated with their training outcomes and how these associated with the dogs' CTAS for each discipline. In trained dogs, it investigated how the tested behaviour related to their regular task performance and their scent sensitivity and specificity levels. Several behavioural components were associated with different success measures. This suggests that the test battery has the potential as a tool to predict MDD task success.

5.4.1 Association between the dogs' behaviour in test battery and training outcome

Three behavioural test components were significantly linked with a favourable training outcome: higher 'Interest in exploring environment', lower 'Playfulness' and lower 'Level of attention to handler' (Figure 5.2).

'Interest in exploring environment' described the dogs' natural tendency to investigate the physical environment, which may keep dogs constantly motivated when searching (Brady et al., 2018). In the current population, dogs that explore the environment more frequently may have a higher natural exploratory tendency, facilitating progress in scent training by reinforcing it. This attribute has been similarly linked to an increased likelihood of training success of explosives and drug detection dogs (e.g. Wilsson and Sinn, 2012; McGarrity et al., 2016; Lazarowski et al., 2019b).

Dogs with higher 'Playfulness' were more likely to fail training. In the test battery, trainee dogs were younger and more playful than trained dogs (4.4.3.2). Playfulness has been linked with enhanced social skills in pet dogs (Rooney and Bradshaw, 2002) and is relevant for the training of dogs for several searching tasks like explosives detection and search and rescue (Jones et al., 2004; Ganitskaya et al., 2020). However, excessive playfulness may come together with high excitability, especially in developmentally immature dogs (McPeake et al., 2019; Wallis et al., 2020). In this sample, more playful dogs had higher impulsivity scores (4.4.4), which at high levels, has been seen to be detrimental to the performance of working tasks, such as explosives detection dogs (Tiira et al., 2020).

Dogs with high levels of 'Level of attention to handler' were also more likely to fail. The survey suggested that the ability to bond with humans was important and ideally high for MDD but required at different levels across tasks (2.4.4; 2.4.5). High dogs' cooperation with their handlers, assessed through social referencing tasks, was linked with favourable task performance in explosives detection dogs (MacLean and Hare, 2018; Lazarowski et al., 2019b) and guide dogs (MacLean and Hare, 2018). However, the current MDD may be a skewed population potentially selected to be more attentive to the handler than other populations. Hence, those rejected are those showing this trait excessively. Dogs that rely overly on their handler may be less able to make independent decisions when working (1.2.6.2). When tested with tasks similar to the S1 ‘Exploring’ (4.2.5.1.1) and S2 'Ignoring' (4.2.5.1.2) subtests of the test battery (Rooney et al., 2003a), explosives detection dogs more inclined to interact with their handler over exploring the environment were significantly less successful in training. Comparable to the current findings. Some dogs may struggle when parted from their handler. For instance, guide dogs have shown changes in physiological stress-related parameters when isolated (Fallani et al., 2007).

The dogs' training success probability was significantly correlated to their CTAS. This suggests that a high CTAS may reasonably predict dogs' training outcomes. However, this is not surprising, as the trainers' opinion of the dogs' performance is critical when deciding the dogs' fate. The decision is subjective, and although the trainers' knowledge of the dog is deep, it may not fully reflect the dog's capacity as the trainer's opinion may be biased (Clark and Rooney, 2021). Even if the CTAS was calculated systematically (3.3.2.8), this measure is still subjective. Therefore, combining subjective and objective measures is fundamental to increasing validity. In future studies, basing the CTAS on the opinion of multiple raters instead of a single one could aid its reliability.

5.4.2 Association between the dogs' behaviour in the test battery and CTAS for each task

I assessed which behavioural components were linked with CTAS for each task. Some relationships coincided with components related to training outcomes (e.g. 'Playfulness' and 'Interest in exploring environment' for Assistance dogs), while others did not. Both success measures are correlated. However, CTAS offers a broader parameter in measuring ability than training outcome and has proven different relationships for each task.

The components' relationship with CTAS varied for each discipline. For prospective Bio-detection dogs, higher 'Food orientation', 'Obedience' and lower 'Interest in exploring environment' were significantly associated with higher CTAS. Whereas in Assistance dogs, older age, higher 'Interest in exploring environment' and 'Reactivity in holding pen', and lower 'Persistence', 'Playfulness' and fewer stress behaviours during the test were significantly linked with higher CTAS. 'Self-control' was significantly associated with CTAS for both disciplines (Figure 5.2).

High dog self-control has been related to good operational performance (e.g. Bray et al., 2015; Tiira et al., 2020). Studies in different dog roles have measured dogs' motor inhibition with detour tasks like the Cylinder test (Bray et al., 2017b; Tiira et al., 2020), similar to the current test battery (4.2.5.1.7). Fewer errors in detour tasks have been correlated with lower arousal in guide dogs' (Bray et al., 2015) and better-searching ability in explosives detection dogs (Tiira et al., 2020). In this study, 'Self-control' was relevant for Bio-detection and Assistance dogs. Less inhibited Bio-detection dogs may tend to make haste decisions leading to false indications in search trials. In Assistance dogs, self-control is fundamental when working in public places.

Like other detection tasks, Bio-detection dogs may be more active and interested in the environment. Hence it is surprising that higher 'Interest in exploring environment' was negatively related to CTAS but positively related for Assistance dogs. These trends may be associated with pre-selection. Extremely explorative Bio-detection dogs may struggle to concentrate when searching. Assistance dogs tend to be generally calm (Bray et al., 2021b) and reared and selected to be less involved with environmental stimuli. Still, a moderate interest in the environment within that threshold may benefit retaining motivation during training.

'Playfulness' was negatively significantly related to CTAS for Assistance dogs but not for Bio-detection dogs. High 'Playfulness' in Assistance dogs may be more disruptive for the task than for more active Bio-detection dogs.

High trainability is essential for dogs' detection training (Bray et al., 2021b) and for learning new scents. 'Obedience' was associated with Bio-detection dogs' ability. 'Food orientation' was related to ability in this population. More food-motivated dogs may be more responsive to training since they are regularly reinforced with food.

Older Assistance dogs had higher CTAS than younger dogs. Developmental maturity increases dogs' self-containment (Wallis et al., 2020) when performing (e.g. Walczak et al., 2012). Dogs showing fewer stress behaviours during the test had increased ability. According to findings from the survey (Chapter 2), Assistance tasks require high adaptability and lack of reactivity. This corresponds with studies identifying successful service dogs were as calmer and more obedient, with low excitability than detection dogs (Bray et al., 2021b). Hence, it is not surprising this study suggests that less playful, more self-controlled dogs who displayed fewer stress signals when tested may show a better aptitude for Assistance tasks.

Lower 'Persistence' was linked to higher CTAS in Assistance tasks. 'Persistence when alerting' was very important for Assistance tasks in the survey (2.4.5.1). However, here 'Persistence' involved dogs focusing on opening a container during the unsolvable task (4.2.5.2.3). It was also associated with impulsivity in trainee dogs (4.4.4). As discussed in Chapter 4 (4.5.6), whether persistence may be a positive or a negative trait for performance may depend on the context and its manifestation (Porritt et al., 2015; Hall, 2017; Tiira et al., 2020). Previous studies in explosives detection dogs and service dogs tested with the unsolvable task found that dogs who tended to pay attention to a human had a higher success probability than those that persevered in obtaining the food (e.g. MacLean and Hare, 2018; Lazarowski et al., 2019b). Here, social referencing may indicate a higher tendency to cooperate with humans (Wallis et al., 2020) and to shift to alternative strategies to obtain food (asking for human help) than persisting in infructuous behaviours. Thus, lower scores of the component 'Persistence' in Assistance dogs (as in here) may be positive since dogs with lower ‘Persistence’ may be more human-focused and more likely to notice changes in, e.g. glycaemia ranges.

Lastly, dogs who were more reactive when held in a pen were more skilled during assistance training. Reactivity in isolation has been considered negative for other roles (Rooney et al., 2003b; Fallani et al., 2007; Rooney et al., 2016). Highly able Assistance dogs tend to bond strongly with humans due to genetics and reinforcement (1.2.6.1). Hence, they may be aroused when isolated from humans. It would be beneficial to enhance habituation to being alone during early socialisation and training so that Assistance dogs can cope better with separation (Rooney et al., 2009).

5.4.3 Association between trained dogs' behaviour in the test battery and their CTAS

In trained dogs, higher 'Food orientation', ‘Interest in exploring environment’ and 'Success in search' but lower 'Obedience' and 'Confidence' and fewer stress behaviours were significantly associated with higher CTAS on their performance in regular detection trials (Figure 5.2).

'Food orientation' and 'Interest in exploring environment' were both associated with favourable training outcomes also in trainee dogs. These may contribute to maintaining dogs' ongoing performance when operational. Reinforcing the dog's exploring interest may stimulate extrinsic motivation to consolidate the memory of learned scents and to learn new ones, while low reinforcement decreases detection performance (e.g. Gazit et al., 2005). 'Energy and interest in the surroundings' were significantly relevant for military dogs' long-term operational effectiveness compared to dogs who were withdrawn early (Brady et al., 2018). This tendency seems similar in this trained Bio-detection dog population.

Several test batteries included searching tasks to assess detection dogs' ability. Dogs’ performance in these subtests was frequently linked with success in detection roles (e.g.

Svartberg, 2002; Rooney et al., 2003a; Maejima et al., 2007). Similarly, here trained dogs with high 'Success in search' were those who found a treat more frequently in a scent searching subtest (4.2.5.2.1). As expected, this component was related to dogs' higher ability in bio-detection scent trials, as high searching ability is fundamental for the performance of detection tasks (1.2.7).

'Obedience' was linked with higher CTAS in trainee Bio-detection dogs but negatively related in trained dogs. This seems counter-intuitive, as it would be anticipated that more obedient dogs would perform better when operational. A possible explanation is that disobedient dogs were filtered out after training, and those who progressed to projects were generally more obedient. However, dogs who rely unconditionally on their handler cues may take fewer chances on independent choices when indicating a scent (1.2.6.2). A certain level of "Selective obedience" is considered necessary for roles such as guide dogs (Knol et al., 1988; Audrestch et al., 2015). This may be relevant for this dog population too.

More confident dogs tend to be better at working tasks (1.2.6.1). Nonetheless, in trained dogs, 'Confidence' was negatively related to CTAS. The component 'Confidence' mainly included the dogs approaching more and displaying more behaviours towards an object in motion (4.2.8.5.8). Dogs approaching the object do not necessarily mean they were confident. Perhaps some were more reactive towards the toy due to negative motivation, e.g. fear. The dog instructors may perceive more reactive dogs as less proficient in their tasks.

Dogs who displayed more stress-related behaviours when tested tended to have lower CTAS. These dogs may have less cognitive and physiological ability to cope with new challenges from the test or to endure the whole procedure without getting anxious or frustrated. High-stress levels in working conditions may diminish task proficiency and harm dogs' well-being (Rooney et al., 2009; Foyer et al., 2016; Perry et al., 2017). The display of stress signals is mediated by different underlying emotions and is context-related (4.5.4). Here some of these behaviours (e.g. vocalising) may also be attributed to displaced behaviours from reward anticipation (Jakovcevic et al., 2013) and not from an underlying negative motivation. However, at high intensity, these may be task disruptive.

5.4.4 Association between behavioural components and scent sensitivity and specificity scores

Sensitivity and specificity indicate how accurate a screening method is in measuring what it is supposed to measure (Swift et al., 2020). These parameters refer to particular forms of error-making in decisions over uncertainty (Pastore and Scheirer, 1974; Heeger and Landy, 1997). High sensitivity indicates a higher tendency to minimise false negatives, and high specificity a higher tendency to minimise false positives (Lalkhen and McCluskey, 2008; Gadbois and Reeve, 2016) (1.2.1.2).

In the current sample, male dogs and those that were older had significantly higher sensitivity. Also, higher 'Success in search' and 'Confidence' were related to scent sensitivity. Factors significantly related to scent specificity included younger age, higher 'Food orientation', and 'Success at problem-solving, but lower 'Level of attention to handler', 'Self-control', ‘Interest in exploring environment’ and 'Success in search' (Figure 5.2).

Younger dogs were good at minimising false positives (specific), whilst older dogs were better at minimising false negatives (sensitive) in search trials. Studies suggest that younger dogs learn new tasks (including new scents) quicker than older dogs as they may be more reward-sensitive (Walczak et al., 2012; Wallis et al., 2020). However, may struggle more when progressing to more complex training (Rooney et al., 2004; Walczak et al., 2012). Hence, younger dogs may be good at identifying S+ and commit fewer false alarms in expectation of a reward in initial training trials where S+ is easier to distinguish from S-. The current sample had eight young dogs added at the end of the data collection, which recently progressed to detection projects. Older dogs may be more experienced with S+ and able to discriminate it even when attenuated or more similar to S- (Lynn and Barrett, 2014). However, since dogs work on different projects of varying complexity, it is difficult to make comparisons. Future research with larger samples trained for the same target scent and focusing on ontogeny changes and detection ability in MDD may help clarify this.

Dogs' scent sensitivity scores were significantly associated with the 'Success in search' component. As expected, more odour-sensitive dogs found treats more frequently in the

Boxes searching subtest (4.2.5.2.1). 'Success in search' was also correlated to higher CTAS. It makes sense that dogs that were faster at finding treats during this subtest may be also better at discriminating S+ and lessening false positives in search trials. Also considered by the instructors to have higher ability. However, more specific dogs had significantly lower scores on the 'Success in search' component. In the S9 Box search subtest, more sensitive dogs were possibly faster and indicated the correct treat location more quickly. Yet, more specific dogs may have been more ‘conservative’ (1.2.1.2), systematic, and took longer to determine the scent location, avoiding false indications.

Dogs who are more olfactorily sensitive had higher 'Confidence' scores. They may be more outgoing, sensorially reactive, and inclined to investigate new situations (Braem et al., 2017; Bray et al., 2021b). Bio-detection dogs with higher 'Confidence' that approach quickly new stimulus may also be more inclined to alert and not ignore S+ but sometimes more predisposed to false alarms looking to maximise rewarded indications (Gadbois and Reeve, 2016).

'Success at problem-solving' was associated with higher specificity. Studies assessing problem-solving aspects, such as short-term memory, motor inhibition and performance in a multistep task, have shown links with better performance in detection tasks (MacLean and Hare, 2018; Lazarowski et al., 2019b; Tiira et al., 2020). Dogs with less tendency to falsely indicate a scent may be better at processing and discriminating information relevant to decision-making.

However, dogs more attentive to their handler tended to be less specific. A dog highly inclined to follow human cues may undermine perceptual information and rely on the handler's behaviour and expectations, leading the dog to error, i.e. false alarms (Pfungst, 1911; Jezierski et al., 2014; Sumegi et al., 2014; Edwards et al., 2017; Lazarowski et al., 2020b). Also, some dogs may be distracted by the presence of people in the trial (Rooney et al., 2003b).

Self-inhibition is integrated by several elements and is considered context-specific (Vernouillet et al., 2018). The negative association between specificity and 'Self-control' was unexpected. It was anticipated that more ‘conservative’ (1.2.1.2) dogs with less inclination to false alarms might have higher self-inhibition. Less inhibited subjects, less aware of consequences when making decisions, maybe more 'liberal' (1.2.1.2), and inclined to err in false alarms (Martin et al., 1994; Lynn and Barrett, 2014; Gadbois and Reeve, 2016). This contradictory finding may be attributed to the high number of young dogs in the sample, which may be less self-controlled and more reactive than older dogs. Perhaps, dogs who progressed to detection projects may already have been selected to be more self-inhibited than those rejected, and less inhibited dogs are within a reasonable threshold in a generally active population such as Bio-detection dogs. Therefore, the relationship between self-inhibition and specificity should be explored further in larger samples tested with different tasks.

Dogs with a higher 'Interest to explore the environment' were significantly less odour-specific. This component has shown to be relevant for trained dogs' general aptitude as it was significantly associated with CTAS. However, dogs that strongly engage with the environment may struggle to focus on the target scents and err more frequently by indicating false positives. Conversely, dogs less inclined to engage with other environmental stimuli might be more focused and accurate when discriminating odours and less likely to alert falsely.

5.4.5 Limitations and future steps

The test battery appears to be an effective tool to objectively measure dog attributes relevant to different medical detection tasks, reflecting traits deemed important by professionals in the field and previous research. The test repeatedly assessed relevant characteristics across tasks and training stages, such as 'Food orientation'. Others were more important for specific populations, like 'Success in search' for trained dogs. Some traits were consistently negatively associated with the dogs' performance, like the display of stress behaviours. However, the direction of the relationship varied for different sub-samples, for traits such as 'Obedience' and 'Interest in exploring environment'. Dogs of different tasks and training stages may require different levels of each component.

The test may help to predict performance for different stages and tasks and to identify natural behavioural tendencies early before being trained for a specific discipline (Harvey et al., 2017; Bray et al., 2021b). For instance, more self-inhibited dogs, which make fewer mistakes in the cylinder task, and those who tend to gaze at humans more during cognitive tasks may be suitable for assistance roles. Dogs who tend to olfactory explore a room, those that find a treat quickly in a simple searching task, and those that retrieve treats from a puzzle may be better at bio-detection tasks. However, some of the subtests may not be useful for predicting performance. Hence these may be excluded. Here I focussed on main effects of the association between each test component and success measures. However, a future step would be to investigate how the interactions of these factors are associated with role success to enhance integrative profiling.

In trained dogs, the performance in the test battery may help to identify aspects to maintain their operational effectiveness, such as those related to 'Search success' and 'Success at problem-solving'. However, the extent to which the test components are associated with success may vary across dogs' scent projects and contexts.

Dog performance is influenced by individual variation and dogs' internal or external aspects, such as genetics (McGreevy et al., 2013; Raffan et al., 2016; Banlaki et al., 2017), early experience (Rooney et al., 2003b; McMillan, 2013); and environmental elements not considered here. These would be valuable for future MDD research.

Scent sensitivity and specificity are valuable objective measures of dogs' performance, each showing different aspects of decision-making relevant to effective scent detection. However, having small samples for the different projects prevents condition standardisation or intercomparison between projects. Therefore, future research may benefit from greater dog samples for different target scents.

The current test battery indicates potential as a predictive tool for MDD behavioural assessment. However, currently would not suffice for this purpose, as the study was exploratory of several measures, not all indicative of performance. Still, a confirmatory refined version, combined with validated subjective measures, assessing test reliability and validity (1.2.5), and evaluating larger MDD samples from different training institutions, may improve this test for practical application.

Another aspect to investigate is how dogs' making optimistic or pessimistic decisions under ambiguity may affect the dogs' task performance and how this is associated with the dogs' behavioural components from the test battery. Therefore, in the next chapter, I investigate this through a cognitive bias test.

5.5 Conclusion

The test battery showed significant associations between the test components and the success measures across training stages and tasks. This tool helps measure MDD behaviour more objectively to improve selection and identify variation in ability, scent sensitivity, and specificity, which is relevant to improving and maintaining working performance.

Considering these findings, a refined version of this test, including those more useful measures, may be applied to different MDD populations with larger samples to improve the usefulness and feasibility of the test.

Chapter 6. Cognitive bias in medical detection dogs:

Is the outcome of a Cognitive Bias test associated with dogs' performance in MDD tasks?

Abstract

This study investigated whether dogs' performance on a cognitive bias test (CBT) is associated with their aptitude as MDD. It focused on the relationship between dogs' 'optimism' or 'pessimism' when making decisions in an ambiguous situation and their behaviour in the test battery (4.4.1), the success measures determined in Chapter 3 (3.5) and their DIAS scores (4.4.4).

The dogs (N=58) were tested with a CBT to assess whether they displayed more 'optimistic' or 'pessimistic' judgements of ambiguous stimuli. Dogs were trained to discriminate between two locations, one being positively associated (bowl baited with food) and the other being negatively associated (empty bowl). Latency to approach an empty bowl placed intermediate to the conditioned locations (i.e. ambiguous cue) was measured. It was assumed that dogs approaching the intermediate cue faster expected a food reward and hence were behaving 'optimistically'.

The dogs completed training relatively quickly (mean=18.42 trials). Dogs with higher Composite total ability scores (CTAS) as rated by their trainers ($p= 0.039$) and those with higher DIAS scores ($p= 0.04$) had significantly shorter latencies to approach ambiguous locations. However, dogs that took longer to approach the bowl were more odour specific in their detection tasks ($p= 0.021$). There was no significant association with training outcome or the dogs' prospective discipline.

Older dogs ($p <0.001$), those with higher 'Confidence' ($p= 0.016$) and those with higher 'Playfulness' in the test battery approached ambiguous locations significantly faster. This tendency was marginally significant in dogs with higher 'Success at problem-solving' (p=

0.053) and 'Food orientation' (p= 0.058). Meanwhile, dogs with higher 'Total stress behaviours' took significantly longer to approach the tested locations (p= 0.012).

These findings suggest a relevant influence of dogs' 'optimism' / 'pessimism' when making decisions on different MDD performance aspects. Higher latencies in dogs who showed greater specificity in odour detection tasks may be attributed to their inherited tendency to discriminate odours, reinforced by their scent discrimination training and being more cautious when indicating an odour presence, i.e. erring on the side of a low level of false alerts. It would be worth further exploring the link between MDD ability and variation in 'optimism/pessimism' on decision-making using larger sample sizes

6.1 Introduction

Cognitive bias or judgement bias refers to the impact of affective tendencies (i.e. positively or negatively valenced emotional states) on information processing (Mendl et al., 2010b). Cognitive bias testing (CBT) focuses on an individual's propensity to judge an ambiguous situation as positive or negative depending on internal or external factors. This is thought to reflect their propensity to be more 'optimistic' (the tendency of an animal to treat an ambiguous location as likely to be rewarding) or 'pessimistic' (likely to be not rewarding or punishing) when making decisions under ambiguity (1.2.6.11).

The impact of emotions on different cognitive functions (i.e. attention, memory, and decision-making) has been explored in humans and various animal species (Mogg and Bradley, 1998; Harding et al., 2004; Paul et al., 2005; Mendl et al., 2009), including dogs (e.g. Mendl et al., 2010a; Burman, 2014; Kis et al., 2015). Different studies have investigated the relationships between decision-making under ambiguity and problems related to anxiety or stress or situations likely to generate these states (Mendl et al., 2010b; Titulaer et al., 2013), and affective states have been inferred from how the animals make a judgement. Studies have assessed changes in judgement bias in dogs with separation-related behaviours after pharmacological treatment (Karagiannis et al., 2015), in pet dogs after administrating oxytocin (Kis et al., 2017), and in companion dogs and dogs within a shelter after an olfactory enrichment programme (Uccheddu et al., 2018; Duranton and Horowitz, 2019). Overall, study findings showed reduced approach latencies to ambiguous locations suggesting more 'optimistic' like responses in dogs after pharmacological treatments designed to induce a relatively positive affective state, in line with predictions. In addition, other experiments focussed on the effect of chronic health issues on dogs' mood (i.e. longer lasting emotional states), including osteoarthritis, idiopathic epilepsy (Hobbs et al., 2020) and Syringomyelia in Cavalier King Charles Spaniels (Cockburn et al., 2018). These findings suggest how dogs make decisions in CBT is an indicator of wellbeing.

Other research has studied the relationship of CBT with different pet dog behavioural aspects (Duranton and Horowitz, 2019), dogs' age (Piotti et al., 2018), cognitive functions such as spatial memory (Gruen et al., 2019) and laterality (Wells et al., 2017a). However, to my knowledge, CBT has not been investigated in working dogs, including MDD.

In detection dogs, the influence of their cognitive bias may be relevant for decision-making and error tendency during search trials (1.2.1.2). Addressing signal detection theory (McNicol, 2005), a highly scent-sensitive (number of true positives divided by the total number of scent presentations) 'Liberal' dog may frequently anticipate the presence of the target scent and hence maximise alerts to increase the chance of a 'hit' (true positive) (Lalkhen and McCluskey, 2008; Gadbois and Reeve, 2016). They may minimise false negatives but also have a higher tendency for false alarms. A highly scent-specific dog may be more 'Conservative' to avoid errors when indicating scents but also miss targets (Lalkhen and McCluskey, 2008; Gadbois and Reeve, 2016). Potentially, these decision-making tendencies may be reflected in CBT. More 'Liberal' dogs in scent trials may also be more 'optimistic' during CBT trials, anticipating a positive outcome in ambiguous locations and hence approaching faster, while dogs that are more 'Conservative' in detection tasks may take longer to approach ambiguous locations during CBT. If this is the case, CBT may be a valuable paradigm to understand better and predict decision-making inclinations in detections dogs and particularly in MDD. However, this remains unexplored.

Impulsivity plays an important role in dogs' decision-making. In humans, more impulsive individuals tend to make hasty choices to obtain immediate rewards, which is counterproductive in the longer term (Moeller et al., 2001). Higher impulsivity in dogs has been linked with negative emotional states like frustration, sometimes leading to aggressive responses (Fatjó, 2001; Wright et al., 2012). There has been one study investigating how CBT associates with impulsivity in companion dogs (Hale, 2021). It showed that dogs with higher scores of the DIAS factor 'Responsiveness' approached significantly faster to an ambiguous location, and those that scored higher on the 'Aggression/Neophobia' DIAS factor had less likelihood to complete the CBT. In MDD, an increased tendency to be impulsive may affect the dogs' scent alerting decisions in olfactory discrimination tasks. However, to date there has not been any research investigating the link between decision making in CBT and impulsivity in MDD, nor how this relates to detection performance.

In humans, the tendency to take the initiative and try different strategies to achieve goals has been associated with positive achievements (Grant et al., 2011). This cognitive function has been identified to be relevant in working dogs (Bray et al., 2021b), and in Diabetes alert detection dogs (DAD), 'Willingness to try new behaviours' was reported to be linked with more accurate alerting of out-of-range events in clients with type I diabetes (Rooney et al., 2019). Trainers here also rated this trait as highly important for MDD performance (2.4.5; 3.6.9). It may be speculated that dogs that tend to judge ambiguous situations with 'optimistic' outcomes may be more proactive and try new things to achieve a reward than more 'pessimistic' dogs, which may tend to avoid situations perceived as unfavourable. However, these links have not been researched and may have a relevant impact on MDD performance.

The present study aims to assess if there is a relationship between dog learning and decision-making in CBT and their performance as an MDD, regarding their behaviour and scent discrimination decision tendenciessince CBT may be a useful independent measure of learning and decision-making propensities in the dog. For these, I assessed the association between the dogs' performance in CBT and the MDD success measures (3.5). I also investigated how the dogs' demography, behaviour in the test battery (4.4.1), and impulsivity scores from the DIAS (4.4.4) were linked with their performance in CBT. In addition, I assessed if the trainers' ratings for 'Willingness to try new behaviours' were associated with the dogs' performance in CBT. Specifically, I investigated:

a. Dogs with higher CTAS will have shorter latencies to probe locations.

b. Dogs that remained in the system will run faster towards the probe locations than those that failed.

c. Dogs with higher scent sensitivity = higher true positive / (true positive + false negative) will tend to approach faster to the probe locations, which may indicate a willingness to 'risk' false positives and hence be relatively 'optimistic'.

d. Dogs with higher scent specificity = higher true negative / (true negative + false positive) will tend to approach slower to the probe locations, which may indicate a low willingness to 'risk' false positives and hence be relatively 'pessimistic'.

e. Dogs with higher overall impulsivity scores in the DIAS will tend to run faster toward the probe locations.

6.1.1 The cognitive bias paradigm

The CBT was developed initially for rats (Harding et al., 2004), trained to press a lever when hearing a tone related to a positive outcome (obtaining a food reward) or to avoid pushing a lever associated with an aversive one (white noise) when exposed to a different tone. After exposing them to a predictable/unpredictable environment, their tendency to press levers upon ambiguous tones was tested to evaluate whether rats judged the tones as rewarding, suggesting 'optimistic' responses or punishing, reflecting 'pessimistic' judgement. Rats housed in unpredictable environments were significantly slower and performed fewer lever presses under ambiguity, and had fewer pressing responses. Burman et al. (2009) modified this paradigm to a spatial judgement bias task for rats, which was adapted for dogs by Mendl et al. (2010a). Instead of auditory cues, subjects are trained to discriminate between a "positive" baited location (bowl with food) and a "negative" one (empty) in the spatial judgement bias test. Dogs' latencies to approach three ambiguous probes (non-reinforced) located between the conditioned positive and negative locations are measured to test decision-making. If dogs behaved as if the bowls contained food, i.e. exhibiting short latencies to approach, this would indicate an 'optimistic' response. If dogs responded to the ambiguous bowl locations as if the bowls were empty, i.e. showing a longer latency to approach, this would indicate a 'pessimistic' response (Mendl et al., 2010a).

In the original study (Mendl et al., 2010a), the handler led the dog behind a visual barrier while the tester baited and placed the bowl between trials. However, restraining the dog behind the barrier could be stressful or frustrating. Therefore, Hale (2021) designed a wooden screen that allows the tester to move behind it freely and place the plate in different locations so the dogs cannot watch the tester. The screen consisted of three panels, with five small doors at their base for each test location (i.e. two conditioned and three ambiguous locations), allowing the tester to slide the bowl through the door so the dog could see its position. This way, dogs could not see the baiting event whilst being restrained, which reduced the possibility of frustration before testing and avoided dogs using unconscious cues from the tester or the handler. The improved experimental setup shortened the time to prepare each test trial. Hence, the present study used the same apparatus and methods adapted from Hale (2021).

6.2 Methods

6.2.1 The test arena and apparatus

The experiment was conducted in the same experimental room as the test battery (4.2.4.1). The wooden screen structure (Hale, 2021) was 3m in diameter and consisted of three panels with 25 cm wide doors for each of the five possible bowl locations (Figs 6.1 and 6.2). The side panels had two doors at their base. The central panel contained a single door. Each door opened backwards and had a small fabric handle in the middle. A 1.5 m cord was attached to each handle to move the door whilst the operator remained standing.

The negative (N) and positive (P) locations were at the far ends of the screen (Figure 6.2). Their locations were counterbalanced for half of the dogs. The 'Middle location' (M) was in the central panel, the 'near positive' (NP) was by P, and the 'near negative' (NN) was by N. Each bowl location was marked 20 cm in front of the door with a discrete X on the carpet to indicate where the tester had to place the bowl.

The tester remained behind the screen during the whole procedure and was therefore not visible to the dogs during testing. A CCTV system (Swann®) mounted on the ceiling pointing towards the CBT screen allowed the tester to observe the dogs on a monitor recording the dogs' latencies to reach the bowls in real time. Another camera recorded a central room view, and a Canon® video camera was placed as in chapter 4.2.4.1 to register the dogs' behaviour during the test and revisit recordings for corrections after data collection was completed. The handler (same dog trainer present in the test battery) and dog waited in the holding area 4m from the test apparatus (thus 4m distance between the dog and the bowl locations) before each trial (Figure 6.1 and 6.2).

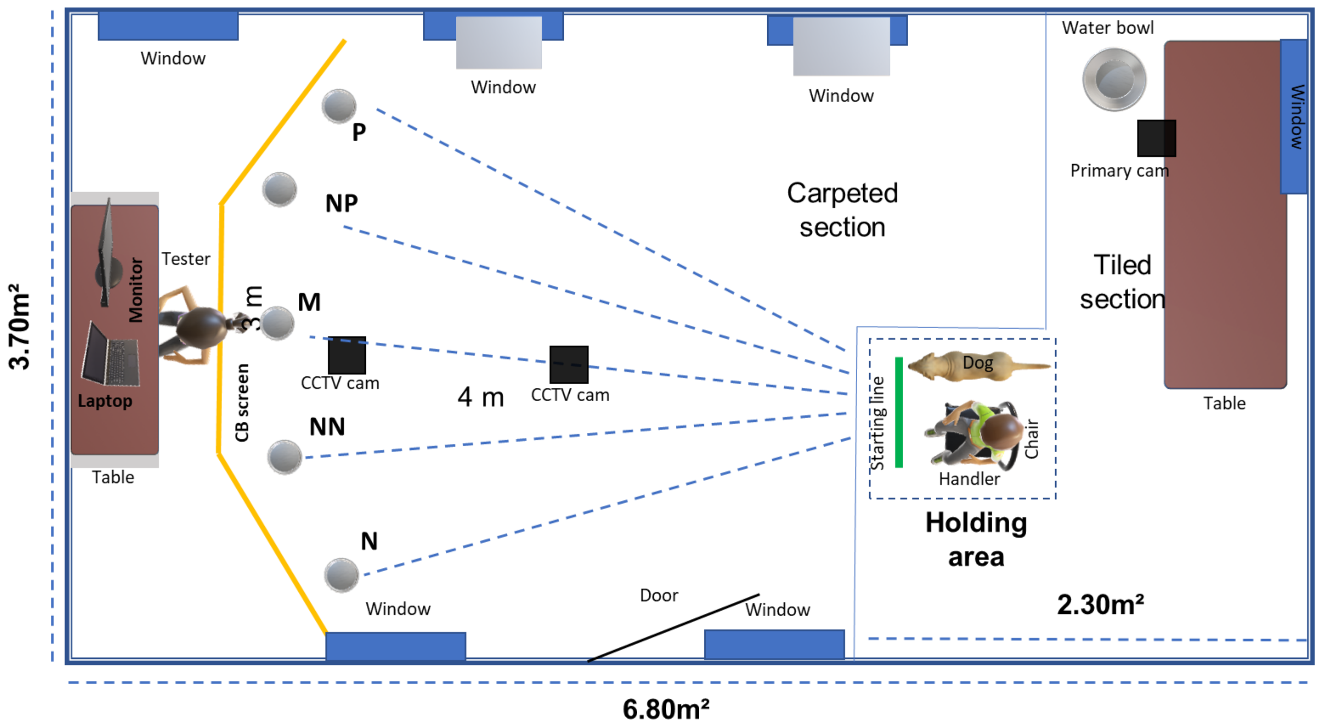


__Figure 6. 1__ Experimental room with the CBT arrangement with room measurements, the items utilised in the CBT and their location, and the participant's initial position. Cam=Camera. Bowl locations on CB screen: P= Positive; NP= Near positive; M=Middle; NN=Near negative; N=Negative.

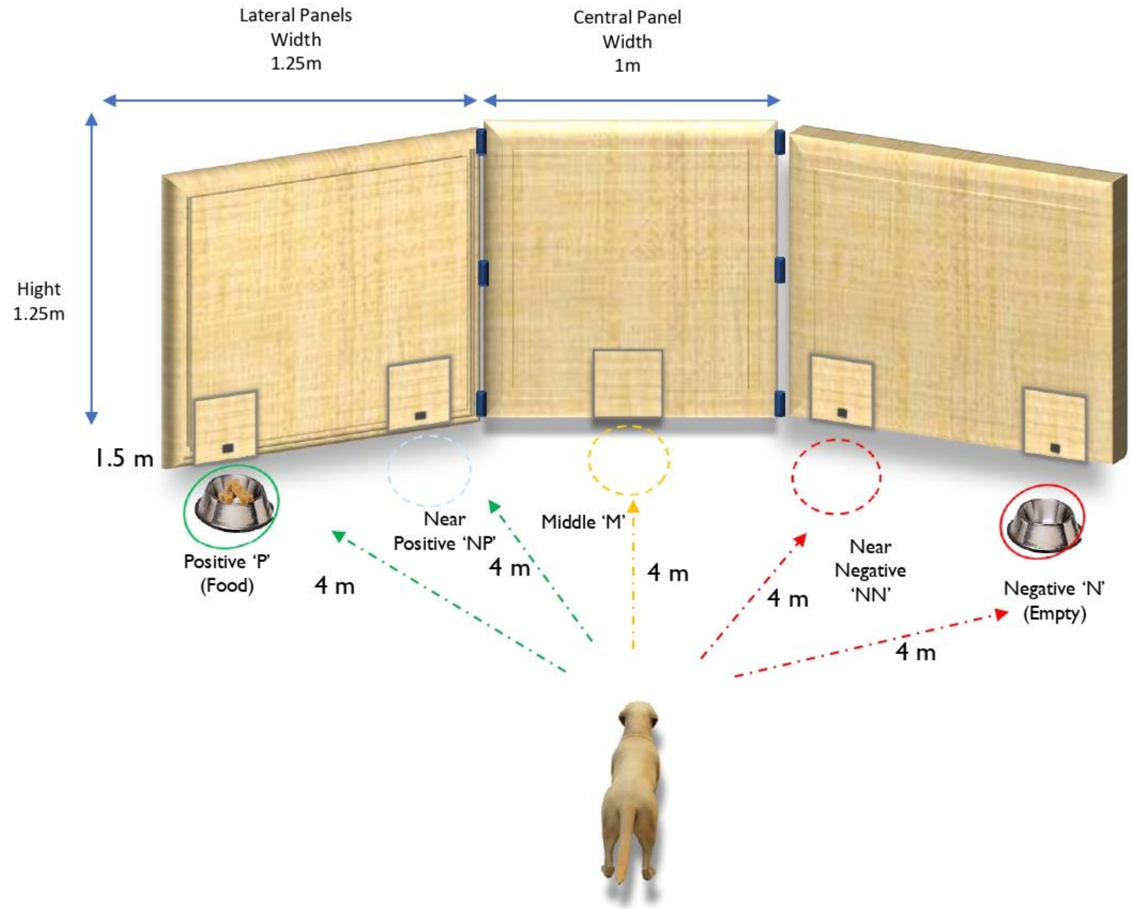


Figure 6. 2 *CBT apparatus containing five possible bowl locations. The positive 'P' (marked in green) and negative N' (marked in red) locations on either side of the wooden panel were the conditioned locations. The other three bowls in between those locations presented ambiguous test locations.*

6.2.2 The CBT procedure

The same dogs described in Chapter 3 (3.2.2; Appendix 4) and tested in the test battery (4.2.3) were also evaluated on the CBT. First trained to discriminate between positive and negative locations before being tested with ambiguous locations. Next, I explain the overall procedure involving the training and testing phases.

CBT always took place after the test battery. There was a minimum of 120 minutes of rest (mean 123, max =132 SD=1.6) between the end of the test battery and the start of the CBT to ensure a break between tests. The CBT lasted 52.4 (min=45.2, max=58.2 SD=1.3) minutes on average.

On arrival, dogs were allowed to freely explore the room for 5 to 8 minutes while the tester explained the test procedure to the handler. However, animals were presumably already habituated to the room before the CBT since they were in the same experimental room during the test battery. Hence, it is likely that the dogs had a positive association with the room since they previously received food rewards there.

The tester remained behind the screen from after the habituation until the end of the procedure. On each training and testing trial, the tester pushed a 20 cm diameter plastic bowl out through one of the doors in the screen while the handler held the dog in the holding area. When placed at the P location, the tester baited the bowl with a single piece of Royal Canin Energy® treat (1.5 cm) as a higher value reward (4.2.4.2). In the N location and the Probe locations, the bowl remained empty. However, for all repetitions during training and testing, the tester placed a treat in the bowl to produce the sound of the baiting event and later gently removed it if the bowl should be empty.

When the bowl was in position, the tester stood behind the middle of the screen, facing away from the dog in front of the monitor. The tester said "Ready", and the handler released the dog with their regular "Go" cue (as in several subtests from the battery test, e.g. Puzzle, Arm-pointing; hence no further familiarisation with this procedure was necessary). However, the handler was told not to encourage the dog any further. The tester timed the dog's latency to reach the bowl with a stopwatch. Each trial was finished by the tester saying "Stop" when the dog reached the bowl (the subject's first approach to <10cm from the bowl) or after 30 seconds from the dog's release, regardless of whether they approached the bowl or not. Then the handler recalled the dog or took it back to the holding area for the subsequent trial. The tester recorded each latency in real-time in an excel sheet modified from Hale (2021).

The CBT involved two phases:

6.2.2.1 Training phase

Dogs were trained to approach from a starting location and discriminate when a food bowl was in a 'Positive' (P) baited (food) or a 'Negative' (N) (empty) associated location. The initial training consisted of a minimum of 15 trials. The original CBT (Mendl et al., 2010) allowed a maximum of 50 attempts to achieve learning. However, this was reduced to 35 to avoid stressing or exhausting the dogs with overlong procedures. The training was considered complete when the dog had shorter latencies for three consecutive P presentations than each of the previous three N trials. Dogs were excluded from the test after the maximum number of trials if they had not reached training criteria or if the trainer considered them overstressed or fatigued.

Each training session started with two bowl presentations in the P location, followed by two in the N location. The subsequent training trials interspersed both locations pseudo-randomly (positions already set in the recording sheet with an equivalent number of both N and P). There were no more than two consecutive presentations for each bowl location (Figure 6.3).

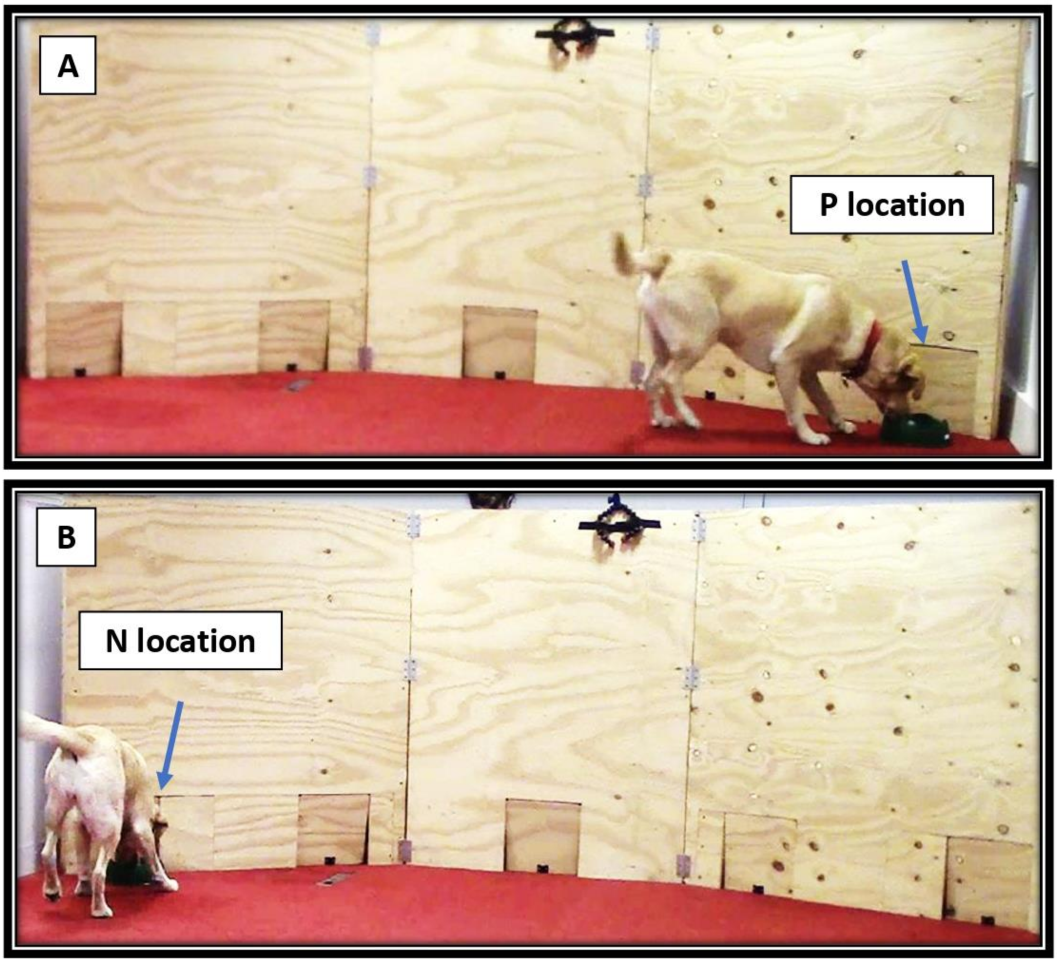


Figure 6. 3 Dog trained to associate the Positive 'P' location with a food reward (A) and the Negative location 'N' with the absence of reward (B). Locations on opposite sides for half of the dogs.

6.2.2.2 Testing phase

After the dogs reached the training criterion, they progressed immediately onto testing. An empty bowl was positioned at one of three ambiguous or probe locations between the P and N 'Near Positive' (NP), 'Middle' (M) and 'Near negative' (NN) to investigate how rapidly dogs approached the location as a measure of their responses to ambiguous cues. There were three trials for each probe location (nine in total) as follows M, NP, NN, NP, NN, M, NN, M, NP, with two P and two N training trials in between each probe cue (no more than two successive for each) to maintain the same reinforcement schedule as applied in Mendl et al. (2010a). These trials had a pseudo-random order also set in the recording sheet. In total, there were 41 test trials and a final trial with an empty P bowl to check whether the dogs were using odours to locate food. All latencies were collected during the test and subsequently analysed statistically. The procedure was stopped if a dog exceeded the maximum of 30 sec without approaching the bowl for more than five sequential trials independently of the location (Figure 6.4).

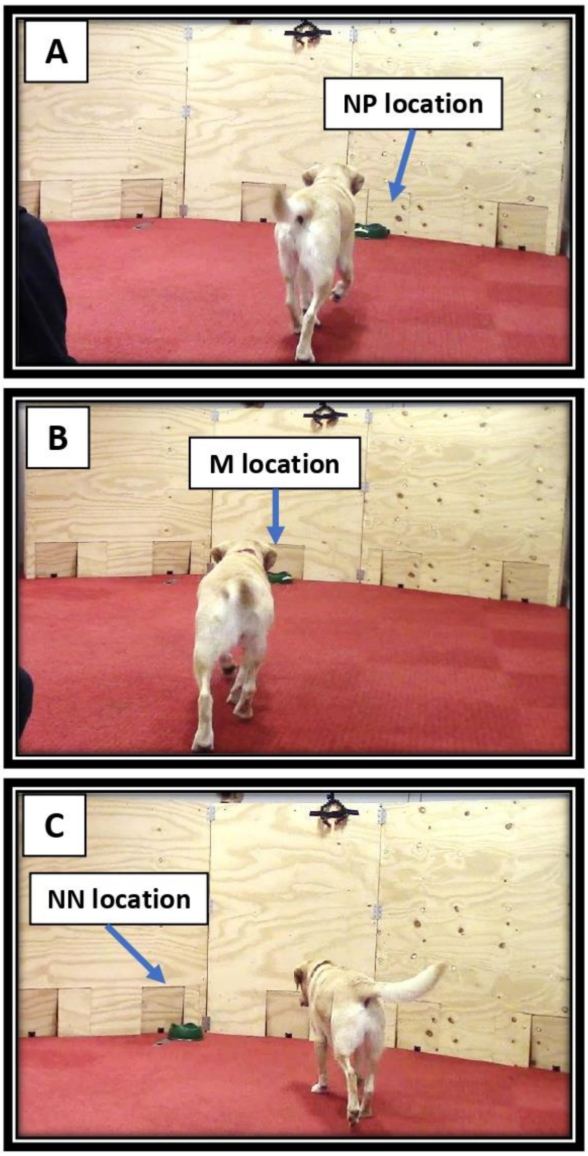


Figure 6. 4 *CBT dog testing Latencies to approach ambiguous locations Near positive (NP), Middle (M) and Near negative (NN).*

6.3 Data analysis

All dogs achieved the training criterion, and all completed the cognitive bias test except for two: one of them stopped approaching the bowl after six probe trials, exceeding the maximum of 30 sec for more than five successive trials. This dog tended to have long latencies or not visit ambiguous locations. The other dog achieved eight probe trials but reacted aggressively towards the handler taking it back to the holding area. These dogs were excluded from further analysis.

The data were analysed statistically using IBM® SPSS® software. The data were not normally distributed according to Shapiro-Wilk normality tests. Therefore, non-parametric tests were performed for the majority of the analyses. As in previous chapters, regression models were conducted to reduce multiple testing. Although this approach may not generally be best suited to non-parametric variables is useful for exploratory examination of the associations between dogs' CBT performance and MDD success and test battery measures, focussing on main effects.

6.3.1 Training phase analysis

The mean latency to reach the food bowls was measured for the last three training trials for P and N locations separately. The difference between P and N mean latencies were calculated to assess the degree to which dogs discriminated between locations after training. I performed Spearman's Rank correlations to investigate if the difference between mean N and P latency for each dog was linked with the number of sessions they took to reach the discrimination criterion. In addition, I tested whether the difference between mean N and P latency for each dog was correlated with their trainers' ratings for 'Ability to learn from being rewarded' (3.4.2).

6.3.2 Testing phase analysis

6.3.2.1 Differences in latencies between locations

I quantified the mean latency to approach the food bowls for each position (P, NP, M, NN and N) during testing. I performed Friedman tests to investigate variations in mean latencies to the different bowl locations, followed by post-hoc pairwise Wilcoxon tests to examine if there were significant differences between each of the bowl locations (The statistical software performed Bonferroni correction for multiple tests automatically).

Since the dogs were individually different, I calculated an adjusted latency score controlling for the dogs' mean latencies to P and N locations during the testing phase. To determine each dog's adjusted score, I used the following formula based on Mendl et al. (2010a):

$$\text{Adjusted score} = \frac{(\text{mean latency to probe location} - \text{mean latency to positive location})}{(\text{mean latency to negative location} - \text{mean latency to positive location})} \times 100$$

6.3.2.2 Effect of scent cues on dogs' latencies to approach locations

To rule out the possibility that dogs approaching the bowls were influenced by scent cues instead of discrimination learning (Mendl et al., 2010a), I performed an additional P trial at the end of the testing phase with an empty bowl. I compared it with the dogs' mean speed to P location during the testing phase using Wilcoxon signed-rank test.

6.3.2.3 Strength of discrimination coefficient

I derived a 'Strength of discrimination coefficient' to quantify the dogs' ability to distinguish between the three probe locations. I calculated a regression equation based on the three adjusted scores for each dog using the following equation (suggested by Mendl, 2021) with (y) at the mid-point:

y= bx + A; where x=location of the ambiguous cue, b=slope, A=intercept at x=0

This equation resulted in a linear regression where the slope of the curve (b) provides information on how sensitive the dog is to proximity to P of the ambiguous locations. A steeper slope means that the dog appears to distinguish locations effectively (e.g. going much slower to NN and much faster to NP). In contrast, a shallow slope indicates the dog does discriminate well between them (e.g. running speed to NN and NP are relatively similar).

I used this strength of discrimination coefficient to determine if the scent specificity score was associated with the dogs' tendency to distinguish across probe locations during the test with Spearman Rank correlations.

6.3.2.4 *Association between dogs' behavioural and demographic characteristics and CBM*

I assessed the association between dogs' cognitive bias measures (CBM) with MDD behavioural and performance aspects with multiple regression models, similar to previous chapters (4.3.2). I used the adjusted latency to each probe location since each might provide different information on the subjects' tendencies over ambiguity (Mendl et al., 2010b).

For all dogs (N=56), I assessed the association between behaviour and task performance and CBM. Each model assessed each CBM separately as the dependent variable. As independent variables, each model included dogs' demography (i.e. sex, age, training status) CTAS (3.4.2.7), DIAS scores (4.4.4), the 11 behavioural components (4.4.1), and total stress behaviours from the test battery (4.4.2). The trainers' ratings of 'Willingness to try new behaviours' were included (3.4.2) since ratings on this trait might be associated with reduced latencies to approach probe locations.

6.3.2.4.1 *Association between CBM and trainee dogs' prospective discipline and training outcomes*

For the trainee dogs (N=39), I assessed if their CBM varied across disciplines and training outcomes. The models included training outcomes (whether dogs remained in the system or failed training; 3.5.1) and prospective tasks (whether dogs were intended for assistance or bio-detection tasks) as independent variables and each CBM as the dependent variable.

6.3.2.4.2 Association between dogs' CBM and trained dogs' scent sensitivity and specificity scores.

I examined if the trained dogs' (N=25; all dogs with scent sensitivity data except two that did not complete the test) scent sensitivity and specificity in their regular detection tasks were related to their CBM. The models included scent sensitivity and specificity scores as independent variables and each CBM as the dependent variable.

6.4 Results

6.4.1 Training phase

The dogs took an average of 18.42 trials (min=15, max=35, SD=4.4) to reach the training criterion and progress to the testing phase. The mode was 15 trials; only two dogs took >30 trials to achieve discrimination criteria. Mean latencies to P ranged from 1.5 to 4.5 sec (mean= 2.46, SD = 0.59) and to N from 2.30 to 29.30 sec (mean= 8.25, SD = 6.78). The mean discrepancy between P and N was 5.8 sec (min= 0.26, max= 26.91, SD= 6.5). The difference between P and N mean latencies was significantly negatively associated with the number of trials to achieve the training criterion (R=-0.437, p=0.001), suggesting that dogs reaching the training criterion faster showed greater discrimination between N and P locations in the training phase. The difference between mean N and P latency was also significantly correlated with their trainers' ratings on 'Ability to learn from being rewarded' (R=0.412, p=0.009) but not with the number of trials to reach criterion.

6.4.2 Testing phase

6.4.2.1 Differences in latencies between locations

The latency to reach the bowl was significantly affected by location (Friedman's test: N=56, $\chi^2(2) = 176.57$, p <0.001), with significant differences between each of the bowl locations, except for P vs NP (Figure 6.5)

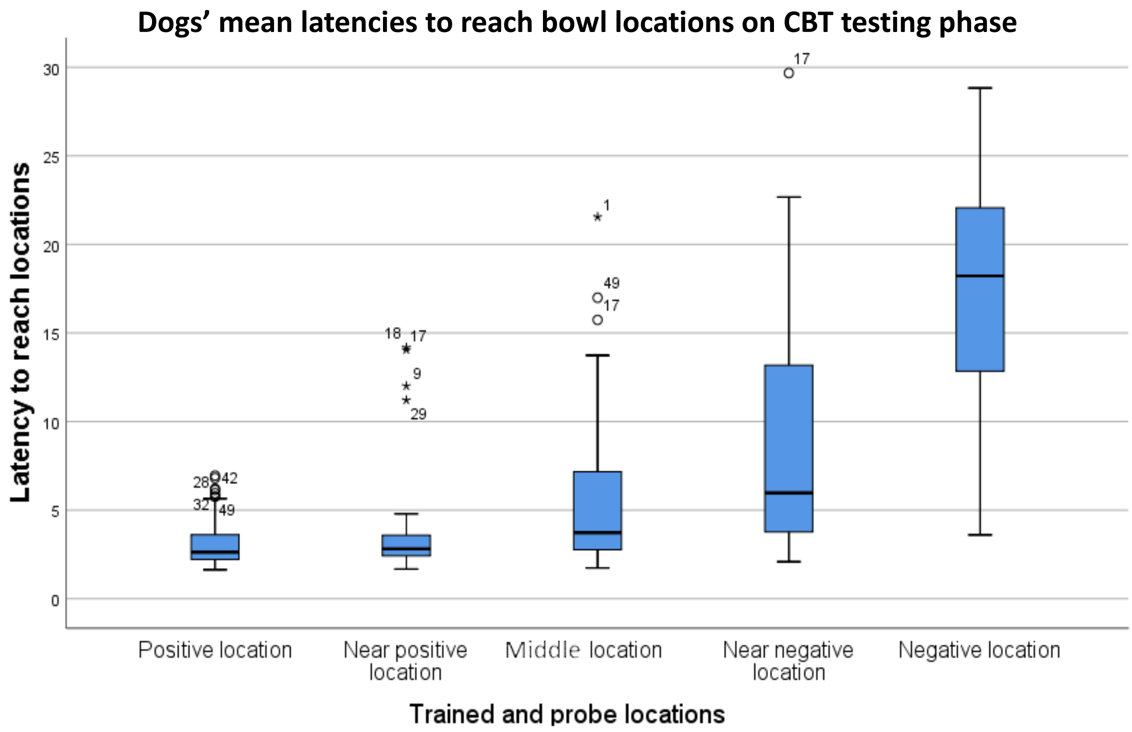


Figure 6. 5 *Distribution of mean latencies to approach each bowl location during testing (in secs) (N=56). Boxes show the median (bar within the box), the 25th interquartile (lower box border) and the 75th interquartile (upper box border). The whiskers indicate the minimum and maximum mean latencies, and the circles and stars represent outliers.*

6.4.2.2 Effect of scent cues on dogs' latencies

When comparing the dogs' latency to approach P with an empty bowl (mean= 2.93sec, +/- 0.8) and the mean latency towards P during the testing phase (mean=3.15 sec, +/- 1.41), no significant difference was found (z= -0.2, p= 0.842), this suggests that there was no influence of odour cues on the dogs' speed to approach locations.

6.4.2.3 *Association between dogs' behavioural and demographic characteristics and CBM*

6.4.2.3.1 NP adjusted latencies scores

The regression model was overall significant ($F_{(12, 43)}$ = 3.204, p < 0.001, adj. R2 = 30%). Older dogs, those with higher DIAS scores, and dogs with higher 'Confidence' scores were quicker to approach the NP location. There were also close to significant tendencies for dogs with stronger 'Food orientation' and those with greater 'Success at problem-solving' to approach NP faster (Table 6.1).

Table 6. 1 *Association between the dogs' demography and behavioural characteristics (N=56) and their latencies to approach NP. ***P <0.001.*

Independent variables	ß	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.37	0.3***
Constant	28.18	10.11	46.24	9.00	0.003			
Age in months	-0.23	-0.35	-0.11	0.06	0.000	-0.46		
Food orientation	-2.98	-6.07	0.11	1.54	0.058	-0.22		
Confidence	-3.84	-6.91	-0.76	1.53	0.016	-0.29		
Success at problem solving	-2.94	-5.91	0.04	1.48	0.053	-0.23		
DIAS	-33.18	-64.78	-1.59	15.73	0.040	-0.25		

Note. B= unstandardised regression coefficient; CI=confidence interval; LL= lower limit; UL=upper limit; SE B= standard error of the coefficient; β=Standardised coefficient; R^2= Coefficient of determination; ΔR^2= Adjusted R2.

6.4.2.3.2 M adjusted latencies scores

The analysis derived a significant model ($F_{(4, 51)} = 4.377$, $p < 0.01$, adj. R2 = 20%), indicating that dogs with a higher CTAS and those with higher 'Playfulness' tended to approach the M location significantly faster (Table 6.2). The association with the remaining factors was not significant.

***Table 6. 2** Association between the dogs' demography and behavioural characteristics (N=56) and their latencies to approach M (Abbreviations meaning in table 6.1.) **P <0.01.*

Independent variables	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.26	0.2**
(Constant)	29.62			6.11	<0.001			
Total ability score	-0.32	17.36	41.87	0.15	0.039			
Playfulness	-8.47	-0.63	-0.02	3.37	0.015	-0.27		
Reactivity in holding pen	-5.55	-15.23	-1.71	3.23	0.092	-0.31		
Food orientation	5.57	-12.04	0.95	3.29	0.096	-0.21		

6.4.2.3.3 NN adjusted latencies scores

The model was significant ($F_{(1, 54)} = 6.771$, $p < 0.05$, adj. R2 = 10%) and indicated that dogs with higher total stress scores moved significantly more slowly to NN (Table 6.3).

***Table 6. 3** Association between the dogs' demography and behavioural characteristics (N=56) and their latencies to approach NN (Abbreviations defined in Table 6.1) *P <0.05.*

Independent variables	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.11	0.1*
(Constant)	11.93	-14.55	38.41	13.21	0.370			
Total stress	1.32	0.30	2.33	0.51	0.012			

6.4.2.4 Association between dogs' CBM and training outcome

Neither the trainee dogs' training outcome nor their prospective task was significantly linked to CBM.

6.4.2.5 Association between trained dogs' CBM and scent sensitivity and specificity scores.

Only the model assessing the link between trained sensitivity and specificity scores and latency to M was significant ($F_{(1, 23)} = 6.126$, $p < 0.05$, adj. $R2 = 18\%$). Dogs that move more slowly towards M tended to have higher scent specificity in their detection tasks (Table 6.4). However, there was no significant relationship between sensitivity and CBM.

Table 6. 4 *Association between trained dogs' scent specificity scores (N=25) and their latency to approach M (Abbreviations meaning in table 6.1) **P <0.01.*

Independent variables	B	95.0% CI for B		SE B	*P*	β	R2	ΔR²
		LL	UL					
Model							0.21	0.18*
(Constant)	-129.73	-257.71	-1.76	61.86	0.047			
Scent specificity	171.57	28.17	314.97	69.32	0.021	0.46		

6.4.2.6 Association between trained dogs' scent specificity and strength of discrimination coefficient

The dogs' scent specificity scores were significantly associated with their strength of discrimination coefficient (R= 0.438, p= 0.028)

6.4.3 How the test components and the CBM are associated with MDD success measures?

To assess what best predicts overall MDD performance, I carried out multiple regression models with each success measure: Training outcome, CTAS, Scent Sensitivity and Specificity scores as the dependent variables, to investigate how the 11 behavioural components from the test battery (4.4.1), the DIAS and the CBM interacted to predict overall MDD performance. Each model included as independent variables the test battery components (those significantly associated with each success measure in Chapter 5), the DIAS scores (4.4.4) and CBM. However, those variables with significant links to dependent variables were very similar to those identified in previous chapters and did not include associations with CBM. Thus, these results were not discussed further.

6.5 Discussion

6.5.1 Is the outcome of CBT associated with dogs' performance in MDD tasks?

This chapter investigated the relationship between the behaviour and performance of MDD and their performance in a CBT. All dogs learned the task and achieved generalisation on their response across the different locations. Most dogs completed the test (only two out of 58 were excluded). This contrasts with previous studies reporting high levels of dog exclusion (e.g. Muller et al., 2012; Hobbs et al., 2020; Hale, 2021). The main findings of this study indicate an association between dogs' performance in detection tasks and their latency in approaching ambiguous locations in the CBT, reflecting the propensity of 'optimistic' and 'pessimistic' judgement on uncertainty.

When examining performance measures, dogs with higher CTAS approached the ambiguous M location significantly faster. Dogs with higher scent specificity in the detection tasks took longer to approach M and had higher strength of discrimination coefficients. These associations indicate that dogs' decision-making under ambiguity is associated with some aspects of their performance in MDD tasks. It may be that the way that the dog makes a decision relating to an ambiguous stimulus is directly linked to facets of behaviour that impact performance measures. However, these relationships may also be mediated by the dogs' affective state, unrelated to their ability.

Some test battery measures were associated with the dogs' behaviour in CBT. Older dogs and those with higher 'Confidence' approached NP faster. Dogs with higher 'Playfulness' approached M faster. Also, dogs who showed fewer stress-related behaviours during the test battery ran to NN more quickly.

The dogs' performance in CBT was also significantly linked with their levels of impulsivity. Dogs with higher DIAS scores approached NP faster. More impulsive dogs may respond quicker when expecting an ambiguous location to be rewarding than dogs with lower impulsivity scores.

6.5.2 Association between dog's performance in CBT with measures of success in MDD tasks.

Dogs more likely to judge the ambiguous location as positive also scored higher on CTAS in the MDD task. This could be due to some aspect of 'optimistic' decision-making associated with good performance in detection tasks, a more positive emotional state (as indicated by CBT) linked with good performance, or some unknown factor generating an apparent association between CBT and MDD.

There was no significant association between CBT performance and training outcome in the trainee dogs. Hence CBT may not predict whether they pass or fail training. However, the lack of association might also result from this study's relatively small sample size. CTAS was associated with shorter latencies for the whole dog sample. Still, training outcome may be a less sensitive parameter of aptitude in trainee dogs since it is a binary measure. Future studies using larger samples of trainee dogs may show different results.

More scent-specific trained dogs may have taken longer to approach M since they judged the probe location as negative. However, it is relevant to ask whether this happened because they are more 'pessimistic' or more cautious and systematic when making decisions. This may imply that over-optimistic dogs tend to be more ‘liberal’ when discriminating odours (1.2.1.2), generalise the scent's presence, and have a higher tendency to false alarms. For example, a highly 'optimistic' dog in the CBT will tend to predict the presence of food in different probe locations and will take less time to approach the bowl. Similarly, this dog may generalise the odour presence across sample locations in a detection trial and indicate false positives more frequently to achieve a reward (5.4.4). Therefore, dogs that are slower to approach the ambiguous stimulus may be more ‘conservative’ (1.2.1.2) and weigh up their decision to alert to a scent for longer before acting.

MDD must discriminate scents at small concentrations, and it is desirable to minimise erroneous indications. Therefore, it is pertinent to investigate further to what extent this tendency can be influenced by preselection, individual attributes, and training.

More scent-specific dogs also tended to discriminate better between ambiguous locations in the CBT, which may be reflected in their detection tasks. These individuals may have greater spatial discrimination, similar to a study where dogs that took longer to approach probe locations in the CBT also learned faster in a spatial memory assessment (Gruen et al., 2019). Here the dogs' longer latencies towards M in CBT may not be entirely related to the tendency to treat ambiguity negatively but to better discrimination skills. Further research should explore the link between scent specificity, the strength of discrimination in CBT, and dogs' performance in other discrimination tasks.

6.5.2.1 Association between 'dogs' behavioural measures and demography and CBM

The only demographic factor associated with CBT was age. The relationship between increased age and shorter latencies to NP is somewhat unexpected. It was anticipated that younger dogs, potentially more curious and energetic, would run faster toward the bowl (Wallis et al., 2020). However, older dogs may have had more training and experience, thus, persevering in their task (Walczak et al., 2012). In contrast, younger dogs may approach faster initially but may be less trained and focused, becoming frustrated sooner when not rewarded and stop coming to the probe locations. In a between-subject CBT study, Piotti et al. (2018) compared the performance of older and younger pet dogs, hypothesising that older dogs would be more pessimistic towards ambiguous stimuli. However, the speed to approach ambiguous probes did not vary significantly with age.

6.5.2.2 Association between dogs' behavioural measures in the test battery and CBT measures

It makes sense that dogs with higher 'Confidence' would have run faster to NP since they would be more 'optimistic' towards ambiguous situations when anticipating a favourable outcome since higher confidence has been associated with positive affective states (1.2.6.1). A study assessing the relationship of performance in CBT with personality traits from the Boldness-Shyness continuum and the C-BARQ questionnaire showed that dogs that were more sociable and excitable approached faster to ambiguous positions. In contrast, those with higher separation-related behaviours and aggressiveness/ fearfulness towards other dogs took longer to approach (Barnard et al., 2018).

Human psychology literature describes optimistic individuals as more emotionally stable, extroverted, more likely to show adaptative actions, and more cognitively flexible, with better memory consolidation and quicker decision-making (Isen, 2001; Rasmussen et al., 2009; Sharpe et al., 2011). In contrast, anxious individuals may take longer to react to a neutral stimulus, make more errors or interpret it as more threatening than less anxious subjects (e.g. Eysenck et al., 1991; Mogg and Bradley, 1998; Bishop, 2007). This 'pessimistic' tendency may also relate to the dogs that displayed more stress behaviours in the test battery taking significantly longer to move towards NN when predicting the absence of food, unlike more 'optimistic' animals that kept visiting the bowl. Similarly, in different CBT studies, dogs with fearful tendencies and anxiety-related disorders tended to have longer latencies towards probe locations (Mendl et al., 2010a; Karagiannis et al., 2015; Barnard et al., 2018).

It was expected that dogs with a higher 'Ability to solve problems' would have shorter latencies towards the ambiguous locations since they may tend to get more involved in different situations to achieve a favourable result without giving up (1.2.6.10). Dogs with a higher 'Food orientation' also ran faster towards the bowl, possibly in anticipation of a food reward. This positive expectation may lead reward-sensitive dogs to persist in seeking food (Burman, 2014).

Dogs with higher 'Playfulness' tended to approach M faster. This is likely, as more playful dogs may generally be more ‘optimistic’, active, curious, and extroverted. Play is considered a positive emotional state in animals (Panksepp, 2004). It has been linked with different positive aspects in humans, such as favourable life quality (Proyer et al., 2010), stress coping (Qian and Yarnal, 2011) and educational accomplishment (Proyer, 2011). In dogs, Playfulness has been associated with better performance for highly active detection tasks (1.2.6.4). However, in this research, lower 'Playfulness' in dogs was found to be related to an increased likelihood of training success (5.3.1.1). Hence, an intermediate level of Playfulness may be favourable for MDD performance by increasing underlying positive affective states without been task disruptive.

Considering these findings together, it would make sense that dogs with higher 'Confidence' (which may well increase with age), 'Playfulness', and low-stress tendencies seem more 'relaxed', possibly related to a relatively high 'optimism'. These may be a desirable set of traits when selecting dogs for MDD tasks.

6.5.2.3 Relationship between the dogs' performance on CBT and their DIAS scores

Impulsivity is an inherent multifactorial aspect (Peremans et al., 2002). Working dogs may be more impulsive and have higher arousal than companion dogs (within an acceptable range of arousal for the task) (Brady et al., 2018), given that working dogs are preselected according to their role requirements (Bray et al., 2021b). It was hypothesised that dogs with higher DIAS scores would approach faster to ambiguous locations, and that was seen to be the case. More impulsive dogs may respond quickly and keep going toward the bowl regardless of whether there is food inside. Similarly, in a study, pet dogs with higher scores of the factor ‘Responsiveness’ in the DIAS approached faster towards an ambiguous location in CBT (Hale, 2021). DIAS scores were also linked with higher 'Persistence' in the test battery (4.4.4). Previous studies in rats and dogs related lower tolerance to immediate gratification with the tendency to initially persist in pressing a device in a delayed gratification task regardless of the results (Van den Bergh et al., 2006; Wright et al., 2012). In dogs, a similar tendency was related to higher DIAS scores (Wright et al., 2012). Therefore, higher levels of impulsivity may have contributed to the dogs' shorter latencies and endurance at CBT.

Conversely, impulsivity has been linked to negative affect (Panksepp, 2004; Wright et al., 2012). Research has found that more impulsive dogs eventually abandon repetitive tasks due to a lack of tolerance for a delayed reward (Wright et al., 2012). Fearful/anxious dogs trained in a CBT that did not complete the test tended to have significantly higher scores of the DIAS factor ‘Aggression/Neofobia' (Hale, 2021) than those that completed the CBT. Therefore, the association between dogs' lower latencies towards ambiguous stimuli and DIAS may seem contradictory in this MDD sample. However, these dogs differ from companion dogs in their genetics and breeding. This research investigated general impulsivity, concentrating on the overall DIAS scores but not the different DIAS factors that explain specific impulsivity aspects (4.1.5). Further research could examine the link between CBT and impulsivity factors in MDD and other working populations.

6.5.2.4 The link between dog performance on CBT and trainers' ratings for 'Willingness to try new behaviours'

I assessed if trainers' ratings on 'Willingness to try new behaviours and get it wrong' were associated with CBT, expecting that it may be related to a tendency to run faster to the probe locations and potentially to false alarms. However, this relationship was not significant, although dogs with shorter latencies tended to have higher ratings on this trait. The link between 'Willingness to try new behaviours' and judgement of ambiguity may not be meaningful. However, this relationship might have a more significant effect in a larger sample and should be considered in future studies.

6.5.2.5 Why are some performance measures linked with different probe locations?

The relationships between performance measures and their latencies to approach the bowl varied across probe locations. Each may provide different information about emotional states (Harding et al., 2004; Doyle et al., 2010; Burman, 2014). However, there are several reasons why dogs approach different locations at variable speeds. These may be mainly related to differences in decision-making under uncertainty, which may come from variations in underlying affect. However, these could also be attributed to individual differences in dogs' breed, behavioural characteristics and background (Burman, 2014; Barnard et al., 2018); learning (Doyle et al., 2010) and cognitive aspects such as spatial memory (Gruen et al., 2019) and laterality (Wells et al., 2017b). It might also be a level of randomness in dogs' variations of approach latencies (Burman, 2014), making it difficult to identify clear reasons for response discrepancies. The CBT used here presents three ambiguous locations to the dogs. However, after several repetitions, the dogs may stop being ambiguous due to learning (Burman, 2014). While M remained 'neutral', NP might be more likely associated with food reward and NN with the absence of it due to generalisation.

Most dogs had short latencies to NP, showed slight variation and did not significantly differentiate between P and NP; this suggests a general optimistic tendency or a high inclination to generalise these locations. Dogs with a higher tendency to approach faster to NP may be more aroused and have a positive affective valence upon reward expectation (see Mendl et al., 2010b). Variation in latency increased as the bowl got closer to N. This is frequently observed across CBT studies. However, dogs with short latencies at NN may have been highly 'optimistic' or deficient in discrimination learning. Dogs with higher stress in the test battery took significantly longer to approach NN. Dogs that approach slower or stop running towards NN may have low arousal and negative valence states when unrewarded (see Mendl et al., 2010b). Potentially assuming that this location will likely bring an unpleasant outcome (lack of reward), as in some CBT studies, the subjects may be reluctant to approach or avoid it. Studies applying punishment (e.g. Harding et al., 2004; Doyle et al., 2010; Bethell et al., 2012) often explain subjects' active adverse stimulus avoidance relating it to danger (See Mogg and Bradley, 1998). However, here N was associated not with punishment but with the absence of a reward, which may cause dogs to move slower, eliciting a state similar to 'disappointment or depression' (Nesse, 2000; Burman, 2014). From an adaptive perspective, animals should seek to save energy if resources are insufficient or when goal-oriented efforts are repeatedly unsuccessful. Hence they are likely to disengage from the task (Nesse, 2000). The dogs' decision tendencies during the CBT can be linked with their core affect. However, it is difficult to infer this is true given that a direct measure of animal affect is currently not possible. Therefore, further research should focus on understanding the potential link between dogs' emotional state and MDD performance. One possible approach could be to use physiological parameters or psychometric scales such as the PANAS or The Canine Frustration Questionnaire (McPeake et al., 2019).

6.5.2.6 *Are medical detection dogs different from other populations?*

Different dog samples have had a high level of dropouts in the CBT test. For instance, in Hale (2021), 41.6% of the dogs did not complete the test, possibly resulting in selective subject inclusion (Burman, 2014). However, no dogs were removed from training, and only two did not complete the test in this study. MDD may have engaged more with CBT than other dog populations due to their training discipline, genetics, and preselection. These dogs were used to participate in training exercises for prolonged periods, and they were familiar with the experimental room and the charity's facilities. In other studies with higher levels of subjects exclusion, the dogs could have experienced neophobia because the CBT was carried out in places unknown to the animals or because they were more anxious animals and had less ability to cope with the task. When Hale (2021) assessed a population of 101 fearful dogs, 42 did not complete the CBT, and of these, 25 showed visible stress-related signals. Alternatively, pet and shelter dogs may show varied energy levels, while working dogs tend to have high stamina (e.g. Cobb et al., 2015; Bray et al., 2021b) and may be more able to complete the CBT without getting fatigued compared to other populations.

MDD learned to discriminate between P and N relatively fast, taking 18.6 trials on average. Other dog populations took longer, e.g. shelter dogs mean= 29.42 trials (Mendl et al., 2010a); pet dogs mean= 42 trials (Muller et al., 2012). It was also interesting that dogs who learned faster tended to achieve better discrimination. This was reflected in their MDD task performance since their training discrepancy was positively correlated with their trainers' ratings on 'Ability to learn from being rewarded'. Perhaps CBT training achievement may be a good predictor of dogs' general learning skills in working dog populations and may be worth exploring further e.g. with a reversal learning task.

MDDs may be more 'optimistic' than other dog populations as they are raised under standardised socialisation and training to make them resilient across contexts and endure in their roles. Although their behaviour may vary across disciplines and training stages, these dogs may be more emotionally stable than companion or rehomed dogs from variable origins and previous experiences. However, this is speculative since CBT studies vary considerably in dog samples and experimental designs. It would also be relevant to assess these results with CBT in different working dog populations to understand how their performance varies across roles. In addition, comparative studies between naïve and working dog samples may provide more information on how CBT performance varies across dog populations and whether testing in naïve dogs predicts later performance after training.

6.5.3 Study limitations and future steps

CBT is purported to allow assessment of the affective status of a wide variety of dogs. In this study, their general decision-making under ambiguity may be a valuable indicator of how they perform MDD tasks, being a relatively easy procedure. However, it also has certain limitations.

Firstly, individual variation is associated with various variables, some of which I aimed to standardise or measure from the study's population, but others may be more challenging to control. For instance, inherent traits and dogs' experience/background, such as social interactions, might be harder to quantify. Although MDD constitutes a fairly pre-selected sample, there is still considerable genetic and environmental variation. The dogs are of different breeds and are reared and housed by volunteer puppy walkers. Hence their housing and living conditions are variable. Likewise, the dogs belong to different disciplines and stages of training, and their training methods may differ subtly. The dogs' age and physical capacity might vary, which I did not investigate here. However, adjusting the dogs' latency helped to homogenise the sample's differences in speed. Future studies could standardise sample conditions to reduce individual effects. However, since there are still few institutions working with MDD, few dogs are trained and kept homogeneously compared to other tasks like guide dogs or explosives or drug detection.

Secondly, The test battery preceded the CBT on the same day. This offered advantages as the dogs were familiar with the experimental arena and the general testing dynamics, allowing me to proceed immediately with CBT. However, the test battery, which occurred earlier, could have been exhausting for the dogs, and various treats were provided during the test. These may have reduced the dogs' energy and appetite for the later CBT. However, the dogs were allowed a rest period of at least two hours between tests. Ideally, I would have performed the CBT on a different day than the test battery. However, this was impossible due to logistical and time limitations.

Thirdly, the dog as model species has advantages regarding its use in research because of its inherent social connection with humans. However, the human factor can affect experimental designs since both the tester and the handler could potentially influence the dog with unconcient cues on the bowl content, i.e. clever Hans effect (Pfungst, 1911; Lit et al., 2011). Individual dogs react differently to new people according to their experience (Burman, 2014). Here, all dogs had been well socialised with people and familiarised with the tester from the test battery. At the CBT onset, the tester moved behind the test screen while the dogs were watching because, when piloting the test setup, the dogs seemed nervous about entering the test area if the tester was already behind the screen. The CBT screen was advantageous as it allowed the tester to manipulate the food bowls with little visual access from the dogs or the handler. However, the dogs could still see the tester's head.

In general, the dogs performed the test without interruption. However, three dogs tried to get through the screen by jumping on it or through the doors. Possibly wanted to approach the tester or explore behind the screen. Future applications of this setup could benefit from a screen that completely blocks the dog's visual access or an automated design allowing the tester to coordinate the test from another room.

Finally, among the advantages of CBT are its low cost and the few materials needed. It is possible to attain data in real-time (i.e. measuring latencies using a stopwatch). However, the test is long and arduous, potentially exhausting for the participants. Hence, I reduced the number of trials to reach the training criterion and the maximum latency to consider a

'No Go' (how is the dog supposed to behave to the negative cue) compared to other studies. Although these measures helped, the CBT is still extensive. In the future, shorter adaptations would increase their feasibility, being cautious of not affecting the experimental design or the data quality.

Even if the CBT may not predict training success, it links with various behavioural aspects indicative of task aptitude. However, dogs' personality is multifaceted (Rooney and Clark, 2021), so it is inadequate to attribute dogs' performance in CBT to isolated behavioural aspects. Further investigation into how these interact together may shed light on how dogs' CBM associates with their medical detection ability. Exploring the CBT of dogs that were rejected from training and how it is linked to the causes of rejection (3.4.1) may allow targeting these relationships in MDD candidates to improve their wellbeing early and promote the selection of emotionally balanced dogs.

These findings could aid in developing socialisation and handling protocols based on those implemented in MDD and tailored to the needs of pet dogs or dog rehoming shelters to improve their wellbeing by promoting a higher 'optimitic' tendency.

6.6 Conclusion

CBT is valuable for understanding the relationship between dogs' decision-making under ambiguity and MDD performance. This exploratory study identified significant links between dogs' general ability and behavioural factors and their tendency to be more ‘optimistic’ or ‘pessimistic’ in a CBT. Future confirmatory studies could focus on individual variations in larger dog samples linked to training outcome and CBT performance.

Likewise, the direct relationship between the dogs' olfactory specificity and their latencies towards probe locations could indicate greater precision in medical detection tasks than the tendency to generalise odours. Future research could further assess this finding to understand how it is reflected in the dogs' olfactory performance.

Finally, MDD 's performance on the CBT varies compared to other dog populations. MDD are more likely to participate and persist on the test and learn faster than other dog populations (i.e. shelter dogs, pet dogs). This suggests that this MDD sample tends to be more ‘optimistic’ than other dog groups. Subsequent studies should address the relationship between dogs' characteristics and CBT responses, including emotional states, which facilitate the selection of MDD.

Chapter 7. General Discussion

As the application of MDD is increasing, it is crucial to expand our understanding of the relationship between dogs' individuality and task performance to enhance the selection of suitable animals for these roles. This dissertation aimed to determine the most relevant traits for MDD and how these differ across bio-detection and assistance disciplines and at different training stages.

This research predicted that different traits in dogs were linked with distinct levels of success in their detection roles and variations in performance over training. It also hypothesised that these traits would vary whether dogs were trained for bio-detection or assistance tasks. In addition, it was anticipated that different traits were associated with varying levels of sensitivity and specificity in the trained dogs' detection roles. These hypotheses were addressed across the dissertation's chapters by integrating multiple methodologies validated in previous research.

Overall, this research is the first, to my knowledge, to explore the impact of individual attributes on MDD. The following sections will summarise the main outcomes from each experimental chapter. It will integrate these findings and assess how these potentially answer this research's hypotheses (1.3). Findings from the test battery and cognitive bias test (CBT) are summarised in Figure 7.1. Finally, it will highlight the project's limitations and future research areas.

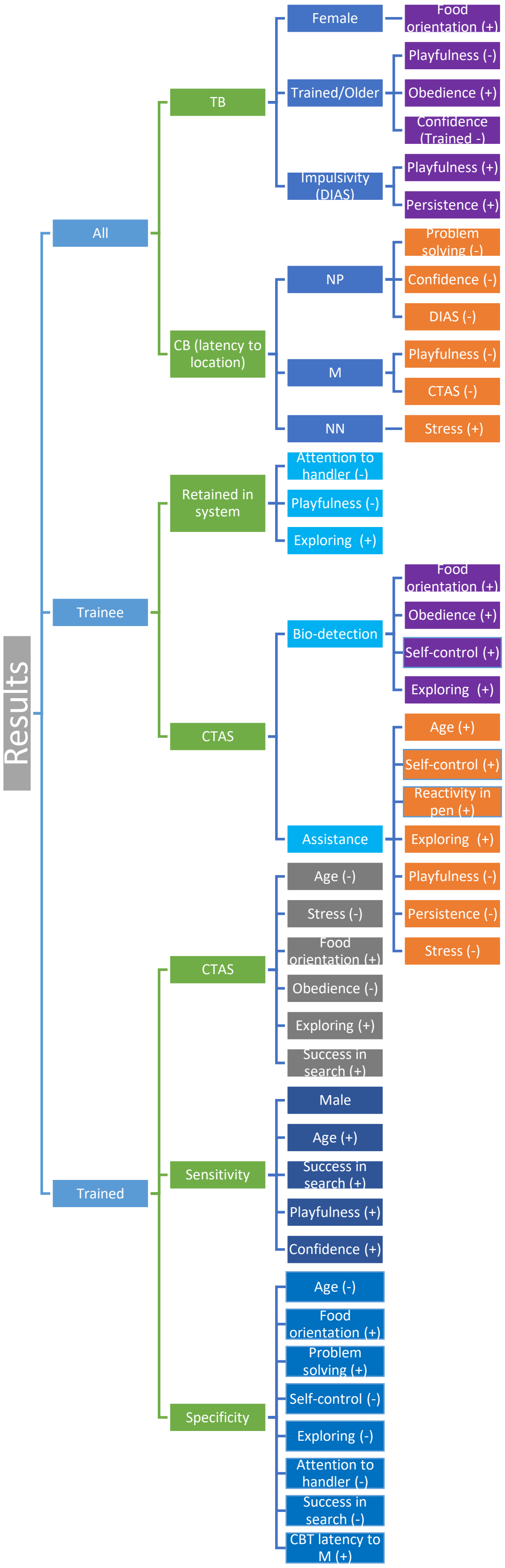


Figure 7. 1 *Main findings from Chapters 4, 5 and 6. (N=58): associations of dogs' test battery components (TB) with demography and DIAS scores; associations between cognitive bias (CB) latencies to ambiguous locations and dogs' demography, TB components and DIAS scores. Trainee dogs: associations between dog demography, TB and CB with training outcome and with CTAS*

(across tasks). Trained dogs: associations between dog demography, TB and CB, CTAS and scent sensitivity and specificity scores. The direction of the associations with each measure is indicated (+)=positive (-)= negative. Exploring=Interest in exploring environment.

7.1 Survey on the importance of different traits for medical detection dogs' performance

In Chapter 2, the MDD survey assessed professionals' opinions on MDD from various countries on the most relevant traits for medical detection tasks, their variation across Bio-detection and Assistance dogs, and what traits in real dogs are far from ideal levels. The study participants rated 40 traits regarding their importance for MDD selection, ideal levels and presence in a dog the participant worked with.

The survey results allow us to visualise the global panorama of MDD, i.e. the demography and the health conditions the dogs are trained to detect, and the perspectives of those who work with the dogs.

The most important traits for selection included 'Level of motivation when working', 'Health', 'Ability to learn from being rewarded', 'Ability to concentrate' and 'Acuity of sense of smell'. Most of these concur with previous studies on explosives and drug detection dogs using a similar methodology (Rooney et al., 2004; Adamkiewicz et al., 2013).

The levels of importance for selection significantly differed across disciplines for ten traits, e.g. 'Tendency to search by smell alone' and 'Ability to solve problems' was considered more important for Bio-detection dogs than Assistance dogs. Whereas 'Attachment to human partner', 'Confidence', and 'Willingness to try new behaviours' were to be significantly more important for Assistance dogs than for Bio-detection dogs. These findings highlight specific requirements for each task and the need for discipline-oriented selection. This is similar to findings from research assessing personality differences across tasks, e.g. explosives and drug detection (Rooney et al., 2004) and explosives detection and service dogs (MacLean and Hare, 2018).

The respondents thought an ideal MDD dog should poses high levels of traits including 'Health', 'Acuity of sense of smell', 'Persistence when alerting', 'Concentration' and 'Confidence'. However, 'Aggression towards people and dogs' and 'Tendency to get distracted' were generally undesirable.

The ideal levels for 16 traits also differed significantly across disciplines: For assistance tasks, traits such as 'Confidence', 'Level of attachment to human partner' and 'Tendency to investigate humans by sniffing' were desired to be significantly higher than for bio-detection tasks. However, for Bio-detection dogs, the ideal levels of traits such as 'Tendency to search by smell alone', 'Ability to concentrate' and 'Independence' exceeded those of the former. Bio-detection dogs also had higher ideal ratings than Assistance dogs for undesirable traits, for instance, 'Tendency to vocalise', 'Fear of specific things' and 'Tendency to be aggressive towards other dogs'. Bio-detection dog handlers may be more tolerant of traits deemed negative since these animals are less exposed to different environments than Assistance dogs.

Nonetheless, the participants rated several traits in the dogs they were working with as higher than ideal, including 'General excitability' and 'Tendency to be distracted when working'. However, traits such as 'Tendency to be aggressive towards people', 'Level of motivation when working' and 'Consistency of behaviour' were close to ideal levels.

The discrepancy from ideal levels of eight traits differed significantly across disciplines. 'Tendency to search by smell alone' was lower than ideal and deviated more in Assistance dogs than in Bio-detection dogs. However, Bio-detection dogs' discrepancy scores were greater than Assistance dogs' scores for several traits like 'General excitability', 'Tendency to vocalise' and 'Tendency to be aggressive towards other dogs', which were all higher than desired. These differences from ideal levels in current dogs may indicate the need to address these traits during dog training and selection.

Overall, this study aided in identifying the most relevant aspects for MDD selection and those that are not as desired, finally deriving 27 traits on which the subsequent studies were based.

7.2 Measurements of medical detection dogs' ability: deriving meaningful measures of MDD performance

Chapter 3 examined what indicates successful performance in medical detection tasks. Overall, it assessed how MDD staff evaluated their dogs (N=58) at different training stages. Ultimately it derived performance measures for the subsequent chapters, including training outcome (remain in the system or fail training), Composite Total Ability Score (CTAS), and for trained dogs, their scent sensitivity and specificity scores. The chapter was divided into two parts:

In Study A, I assessed the reasons for dogs' rejection from the training programme that the trainers mentioned. Training failure was mainly related to a lack of searching ability, anxiety and poor progress. Of 39 trainee dogs, 23 were kept in the programme (59%), and 16 were withdrawn (41%). The rejection proportion is similar to working dog failure rates previously described (Cobb et al., 2015)

The dog trainers evaluated their dogs by rating each of the 27 traits from the survey and the dogs' overall ability (OA) at their tasks. Each rated their dog at the time of the test battery (Chapter 4). I then examined how the traits linked with the dogs' OA and how they varied across bio-detection and assistance disciplines.

Rating for some traits showed high variation over the scale, e.g. 'Frustration during training' and 'Ability to solve problems'. Others, like 'Tendency to bring an object back to a person', scarcely varied.

Ratings for eight traits were significantly higher for Assistance dogs than for Bio-detection dogs and included 'Ability to solve problems' and 'Ability to learn from being rewarded'. Although 'Tendency to vocalise' was higher for Bio-detection dogs than Assistance dogs. Due to task requirements, Assistance dog selection and training may be stricter than for Bio-detection dogs. Also, trained Bio-detection dogs were rated significantly higher than trainees for ten traits and OA. Trained dogs' superior performance may be due to preselection, the effects of training and greater maturity, and trainers may consider them superior as they worked with them for longer.

Dog trainers evaluated their dogs at different time-points in their training for each discipline. The study investigated how consistent the ratings of dogs' behaviour were over training and whether patterns of change were associated with training outcome. Individual progress differed for each dog. Most dogs whose ratings for OA increased or remained the same over time continued in the system. Those with generally poor ratings and declining performance tended to fail training.

Generally, the ratings of Assistance dogs were higher and showed greater consistency over time than those of Bio-detection dogs. Some traits rated consistently for Assistance dogs were 'Confidence in different environments', 'Ability to learn from being rewarded' and 'Obedience to human command', and for Bio-detection, 'Impulsivity' and 'Willingness to bring an object back'.

Certain traits rated were significantly associated with higher OA ability ratings, including higher 'Ability to solve problems', 'Willingness to try new behaviours', and 'Tendency to investigate humans by sniffing', but lower 'Body sensitivity' and 'Adaptation to crate or kennel'.

The dog's discrepancy from ideal was calculated for each trait according to the dog's task. 'Tendency to be distracted when working' was higher than ideal for both disciplines. 'Ability to solve problems' and 'Concentration' were far below expected levels for Bio-detection dogs, and 'Body sensitivity' and 'Tendency to seek human attention' were higher than desired in Assistance dogs. The dogs' OA was integrated with their overall discrepancy from ideal to produce the CTAS, a systematic subjective measure quantifying dog performance (3.2.5.5). Dogs that were retained in the programme had a significantly higher CTAS than those that failed training.

The 27 traits were clustered into five components through PCA: 1.'Responsiveness to training', 2.'Keenness to please', 3.'Human orientation', 4.'Tendency to be impulsive', 5.'Vocal excitability'. These were used in Chapter 4 to investigate the association between the trainers' opinions on dog traits during training and the dogs' performance in the test battery.

As objective measures of trained dog performance, scent sensitivity and specificity scores were collected from 27 trained dogs working on eight detection projects on different target scents. Overall, dogs' sensitivity and specificity were high but varied with the target scent. These were used independently as success measures in subsequent chapters.

In Study B, I examined how assessors rated 13 prospective MDD during an internal aptitude test. They rated 20 traits and their overall ability. Several traits showed high variation, but six did not, e.g. 'Frustration during training' and 'Tendency to explore areas by sniffing'. For most assessors, several traits were highly correlated with others, such as 'Motivation during training' and 'Concentration', indicating that they may have struggled to distinguish them when rating the dogs. Assessors generally agreed with one another when rating 'Motivation during training' and 'Tendency to search by smell alone' but not other traits. Dogs' OA during the test was not associated with training outcome. The in-house test may be improved with greater standardisation, combining subjective and objective measures, and assessors' training on which behaviours to focus on.

7.3 Development of a test to measure individual attributes in MDD dogs: Which variables can be measured? And how do they associate with each other?

Chapter 4 assessed the development of a test battery based on the survey's results (Chapter 2) to quantify the most relevant traits for MDD in the above-mentioned sample (N=58). The test battery involved 18 cognitive and temperament subtests. The trainers also completed the DIAS questionnaire. It assessed how the variables extracted were associated with each other and combined into components, how these varied with the dogs' demography, how these associated with the dogs' DIAS scores and if the dogs' behaviour when tested was associated with the dog trainers' opinion of their general performance.

The test produced 98 variables, clustered into 11 principal components using PCA. These were labelled from their major loadings 1.'Playfulness', 2.'Persistence', 3.'Reactivity in holding pen', 4.'Food orientation', 5.'Obedience', 6.'Level of attention to handler', 7.'Self-control', 8.'Confidence', 9.'Success at problem-solving', 10.'Interest in exploring environment' and 11.'Success in search'. It also quantified the total stress behaviours seen during the test.

I evaluated whether the components were linked with the dogs' demography: 'Food orientation' was higher in females. Trained dogs scored significantly higher for 'Obedience' and lower for 'Playfulness' than trainee dogs. 'Obedience' increased with a dog's age, and 'Playfulness' and 'Confidence' declined. In addition, dogs with higher DIAS scores tended to have higher scores for 'Playfulness' and 'Confidence'.

The dogs' behaviour during the test was compared to their trainers' opinion of their regular performance. Several components were significantly associated with components from the trainers' ratings (3.4.2.8): 'Keenness to please' was correlated with 'Interest in exploring environment' and 'Success in search'. 'Human orientation' was negatively linked with 'Confidence' and 'Vocal excitability' with 'Success at problem-solving' (Figure 7.1).

7.4 Associations between medical detection dogs' behaviour in the test battery and their performance: How do these vary for different training stages and tasks?

Chapter 5 assessed the association of the dog behavioural components from the test battery and the success measures from Chapter 3. In trainee dogs (N=39), I investigated their relationship with the dogs' training outcome and their CTAS over training and if these varied with the dogs' prospective tasks.

Dogs that remained in the system tended to have significantly higher 'Interest in exploring environment', lower 'Playfulness' and 'Level of attention to handler' than those that failed training. Training outcomes did not vary between dog disciplines. However, CTAS was higher in Assistance dogs than in Bio-detection dogs. For each discipline, the dogs' levels of different components were significantly associated with their CTAS. For both disciplines, higher 'Self-control' was linked with higher CTAS. However, other components differed. For Bio-detection dogs, higher 'Food orientation', higher 'Obedience' and lower ‘Interest in exploring environment’ were significantly related to higher CTAS. For Assistance dogs, older age, higher 'Interest in exploring environment' and 'Reactivity in holding pen' but lower 'Playfulness' and 'Persistence', and dogs showing fewer stress behaviours generally had higher CTAS.

In trained dogs (N=19), it examined whether their behaviour in the test battery was associated with their CTAS. Younger age, higher 'Food orientation', ‘Interest in exploring environment’ and 'Success in search', and lower 'Confidence', 'Obedience' and fewer stress behaviours were linked with higher CTAS in their detection projects.

The study also assessed the association between trained dogs' behaviour on the test and their sensitivity and specificity in their projects at the data cut-point (N=27). Dogs' scent sensitivity was significantly linked with age, 'Playfulness', 'Confidence' and 'Success in search'. Dogs' scent specificity was significantly associated with 'Food orientation' and 'Success at problem-solving' and negatively correlated with age, 'Level of attention to handler', 'Self-control', 'Interest in exploring environment' and 'Success in search' (Figure 7.1).

7.5 The cognitive bias of medical detection dogs: Does the outcome of a Cognitive Bias test associate with dog performance in medical detection tasks?

Chapter 6, investigated whether the same dogs tended to anticipate a reward or a negative outcome when making decisions in an ambiguous context with a modified version of the CBT (Mendl et al., 2010a). It specifically assessed the cognitive bias relationship with the dogs' success measures, behaviour in the test battery and their DIAS scores.

Dogs that distinguished better between P and N locations after training tended to achieve learning criteria in the CBT in significantly fewer trials than those that did not, and they also had significantly higher trainers' ratings for 'Ability to learn from being rewarded' in their regular training and work.

Regarding the ambiguous locations, dogs that approached NP faster were significantly older and had significantly higher DIAS scores and higher 'Confidence' and 'Success at problem-solving in the test battery. Shorter latencies to M were significantly associated with higher CTAS and higher 'Playfulness'. Also, dogs that approached faster to NN had fewer stress behaviours in the test battery. Cognitive bias was not significantly linked with training success in trainee dogs or scent sensitivity in trained dogs'. However, scent specificity was significantly associated with latencies to M and their strength of discrimination between ambiguous locations (Figure 7.1).

7.6 Hypotheses addressed

In this PhD project, It was hypothesised that different traits in dogs are linked with different levels of success in their detection roles, that specific traits in dogs are associated with differences in performance over the training programme and training outcomes, and that these will vary between dogs trained for bio-detection or assistance tasks and will be associated with different levels of sensitivity and specificity in their detection roles.

7.6.1 Different dog traits (as tested) are linked with different levels of success in their detection roles

Overall, these research results suggest that several individual differences in MDD are significantly related to subjective and objective measures of task success. Many traits considered relevant to medical detection in the survey (Chapter 2) were tested in current MDD with the test battery (Chapter 4). However, only certain characteristics were repeatedly associated with MDD performance, i.e. 'Playfulness', 'Confidence', 'Success at problem-solving', 'Self-control', 'Obedience', 'Interest in exploring environment', 'Persistence', 'Level of attention to handler', and 'Success in search'. Previous research found strong relationships between these characteristics and successful performance across working roles.

The relevance of confidence and trainability for working dogs has been repeatedly evidenced in different service roles, e.g. guide dogs and detection dogs (e.g. Svartberg, 2002; Harvey et al., 2016). The survey participants deemed these traits among the most important for MDD performance (2.4.5). These also had high loadings within two components from the traits trainers' ratings in Chapter 3 (3.2.2.8): Component 1: 'Responsiveness to training' and Component 2: 'Keenness to please'.

'Interest in exploring environment' was frequently associated with success measures here and highly relevant for detection roles in the scientific literature (e.g. Jamieson et al., 2017). However, 'Tendency to explore areas by sniffing' only had intermediate importance in the survey (2.4.5) and was not particularly relevant in Chapter 3. Suggesting that even if this trait may be necessary to consider for MDD selection, perhaps MDD institutions are not addressing it enough and may need more attention.

Dogs' tendency to play has been deemed particularly relevant for detection tasks (e.g. Svartberg, 2002; Ganitskaya et al., 2020). Here, 'Playfulness' was significantly associated with success measures. The relationship direction varied across samples, although they were mostly negative. 'Motivation to play with toys' was not especially relevant in Chapters 2 and 3, although, in the latter, it was clustered within component 5. 'Vocal excitability' (3.2.2.8). 'Playfulness' was also significantly associated with impulsivity scores in Chapter 4. However, it was also linked with shorter latencies to ambiguous locations in the CBT, indicating that more playful dogs may be more 'optimistic' when making decisions. Therefore, excessive playfulness may be detrimental to performance, but moderate levels could be beneficial. Paying more attention to this trait, which is easy to test and observe, may aid in diminishing rejection rates.

A high 'Level of attention to handler' was negatively linked with MDD performance, similar to previous research in tasks such as guide dogs, military, and search and rescue dogs

(Rooney et al., 2003a; Harvey et al., 2016; Hare et al., 2018). Although good communication with the handler is desirable, an excessive attachment may affect their decision-making ability and lead to behavioural issues. Exercises to increase independence during socialisation and training may aid in promoting a cooperative relationship without getting into an excessive dog reliance on humans.

'Success in search' was particularly associated with trained dogs' performance measures (Figure 7.1). Previous findings in detection roles indicate that high olfactory search ability is fundamental for successful operation (e.g. Rooney et al., 2004; Lazarowski et al., 2019a). In the survey, ‘Acuity of sense of smell’ and ‘Tendency to search by smell alone’ were among the most important. But in Chapter 3 were lower than ideal levels (3.4.2.6), and searching skills deficiencies were the leading cause of failure (3.4.1) 'Success in search' was positively linked with CTAS and scent sensitivity score but negatively with specificity in Chapter 5 (Figure 7.1). Suggesting that this characteristic is crucial for Bio-detection dogs but may vary with their tendency to make decisions in searching trials. Search ability involves physiological and behavioural elements (i.e. olfaction acuity and search motivation etc.). Therefore, it is difficult to isolate the aspects involved with optimal levels. A call for research on differences in sensory perception across conditions’ VOCs and their association with personality variation and decision-making may aid in optimising searching performance.

Research on cognitive traits has been more limited in working dogs. Still, findings support links between problem-solving skills, persistence, social referencing and self-control with operational success in, e.g. explosives detection, guide dogs and disability service dogs (Bray et al., 2017b; MacLean and Hare, 2018; Lazarowski et al., 2019b; Tiira et al., 2020). Here, 'Success at problem-solving' was repeatedly associated with performance measures in the test battery and with shorter latencies in the CBT and 'Ability to solve problems when working' was among the most relevant traits in Chapters 2 and 3. 'Self-control', was linked with CTAS of trainee dogs. ‘Persistence’ was correlated with Impulsivity scores and negatively related to Assistance dogs' CTAS. Increasing attention to cognitive traits during socialisation and training (e.g. implementing problem-solving exercises) and using cognitive tasks would allow selection by focusing on internal flexibility (Figure 7.1).

The latencies to ambiguous locations in CBT varied with different dog traits and performance and the dogs' CTAS. This indicates that MDD's tendency to make 'optimistic' or 'pessimistic' decisions over ambiguity varied in relation to certain behavioural aspects. Variations in CBT performance have been observed across studies of pet and sheltered dogs (Burman, 2014) related, e.g. to behavioural problems (Mendl et al., 2010a; Karagiannis et al., 2015) and personality traits (Barnard et al., 2018). However, this has been mostly unexplored in working dogs. Further studies of CBT in different working populations would provide more information on how decision-making tendencies in ambiguity are associated with differences in operating performance.

The DIAS scores were linked with the test components 'Playfulness' and 'Persistence' and with shorter latencies in CBT. However, high 'Playfulness' was associated with training failure and 'Persistence' with lower ability in Assistance dogs. Hence high levels of certain traits related to impulsivity may be associated with MDD performance.

When tested (Chapters 4, 5 and 6), demography was linked with some success measures. Sex was only significantly associated with sensitivity, being higher in males. Age associations were variable for different performance measures; this may be attributed to dogs' developmental changes and their influence on how they learn as tasks become more complex (e.g. Walczak et al., 2012; Fratkin et al., 2013).

7.6.2 Specific traits in trainee dogs are associated with differences in performance over the training programme and outcome.

There were variations across training stages in specific traits associated with performance. In Chapter 5, some behavioural components were significantly linked with prospective MDD training outcome and their CTAS (Figure 7.1). This suggests that attention to these traits in young dogs during training and selection may increase the likelihood of success in the training programme. The dogs' performance in CBT (Chapter 6) was not significantly linked with their training outcome but may be worth exploring further in larger samples.

7.6.3 Specific traits in dogs (as tested) will vary between dogs trained for bio-detection or assistance tasks.

The results suggested that traits vary across disciplines. The survey indicated that several traits' relevance differed between bio-detection and assistance tasks, and in Chapter 3, trainers rated current dogs differently between disciplines. These may reflect actual differences across samples due to variations in rearing, training and selection for each task but also may be influenced by trainers' role expectations and bias (Clark et al., 2020b; Bray et al., 2021b).

In the test battery, different behavioural components were significantly linked with the CTAS for each discipline. These findings suggest specific behavioural phenotypes needed for each. Here, successful Assistance dogs may be confident in different scenarios, calmer, and bond with humans quickly, similar to what has been characterised in other service dogs, e.g. guide dogs (Bray et al., 2021b). Meanwhile, Bio-detection dogs may be more active, with high problem-solving skills and with high olfactory involvement with their surroundings, similar to findings relevant to other detection tasks (Jamieson et al., 2017; Lazarowski et al., 2020b). Based on this evidence, training and selection for each task may focus on traits more relevant for each. However, testing larger samples of each discipline may provide more information for accurate profiling.

7.6.4 Different dog traits are associated with different levels of sensitivity and specificity in their detection roles

Several traits tested in trained dogs were associated with the dogs' scent sensitivity (higher tendency to alert S+ and fewer false negatives) and specificity (Tendency to ignore S- and avoid false positives). This suggests that some characteristics of trained MDD (Figure 7.1) may be related to their tendency to make decisions on scent alerting over uncertain sensorial information (Heeger and Landy, 1997; Lalkhen and McCluskey, 2008). Previous studies have assessed the association of dog traits with search success in explosives and drug detection dogs (e.g. Rooney et al., 2007b; Ganitskaya et al., 2020; Tiira et al., 2020), and some findings concur with the current research, e.g. the association between 'Playfulness' and sensitivity (5.4.4). However, behavioural assessment methods and measures of search ability varied across studies, making comparisons challenging.

The CBT showed that dogs with higher scent specificity generally took longer to approach an ambiguous location. This suggests that dogs with less 'Optimistic' tendencies may be more cautious when making decisions when searching. Scent sensitivity and specificity are jointly necessary for an optimal detection dog operation (Lazarowski et al., 2020a). However, different detection tasks may require them at different levels (1.2.1.2). Aiming to select dogs best suited for each detection project, achieving scent generalisation quicker, and limiting operational errors. In practice, individual attributes influencing these tendencies may be linked with the dogs' propensity to generalise and identify, e.g. reduced scent concentrations, which is often desirable for ultimate performance (e.g. Walczak et al., 2012; Davies et al., 2019). However, it also may increase false alarms tendency (Lazarowski et al., 2020a). Targeting these traits can be helpful during selection and ongoing training. CBT can be useful for identifying these decision-making trends.

7.7 Limitations and future research

This research provides initial evidence suggesting that MDD individual attributes are associated with different measures of task success. However, these should be considered cautiously due to the research's exploratory nature. Replication is necessary to determine their validity and interpret them more accurately.

These studies were conducted on a relatively small number of dogs. However, the sample size was considerable compared to earlier MDD studies since few dogs are trained for these tasks. However, larger numbers of dogs for different tasks may provide more insights into these relationships.

There were only Assistance dogs in the trainee sample since operational ones were unavailable. Previous research has focussed on Assistance dogs' alerting behaviours and task effectiveness (e.g. Wilson et al., 2019; Wilson et al., 2020). Further studies assessing operational Assistance dogs could explore the link between their attributes (tested here) and alerting behaviour recorded *in-vivo*.

Few Bio-detection dogs were trained for each target scent, and several conditions were investigated. Hence, it is challenging to take the data as a whole. Larger samples for each scent and greater standardisation (e.g. number and sources of scent samples, training trials, etc.) may increase evidence to determine the suitable balance of sensitivity and specificity for each condition and how these may link with dogs' behaviour and their tendency to make decisions when indicating scents.

This project assessed dogs' performance at different training stages when most studies have focussed on trainee dogs' success in training (e.g. Maejima et al., 2007; Harvey et al., 2016). Ideally, I would have followed the dogs' career progress. However, this was not possible during the limited PhD period. Longitudinal studies similar to Riemer et al. (2014c) and Bray et al. (2021a) could monitor performance consistency during a more extended period of the dogs' active life.

The test battery identified behavioural aspects potentially related to performance. Some subtests seem to be more relevant than others. Confirmatory studies with a refined shorter test involving the most effective subtests may aid in adapting this method for practical purposes, i.e. MDD selection or operational monitoring. Due to logistics and time limitations, it was not viable to carry out reliability assessments. However, replication would be appropriate to assess intra and inter-observer reliability and test re-test consistency.

There were no adjustments made for repeated statistical testing as this was an exploratory study to derive initial evidence in MDD behaviour. However, in confirmatory studies, this would be critical to take some of these findings forward and test more specific hypotheses.

This research focused on main associations of each behavioural component with success measures to understand the role of individual factors over performance. However, future work could investigate with linear models the combined effect of independent variables on the outcome variable and their interactions within to integrate these traits for better dog profiling.

Chapter 7

This research derived several measures of dogs' task effectiveness (Chapter 3). However, gathering the trainers' assessments remotely (due to COVID-19) was challenging, and several dog ratings were incomplete. Also, for several dogs, their training was interrupted during the pandemic. It was not possible to assess if these dogs succeeded in their training, just that they were still in the system. In future, more consistent subjective data and certainty on the dogs' fate may enhance the soundness of success parameters since validating evaluation tools with multi-dimensional performance measures is crucial for dogs' practical assessments (Bray et al., 2021b; Rooney and Clark, 2021).

Medical Detection Dogs® charity's staff work thoroughly to train the dogs with frequent success. In Chapter 3, was found low agreement when trainers rated most traits and possibly difficulty distinguishing them (3.8.2). Assessment standardisation and staff education on what to observe during training and testing may help increase evaluation reliability and content validity. Further revised versions of our rating sheets may eventually be employed.

The DIAS has been used to assess impulsivity in companion dogs but has rarely been implemented in working dog populations (Brady et al., 2018). This project showed associations between DIAS scores and some characteristics of MDD. However, only the DIAS overall score was examined. Previous studies on companion and working dogs have found relevant findings when assessing each of the three DIAS factors (Wright et al., 2012; Brady et al., 2018; Hale, 2021). Future research could explore in-depth impulsivity and its association with MDD performance by assessing each DIAS factor independently.

MDD is a novel working-dog area, granting opportunities for new research. This project selected certain aspects deemed relevant for MDD, also linked with other working tasks' performance (e.g. Rooney et al., 2004; Lazarowski et al., 2020c; Bray et al., 2021b). In addition, it evaluated the impact of cognitive functions, which reflect behavioural flexibility. Some are evident from dogs' early life and are consistent over development (Bray et al., 2021a). Several cognitive traits have shown significant links with working dog performance (e.g. Bray et al., 2017b; MacLean and Hare, 2018; Tiira et al., 2020), some similar to the current research. However, others not explored here may be important too. For instance, working memory and laterality were previously associated with task success for detection and service dogs (e.g. Batt et al., 2009; MacLean and Hare, 2018). It is worth further examining the association of cognitive dimensions in future MDD studies.

7.8 Conclusion

These research findings suggest that individual differences in MDD are linked with success in their tasks across disciplines and training stages. This knowledge may help improve MDD training methods and selection and optimise their performance during their working life to aid human health. Further confirmatory research with larger samples and adjusted experimental designs is essential to confirm these results.

APPENDIX

Chapter 2

Appendix 1. Survey preliminary interview

Participant: Date: Time:

1. How long have you been working with detection dogs?
2. Have you worked with dogs performing a different working task before?
3. Why did you come to work with medical detection dogs?
4. Have you currently or in the past worked with both medical bio-detection and medical alert Assistance dogs?
5. Which characteristics are important in a detection dog in general (at list five, but you can add as many as you consider relevant)?
6. Describe your ideal medical detection dog
7. Describe the best medical detection dog you have ever worked with and which attributes you consider to make them outstanding.
8. What traits do you consider less desirable in a medical detection dog (at list five, but you can add as many as you consider relevant)?
9. Describe the differences in behaviour traits you found when comparing medical bio-detection and medical alert Assistance dogs.
10. How do you know when a dog is a good candidate to be trained for medical detection tasks?
11. Do you have a prefered sex for medical detection tasks?
12. Do you have a a prefered breed for medical detection tasks?

Appendix 2. Facebook post advertising medical detection dogs survey

(Illustration by the author SBD)

Have you ever worked with detection dogs in medical roles?

If so, we will really appreciate your help with an exciting new University of Bristol study to help understand the characteristics that make the best dogs for these tasks.
If you currently or in the past have trained, handled or selected bio-detection (searching remote samples) or medical assistance dogs (e.g. diabetes and epilepsy alerts) please, could you spare 20 minutes to fill in our survey?https://goo.gl/forms/tShW8dUsDPSpMGy22

Appendix 3. Medical detection dogs survey

Survey on the importance of different traits for medical detection dog performance

* Required

Participant information sheet

Many thanks for taking the time to complete this survey.

It is part of a PhD research project at the University of Bristol Veterinary School. This study aims to understand what makes a good medical detection dog and what factors influence their working performance. We would very much value your opinions, if you can take the time to tell us.

Your participation in this research is completely voluntary, and your answers will remain anonymous. Only researchers closely involved with this study will have access to the information provided. It will take no more than 20 minutes of your time. This study has the approval of University of Bristol Faculty of Health Sciences Research Ethics Committee.

If you are over 18, a member of an organisation working with medical detection dogs (bio-detection or assistance), and have trained or worked closely (currently or in the past) with at least one dog, please fill in the survey. The length of time you have worked with these dogs does not matter, your contribution of knowledge and expertise will be greatly appreciated.

The survey is divided into three sections. Firstly, you will be asked to rate a range of individual traits according to what level of each trait you think the ideal medical detection dog would have. Secondly, you will rate the same traits according to how important you think these are when selecting a dog for medical detection tasks. Thirdly, you will need to rate a dog you have worked with. Finally, you will be asked to answer questions about yourself and your working experience (for demographic purposes only). Specific instructions will be provided in each section. please read them carefully, although the traits to be rated are repeated in every section, the purpose of each section is different.

1. **If you are happy to participate in this survey, please indicate your consent for each of the statements below:** *

   Check all that apply.

   - [ ] I certify that I am over 18
   - [ ] I have read the information above
   - [ ] I agree for my information to be processed by researchers involved in this study as mentioned above
   - [ ] I agree to participate in this study

Many thanks for your help

Ideal Medical Detection Dog

In this section, a range of 40 different individual traits, potentially important for medical detection tasks, will be presented to you. Please indicate the level of each trait you think the IDEAL dog used for the discipline you work in would have. Please use the scale of 1 to 5: As high as possible (5), High (4), Intermediate (3), Low (2), As low as possible (1).

For example, if you consider that your ideal dog would never be distracted when working you should choose 1 (As low as possible) for "Tendency to be distracted when working".

2. **A) Ideal levels of each trait** *

Mark only one oval per row.

	1. As low as possible	2. Low	3. Intermediate	4. High	5. As high as possible
Acuity of sense of smell - sharpness of nose					
Tendency to search by smell alone					
Tendency to investigate humans by sniffing					
Tendency to explore areas by sniffing					
Level of motivation when working					
Willingness to try new behaviours even if they are wrong					
Persistence when alerting the presence of a target odour					
Tendency to remain specific to the target odour - only respond to precise odour trained					
Tendency to generalise alerts to similar scents					
Consistency of behaviour from day to day					

3. **B) Ideal levels of each trait** *

Mark only one oval per row.

	1. As low as possible	2. Low	3. Intermediate	4. High	5. As high as possible
Tendency to become frustrated when working					
Health - likelihood of having a long healthy working life					
Stamina - endurance when working					
General excitability - tendency to become highly aroused					
Ability to remain calm when not working					
Confidence in different environments					
Travel ability					
Ease of adaptation to crate or kennel					
Friendliness to new people					
Friendliness to other dogs					

4. **C) Ideal levels of each trait** *

Mark only one oval per row.

	1. As low as possible	2. Low	3. Intermediate	4. High	5. As high as possible
Level of attachment to human partner	◯	◯	◯	◯	◯
Obedience to human commands	◯	◯	◯	◯	◯
Ability to learn from being rewarded	◯	◯	◯	◯	◯
Motivation to obtain food	◯	◯	◯	◯	◯
Motivation to play with toys	◯	◯	◯	◯	◯
Willingness to bring an object back to a person	◯	◯	◯	◯	◯
Motivation to retain possession of an object	◯	◯	◯	◯	◯
Independence - ability to work without constant guidance	◯	◯	◯	◯	◯
Ability to solve problems when working	◯	◯	◯	◯	◯
Ability to concentrate during a trained task	◯	◯	◯	◯	◯

5. **D) Ideal levels of each trait** *

Mark only one oval per row.

	1. As low as possible	2. Low	3. Intermediate	4. High	5. As high as possible
Tendency to vocalise in public places	◯	◯	◯	◯	◯
Tendency to seek human attention	◯	◯	◯	◯	◯
Tendency to chase an object	◯	◯	◯	◯	◯
Tendency to be distracted when working	◯	◯	◯	◯	◯
Fear of specific things (e.g. litter bags, brooms)	◯	◯	◯	◯	◯
Reaction to sudden loud noises	◯	◯	◯	◯	◯
Body sensitivity - reactivity to touch and contact with objects	◯	◯	◯	◯	◯
Impulsivity - tendency to make hasty choices when working	◯	◯	◯	◯	◯
Tendency to be aggressive towards people	◯	◯	◯	◯	◯
Tendency to be aggressive towards other dogs	◯	◯	◯	◯	◯

Importance of different traits

Please rate the same traits according to how IMPORTANT you think it is to consider each when choosing a medical detection dog.

Please rate the importance of each trait as: Vitally important (5); Important (4) Intermediate (3), Slightly important (2) or Not at all important (1).

Try to use the whole scale and answer independently of the ideal level you described in Section 1.

❢PLEASE NOTE ❢ that the IMPORTANCE OF TRAITS (in this section) it is NOT THE SAME as the IDEAL LEVEL of traits (in Section 2). Some traits may be ideally needed at low levels, but can still be very important to consider when selecting dogs. So, for instance, if you think that "Fear of specific things" should ideally be very low in a medical detection dog (rated as 1 in the previous section), but you think that this is highly important trait to consider, and select against, when choosing a dog, you should select 5 (Vitally important) in this section. So even undesirable traits needed at very low levels can still be vitally important to consider, and so should be rated as 5 here.

6. **A) Importance of different traits** *

Mark only one oval per row.

	1. Not at all important	2. Slightly important	3. Intermediate	4. Important	5. Vitally important
Acuity of sense of smell - sharpness of nose	◯	◯	◯	◯	◯
Tendency to search by smell alone	◯	◯	◯	◯	◯
Tendency to investigate humans by sniffing	◯	◯	◯	◯	◯
Tendency to explore areas by sniffing	◯	◯	◯	◯	◯
Level of motivation when working	◯	◯	◯	◯	◯
Willingness to try new behaviours even if they are wrong	◯	◯	◯	◯	◯
Persistence when alerting the presence of a target odour	◯	◯	◯	◯	◯
Tendency to remain specific to the target odour - only respond to precise odour trained	◯	◯	◯	◯	◯
Tendency to generalise alerts to similar scents	◯	◯	◯	◯	◯
Consistency of behaviour from day to day	◯	◯	◯	◯	◯

7. **B) Traits importance to consider for selection** *

Mark only one oval per row.

	1. Not at all important	2. Slightly important	3. Intermediate	4. Important	5. Vitally important
Tendency to become frustrated when working	◯	◯	◯	◯	◯
Health - likelihood of having a long healthy working life	◯	◯	◯	◯	◯
Stamina - endurance when working	◯	◯	◯	◯	◯
General excitability - tendency to become highly aroused	◯	◯	◯	◯	◯
Ability to remain calm when not working	◯	◯	◯	◯	◯
Confidence in different environments	◯	◯	◯	◯	◯
Travel ability	◯	◯	◯	◯	◯
Ease of adaptation to crate or kennel	◯	◯	◯	◯	◯
Friendliness to new people	◯	◯	◯	◯	◯
Friendliness to other dogs	◯	◯	◯	◯	◯

8. **C) Traits importance to consider for selection** *

Mark only one oval per row.

	1. Not at all important	2. Slightly important	3. Intermediate	4. Important	5. Vitally important
Level of attachment to human partner	◯	◯	◯	◯	◯
Obedience to human commands	◯	◯	◯	◯	◯
Ability to learn from being rewarded	◯	◯	◯	◯	◯
Motivation to obtain food	◯	◯	◯	◯	◯
Motivation to play with toys	◯	◯	◯	◯	◯
Willingness to bring an object back to a person	◯	◯	◯	◯	◯
Motivation to retain possession of an object	◯	◯	◯	◯	◯
Independence - ability to work without constant guidance	◯	◯	◯	◯	◯
Ability to solve problems when working	◯	◯	◯	◯	◯
Ability to concentrate during a trained task	◯	◯	◯	◯	◯

9. **D) Traits importance to consider for selection** *

Mark only one oval per row.

	1. Not at all important	2. Slightly important	3. Intermediate	4. Important	5. Vitally important
Tendency to vocalise in public places	◯	◯	◯	◯	◯
Tendency to seek human attention	◯	◯	◯	◯	◯
Tendency to chase an object	◯	◯	◯	◯	◯
Tendency to be distracted when working	◯	◯	◯	◯	◯
Fear of specific things (e.g. litter bags, brooms)	◯	◯	◯	◯	◯
Reaction to sudden loud noises	◯	◯	◯	◯	◯
Body sensitivity - reactivity to touch and contact with objects	◯	◯	◯	◯	◯
Impulsivity - tendency to make hasty choices when working	◯	◯	◯	◯	◯
Tendency to be aggressive towards people	◯	◯	◯	◯	◯
Tendency to be aggressive towards other dogs	◯	◯	◯	◯	◯

Information about a medical detection dog you are working with

In this section, you will be asked to score a medical detection dog you are currently working with or have worked with, using the same traits as before. If you are not currently working with any dog, think of the most recent you worked with. Ideally, this dog should have completed training or in the last stages of training. If you are working with more than one dog, select the one whose initials are first in the alphabet.

10. **What task is the dog trained to perform?** *

Mark only one oval.

◯ Bio-detection, detecting remote odour samples for disease diagnosis (e.g. cancer detection within a scent platform)

◯ Medical alert assistance for clients with certain conditions (e.g. diabetes, Addison’s disease, epilepsy)

◯ More than one task

11. **If the dog is trained to perform a medical alert assistance task(s), which condition(s) is he/she trained to detect (Select all that apply)?**

Check all that apply.

- [ ] Hypoglycaemia
- [ ] Hyperglycaemia
- [ ] Epileptic seizures
- [ ] Addison's disease
- [ ] Allergies
- [ ] Narcolepsy
- [ ] Other

12. **If you selected "Other", please specify.**

____

13. **If the dog is trained to detect remote odour samples for disease diagnosis, which condition(s) is he/she trained to detect ?**

____

14. **What is the dog's name?**

____

15. **Please select your dog's age** *

Mark only one oval.

- Not sure
- Under 6 months
- 6 months
- 7 months
- 8 months
- 9 months
- 10 months
- 11 months
- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- 6 years
- 7 years
- 8 years
- 9 years
- 10 years
- 11 years
- 12 years
- 13 years
- 14 years
- 15 years
- Other

16. **If you selected "Other", please specify.**

17. **What is the dog's sex?** *

Mark only one oval.

- Male
- Female

18. **Is the dog neutered/spayed?** *

Mark only one oval.

- Yes
- No
- Not sure

19. **What breed or type is your dog?** *
Mark only one oval.

- Pedigree/pure breed
- First - cross
- Mixed breed
- Not sure

20. **If your dog is a pure breed, what breed is he/she?** *
Mark only one oval.

- Not applicable
- Labrador Retriever
- Golden Retriever
- German Shepherd Dog
- Belgian Mallinois
- Border Collie
- Cocker Spaniel
- English Springer Spaniel
- Poodle
- Other

21. **If you selected "Other", please specify.**

22. **Please select the dog's size** *
Check all that apply.

- [ ] Extra small (e.g. Chihuahua, Yorkshire Terrier)
- [ ] Small (e.g. Jack Russell Terrier, Miniature Schnauzer)
- [ ] Medium (e.g. Border Collie, English Springer Spaniel)
- [ ] Large (e.g. German Shepherd Dog, Labrador Retriever)
- [ ] Extra-large (e.g. Great Dane, St Bernard)

23. **From where was this dog bred by your organisation?** *
Mark only one oval.

- Breeder
- Dog-rehoming shelter
- Personal donation
- Reassigned after not being suitable for a previous working task
- Not sure
- Other

24. **If you selected "Other", please specify.**

25. **Approximatelly what age was the dog when he/she was obtained?** *

Mark only one oval.

- Unknown
- 1 month
- 2 months
- 3 months
- 4 months
- 5 months
- 6 months
- 7 months
- 8 months
- 9 months
- 10 months
- 11 months
- 12 months (1 year)
- 13 months
- 14 months
- 15 months
- 16 months
- 17 months
- 18 months
- 19 months
- 20 months
- 21 months
- 22 months
- 23 months
- 24 months (2 years)
- 25 months
- 26 months
- 27 months
- 28 months
- 29 months
- 30 months
- 31 months
- 32 months
- 33 months
- 34 months
- 35 months
- 36 months (3 years)
- 37 months
- 38 months

- 39 months
- 40 months
- 41 months
- 42 months
- 43 months
- 44 months
- 45 months
- 46 months
- 47 months
- 48 months (4 years)
- 49 months
- 50 months
- 51 months
- 52 months
- 53 months
- 54 months
- 55 months
- 56 months
- 57 months
- 58 months
- 59 months
- 60 months (5 years)
- More than 5 years

26. **If you selected "More than 5 years", please specify how many.**

27. **How long have you been working with this dog?** *

Mark only one oval.

- 1 month
- 2 months
- 3 months
- 4 months
- 5 months
- 6 months
- 7 months
- 8 months
- 9 months
- 10 months
- 11 months
- 12 months (1 year)
- 13 months
- 14 months
- 15 months
- 16 months
- 17 months
- 18 months
- 19 months
- 20 months
- 21 months
- 22 months
- 23 months
- 24 months (2 years)
- 25 months
- 26 months
- 27 months
- 28 months
- 29 months
- 30 months
- 31 months
- 32 months
- 33 months
- 34 months
- 35 months
- 36 months (3 years)
- 37 months
- 38 months
- 39 months

- 40 months
- 41 months
- 42 months
- 43 months
- 44 months
- 45 months
- 46 months
- 47 months
- 48 months (4 years)
- 49 months
- 50 months
- 51 months
- 52 months
- 53 months
- 54 months
- 55 months
- 56 months
- 57 months
- 58 months
- 59 months
- 60 months (5 years)
- More than 5 years

28. **If you selected "More than 5 years", please specify how many.**

29. **Does this dog have any health problem(s)?** *

Mark only one oval.

- Yes
- No

30. **If you selected "Yes", please specify.**

31. **How does your dog indicate when he/she alerts the presence of a scent (select all that apply)?** *

Check all that apply.

- [ ] Scratch
- [ ] Bite
- [ ] Sit
- [ ] Stand and stare
- [ ] Jump
- [ ] Vocalise
- [ ] Retrieve an article
- [ ] Other

32. **If you selected "Other", please specify.**

____

33. **What method do you use to reward your dog when training/ working scent (select all that apply)?** *

Check all that apply.

- [ ] Food
- [ ] Toy
- [ ] Stroking or petting (physical contact)
- [ ] Verbal praise (E.g. "Good boy/girl")
- [ ] Combination
- [ ] Other

34. **If you selected "Other", please specify.**

____

35. **If your dog is trained for medical alert assistance, has he/she already been paired with a human client/partner/patient?**

Mark only one oval.

- ( ) Yes
- ( ) Not yet

36. **How satisfied are you with this dog's overall performance in medical detection tasks?** *

Mark only one oval.

- ( ) Very satisfied
- ( ) Satisfied
- ( ) Neutral
- ( ) Not very satisfied
- ( ) Not at all satisfied

Rate your current dog

Please rate your current dog (the one your described above) for each trait. Indicate the level of each individual attribute using a scale of 1 to 5, in which the options are as extremely high (5), high (4), intermediate (3), low (2), extremely low (1). Please try to consider the level of each trait individually.

37. **A) Your current dog** *

Mark only one oval per row.

	1. Extremely low	2. Low	3. Intermediate	4. High	5. Extremely high
Acuity of sense of smell - sharpness of nose	⬭	⬭	⬭	⬭	⬭
Tendency to search by smell alone	⬭	⬭	⬭	⬭	⬭
Tendency to investigate humans by sniffing	⬭	⬭	⬭	⬭	⬭
Tendency to explore areas by sniffing	⬭	⬭	⬭	⬭	⬭
Level of motivation when working	⬭	⬭	⬭	⬭	⬭
Willingness to try new behaviours even if they are wrong	⬭	⬭	⬭	⬭	⬭
Persistence when alerting the presence of a target odour	⬭	⬭	⬭	⬭	⬭
Tendency to remain specific to the target odour - only respond to precise odour trained)	⬭	⬭	⬭	⬭	⬭
Tendency to generalise alerts to similar scents	⬭	⬭	⬭	⬭	⬭
Consistency of behaviour from day to day	⬭	⬭	⬭	⬭	⬭

38. **B) Your current dog** *

Mark only one oval per row.

	1. Extremely low	2. Low	3. Intermediate	4. High	5. Extremely high
Tendency to become frustrated when working	⬭	⬭	⬭	⬭	⬭
Health - likelihood of having a long healthy working life	⬭	⬭	⬭	⬭	⬭
Stamina - endurance when working	⬭	⬭	⬭	⬭	⬭
General excitability - tendency to become highly aroused	⬭	⬭	⬭	⬭	⬭
Ability to remain calm when not working	⬭	⬭	⬭	⬭	⬭
Confidence in different environments	⬭	⬭	⬭	⬭	⬭
Travel ability	⬭	⬭	⬭	⬭	⬭
Ease of adaptation to crate or kennel	⬭	⬭	⬭	⬭	⬭
Friendliness to new people	⬭	⬭	⬭	⬭	⬭
Friendliness to other dogs	⬭	⬭	⬭	⬭	⬭

39. **C) Your current dog** *

Mark only one oval per row.

	1. Extremely low	2. Low	3. Intermediate	4. High	5. Extremely high
Level of attachment to human partner	◯	◯	◯	◯	◯
Obedience to human commands	◯	◯	◯	◯	◯
Ability to learn from being rewarded	◯	◯	◯	◯	◯
Motivation to obtain food	◯	◯	◯	◯	◯
Motivation to play with toys	◯	◯	◯	◯	◯
Willingness to bring an object back to a person	◯	◯	◯	◯	◯
Motivation to retain possession of an object	◯	◯	◯	◯	◯
Independence - ability to work without constant guidance	◯	◯	◯	◯	◯
Ability to solve problems when working	◯	◯	◯	◯	◯
Ability to concentrate during a trained task	◯	◯	◯	◯	◯

40. **D)Your current dog** *

Mark only one oval per row.

	1. Extremely low	2. Low	3. Intermediate	4. High	5. Extremely high
Tendency to vocalise in public places	◯	◯	◯	◯	◯
Tendency to seek human attention	◯	◯	◯	◯	◯
Tendency to chase an object	◯	◯	◯	◯	◯
Tendency to be distracted when working	◯	◯	◯	◯	◯
Fear of specific things (e.g. litter bags, brooms)	◯	◯	◯	◯	◯
Reaction to sudden loud noises	◯	◯	◯	◯	◯
Body sensitivity - reactivity to touch and contact with objects	◯	◯	◯	◯	◯
Impulsivity - tendency to make hasty choices when working	◯	◯	◯	◯	◯
Tendency to be aggressive towards people	◯	◯	◯	◯	◯
Tendency to be aggressive towards other dogs	◯	◯	◯	◯	◯

About yourself

In this section you will be asked some questions about yourself, your professional experience and your current employment.

41. **Please select your age range**
Mark only one oval.

- 18 to 20
- 21 to 29
- 30 to 39
- 40 to 49
- 50 to 59
- 60 or over
- Prefer not to say

42. **Which gender do you most identify as?**
Mark only one oval.

- Female
- Male
- Prefer not to say
- Other: ____________________

43. **In Which country do you currently live?**

____

44. **In which institution/organisation are you currently employed?** *

____

45. **What is your position in this institution/organisation?** *

____

46. **Please select your main roles in your current job (you can select more than one option)**

Check all that apply.

- [ ] Manager/director
- [ ] Client support
- [ ] Client-dog instructor
- [ ] Dog trainer
- [ ] Dog groomer
- [ ] Handler
- [ ] Dog walker
- [ ] Dog behaviour specialist
- [ ] Animal care assistant
- [ ] Veterinarian
- [ ] Veterinary nurse
- [ ] Dog selection/ procurement
- [ ] Lecturer
- [ ] Funds raiser
- [ ] Public representative
- [ ] Other

47. **If you selected "Other", please specify.**

48. **Approximately how many dogs do your institution/organization currently works with?**

Mark only one oval.

- 1 dog
- 2 dogs
- 3 dogs
- 4 dogs
- 5 dogs
- 6 dogs
- 7 dogs
- 8 dogs
- 9 dogs
- 10 dogs
- 11 dogs
- 12 dogs
- 13 dogs
- 14 dogs
- 15 dogs
- 16 dogs
- 17 dogs
- 18 dogs
- 19 dogs
- 20 dogs
- 21 dogs
- 22 dogs
- 23 dogs
- 24 dogs
- 25 dogs
- 26 dogs
- 27 dogs
- 28 dogs
- 29 dogs
- 30 dogs
- 31 dogs
- 32 dogs
- 33 dogs
- 34 dogs
- 35 dogs
- 36 dogs
- 37 dogs
- 38 dogs
- 39 dogs

- 40 dogs
- 41 dogs
- 42 dogs
- 43 dogs
- 44 dogs
- 45 dogs
- 46 dogs
- 47 dogs
- 48 dogs
- 49 dogs
- 50 dogs
- More than 50 dogs

49. **If you selected "More than 50 dogs", please specify how many.**

50. **How long have you been working with medical detection dogs?**

Mark only one oval.

- 1 month
- 2 months
- 3 months
- 4 months
- 5 months
- 6 months
- 7 months
- 8 months
- 9 months
- 10 months
- 11 months
- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- 6 years
- 7 years
- 8 years
- 9 years
- 10 years
- 11 years
- 12 years
- 13 years
- 14 years
- 15 years
- 16 years
- 17 years
- 18 years
- 19 years
- 20 years
- More than 20 years

51. **If you selected "More than 20 years", please specify how many.**

52. **How long have you been involved with working dogs in general (including other tasks different than medical detection)?**

Mark only one oval.

- 1 month
- 2 months
- 3 months
- 4 months
- 5 months
- 6 months
- 7 months
- 8 months
- 9 months
- 10 months
- 11 months
- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- 6 years
- 7 years
- 8 years
- 9 years
- 10 years
- 11 years
- 12 years
- 13 years
- 14 years
- 15 years
- 16 years
- 17 years
- 18 years
- 19 years
- 20 years
- More than 20 years

53. **If you selected "More than 20 years", please specify how many.**

54. **What is the main function of the medical detection dog(s) you work with (please select all that apply)?**

Check all that apply.

- [ ] Medical alert assistance for clients with specific conditions (e.g. diabetes, Addison's disease, epilepsy)
- [ ] Bio-detection, detecting remote samples (e.g. cancer detection within a scent platform)

55. **If you work with medical alert assistance dogs, which condition(s) are they trained to detect (Select all that apply)?**

Check all that apply.

- [ ] Hypoglycaemia
- [ ] Hyperglycaemia
- [ ] Epileptic seizures
- [ ] Addison's disease
- [ ] Allergies
- [ ] Narcolepsy
- [ ] Other

56. **If you selected "Other", please specify.**

____

57. **If you work with bio-detection dogs, which condition(s) are they trained to detect?**

____

____

____

____

____

58. **How many medical detection dogs are you currently working with (if you are not working with any dogs at the moment select "none")?**

Mark only one oval.

- ( ) None
- ( ) 1
- ( ) 2
- ( ) 3
- ( ) 4
- ( ) 5
- ( ) 6
- ( ) 7
- ( ) 8
- ( ) 9
- ( ) 10
- ( ) More than 10

59. **If you selected "More than 10", please specify how many.**

60. **How many medical detection dogs have you worked with in the past?**
Mark only one oval.

- None
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- More than 10

61. **If you selected "More than 10" please specify how many.**

62. **Have you been involved in the selection and/or procurement of medical detection dogs?**
Mark only one oval.

- Yes, currenty
- Yes, in the past
- No

63. **Have you worked previously with dogs intended for tasks other than medical detection?**
Mark only one oval.

- Yes
- No

64. **If yes, please select the dog area(s) you have had experience with (please select all that apply)**

Check all that apply.

- [ ] Companion
- [ ] Explosives detection
- [ ] Narcotic/drug detection
- [ ] Dual tasks (e.g. Patrol and explosive detection)
- [ ] Contraband of food or forbidden items detection
- [ ] Conservation roles (e.g. scat detection)
- [ ] Insect detection
- [ ] Oestrus in cattle detection
- [ ] Military tasks
- [ ] Patrolling or guarding tasks
- [ ] Guiding for the visually impaired
- [ ] Mobility assistance
- [ ] Emotional assistance
- [ ] Hearing assistance
- [ ] Animal assisted therapy
- [ ] Herding
- [ ] Agility
- [ ] Racing
- [ ] Sledding
- [ ] Other

65. **If you selected "Other", please specify.**

Please select up to three dog breeds you have the most experience working with for medical detection tasks.

66. *

Check all that apply.

- [ ] Mixed breed
- [ ] Labrador Retriever
- [ ] Golden Retriever
- [ ] German Shepherd Dog
- [ ] Belgian Mallinois
- [ ] Border Collie
- [ ] Cocker Spaniel
- [ ] English Springer Spaniel
- [ ] Poodle
- [ ] Other:

67. **If you selected "Other", please specify.**

68. **Please select which of these dog breeds you consider as the best suited to medical detection work (If you do not have a specific preference, select "None").** *
Mark only one oval.

- None
- Mixed breed
- Labrador Retriever
- Golden Retriever
- German Shepherd Dog
- Belgian Mallinois
- Border Collie
- Cocker Spaniel
- English Springer Spaniel
- Poodle
- Other

69. **If you selected "Other", please specify.**

70. **What sex do you think makes the best medical detection dog?** *
Mark only one oval.

- Any male
- Any female
- Neutered male
- Non-neutered male
- Spayed female
- Non-neutered female
- No difference

71. **In your opinion, which are the three main traits that should be considered when selecting a medical detection dog (in your own words and order of importance)?**

72.

73.

74. **We have tried to include all important traits in the previous sections, but if are there any vital traits that you feel we have missed, please list them here.**

Many thanks for your time!

Just to remind you that your answers will remain anonymous. Your participation has been highly valuable to gain greater understanding about which aspects of medical detection dogs' behaviour make them more successful. In addition, we can explore how individual differences vary between different medical detection roles. Potentially, this information could help improve selection, getting more successful medical detection dogs and enhancing their well-being.We want to thank you again to taking the time to complete this survey

75. **Would you like to hear the results of this survey?**
Check all that apply.

- [ ] I would like to hear about the results of this survey

If you would be happy to be contacted for further research projects over the next 10 years, please provide your name and email address in the spaces below.

76. **Name:**

77. **E-mail:**

Powered by
Google Forms

Appendix

Chapter 3

Appendix 4. MDD sample

Table. MDD sample demographic details, training status when tested and at data collection end point and training outcome

Dog ID	Sex	Sexual status	Dog breed	Age in months	Task	Trainin stage	Status when tested	Current fate	Training outcome
1	Male	Neutered	LR	19	Bio	Trainee	Socialiser	Traning	Success
2	Female	Neutered	LG	17	Assis	Trainee	Socialiser	Operational	Success
3	Female	Neutered	LR	29	Assis	Trainee	Training	Traning	Success
4	Male	Neutered	LR	29	Bio	Trainee	Socialiser	Operational	Success
5	Male	Neutered	LR	24	Bio	Trainee	Socialiser	Failed	Failed
6	Male	Neutered	LR	18	Bio	Trainee	Training	Failed	Failed
7	Female	Neutered	LG	25	Assis	Trainee	Training	Operational	Success
8	Male	Neutered	LR	20	Bio	Trainee	Training	Operational	Success
9	Female	Neutered	LR	14	Bio	Trainee	Socialiser	Operational	Success
10	Male	Neutered	BC	113	Bio	Trained	Retired	Retired	Success
11	Female	Neutered	LR	29	Assis	Trainee	Training	Operational	Success
12	Male	Neutered	LR	13	Bio	Trainee	Socialiser	Training	Success
13	Female	Neutered	LR	21	Bio	Trainee	Socialiser	Traning	Success
14	Female	Neutered	CS	15	Bio	Trainee	Training	Failed	Failed
15	Male	Neutered	LR	13	Bio	Trainee	Socialiser	Failed	Failed
16	Male	Neutered	LR	15	Bio	Trainee	Socialiser	Operational	Success
17	Male	Neutered	LG	18	Assis	Trainee	Socialiser	Failed	Failed
18	Male	Entire	GR	20	Bio	Trainee	Socialiser	Operational	Success
19	Female	Neutered	LR	17	Assis	Trainee	Socialiser	Failed	Failed
20	Male	Neutered	LR	28	Bio	Trainee	Socialiser	Traning	Success
21	Male	Neutered	GR	20	Bio	Trainee	Socialiser	Failed	Failed
22	Female	Entire	LR	15	Bio	Trainee	Socialiser	Operational	Success
23	Female	Entire	CS	16	Bio	Trainee	Socialiser	Operational	Success
24	Female	Neutered	LR	31	Assis	Trainee	Socialiser	Operational	Success
25	Male	Entire	LD	14	Assis	Trainee	Socialiser	Operational	Success
26	Female	Neutered	LR	54	Bio	Trained	Operational	Operational	Success
27	Female	Neutered	LG	16	Assis	Trainee	Socialiser	Operational	Success
28	Female	Neutered	LR	38	Bio	Trained	Failed	Failed	Failed
29	Male	Entire	LG	14	Assis	Trainee	Socialiser	Failed	Failed
30	Male	Neutered	LR	26	Bio	Trainee	Socialiser	Failed	Failed
31	Female	Neutered	LR	23	Bio	Trainee	Socialiser	Failed	Failed
32	Male	Neutered	LG	14	Assis	Trainee	Socialiser	Failed	Failed
33	Female	Neutered	CS	17	Bio	Trainee	Training	Failed	Failed
34	Male	Neutered	LR	20	Bio	Trainee	Failed	Failed	Failed
35	Male	Neutered	LG	31	Assis	Trainee	Failed	Failed	Failed
36	Male	Neutered	LR	21	Bio	Trained	Operational	Failed	Failed
37	Female	Neutered	GR	26	Bio	Trained	Operational	Operational	Success
38	Female	Entire	LR	13	Assis	Trainee	Socialiser	Failed	Failed
39	Male	Neutered	LR	57	Bio	Trained	Operational	Operational	Success
40	Female	Neutered	SS	52	Bio	Trained	Operational	Operational	Success
41	Male	Neutered	CS	12	Bio	Trainee	Socialiser	Traning	Success

42	Female	Neutered	CS	134	Bio	Trained	Operational	Retired	Success
43	Female	Entire	LR	12	Assis	Trainee	Socialiser	Operational	Success
44	Female	Neutered	LG	81	Bio	Trained	Operational	Retired	Success
45	Female	Neutered	LR	64	Bio	Trained	Operational	Operational	Success
46	Female	Neutered	LR	81	Bio	Trained	Operational	Retired	Success
47	Female	Neutered	LR	36	Bio	Trained	Operational	Operational	Success
48	Male	Neutered	CS	90	Bio	Trained	Operational	Operational	Success
49	Male	Neutered	CS	77	Bio	Trained	Operational	Operational	Success
50	Female	Neutered	HV	104	Bio	Trained	Operational	Operational	Success
51	Male	Neutered	LG	40	Bio	Trained	Operational	Operational	Success
52	Male	Neutered	LD	18	Assis	Trainee	Training	Operational	Success
53	Male	Neutered	CS	20	Assis	Trainee	Training	Operational	Success
54	Female	Neutered	LG	50	Bio	Trained	Operational	Operational	Success
55	Male	Neutered	GR	11	Bio	Trainee	Socialiser	Operational	Success
56	Male	Entire	LR	42	Bio	Trained	Operational	Traning	Success
57	Female	Entire	LR	13	Assis	Trainee	Socialiser	Failed	Failed
58	Female	Neutered	LR	35	Bio	Trained	Operational	Operational	Success

Note. Bio=Bio-detection dogs, Assis= Assistance dogs, LR= Labrador Retriever, LG= Labrador/Golden cross, GR= Golden Retriever, CS= Cocker Spaniel, SS= Springer Spaniel, LD=Labradoodle, BC=Border Collie, HV= Hungarian wire hair Visla.

Appendix 5. Rating sheet for trainer's assessment

Trainer's Assessment of Dog

Date:		**Location:**		**Start time:**	
Dog's name:		**Dog's Breed:**		**Dog's Age:**	
Dog's Sex:	**F** ☐ **M** ☐	**Neutered/ Spayed:**	**Y** ☐ **N** ☐		
Your name:		**Approx. How long have you known the dog?**		**Approx. How many times have you work with the dog?**	

Based on **your general impression** of this dog please rate the following traits:

Trait	NA	1 *Very Low*	2 *Low*	3 *Interme-diate*	4 *High*	5 *Very High*	6 *Extremely High*	***Comments***
1. Level of motivation *during training* *(Drive)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
2. Tendency to search by smell alone *(Use of Nose)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
3. Ability to concentrate during task *(Focus)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
4. Confidence in different environments *(Resilience)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
5. Tendency to become frustrated during training	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
6. Ease of adaptation to crate or kennel	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
7. Level of attachment to human partner	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
8. Acuity of sense of smell *(Olfactory acuity)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
9. Tendency to remain specific to the target odour	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
10. Persistence when alerting the presence of a target odour	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
11. Tendency to investigate humans by sniffing	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
12. Tendency to explore areas by sniffing	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
13. Body sensitivity - *reactivity to touch and contact with objects*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
14. Obedience to human command	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
15. Impulsivity - *tendency to make hasty choices during task*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	

Appendix

Trainer Assessment of Dog

Based on **your general impression** of this dog please rate the following traits:								
Trait	NA	1 Very Low	2 Low	3 Interme-diate	4 High	5 Very High	6 Extremely High	**Comments**
16. Independence - ability to work without constant guidance	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
17. Consistency of behaviour from day to day	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
18. Ability to solve problems during training	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
19. Ability to learn from being rewarded	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
20. Willingness to try new behaviours even if they are wrong	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
21. Tendency to vocalise in public places	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
22. Motivation to play with toys	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
23. Motivation to obtain food	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
24. Willingness to bring an object back to a person	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
25. Tendency to seek human attention	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
26. Tendency to be distracted when working	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
27. General excitability - tendency to become highly aroused	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	

28. Please rate this dog's overall ability for medical detection tasks										
One of the worse dogs I have ever seen									One of the best dogs I have ever seen	
	1	2	3	4	5	6	7	8	9	10

Any potential details or observations:

Appendix 6. Dog training rejection form

Any potential problems or general comments?

Approx. how much training have you completed with this dog? ? ☐ sessions ☐ hours

Is the dog being withdrawn?

If so, as of what date?

What is the main reason for its withdrawal?

Appendix 7**.** Rating sheet for internal test

Assessment of Detection Task

Date:		**Location:**		**Start time:**	
Dog's name:		**Dog's Breed:**		**Dog's Age:**	
Dog's sex:	**F** ☐ **M** ☐	**Neutered/ Spayed:**	**Y** ☐ **N** ☐		
Your name:		**Your role:**	**Assessor** ☐	**Trainer** ☐	**Observer** ☐
Assessed in this way before?	**Y**☐ **N**☐	**Previous Assessment dates:**			

Based on the dog's performance **in this test only**, please rate the following traits:

Trait	NA	1 Very Low	2 *Low*	3 *Interme-diate*	4 *High*	5 *Very High*	6 *Extremely High*	***Comments***
1. Level of motivation *during task* *(Drive)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
2.Tendency to search by smell alone *(Use of Nose)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
3. Ability to concentrate during task (*Focus)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
4. Confidence in different environments (*Resilience)*	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
5. Tendency to become frustrated during task	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
6. Acuity of sense of smell (*Olfactory acuity*)	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
7. Tendency to remain specific to the target odour	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
8. Persistence when alerting the presence of a target odour	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
9. Tendency to investigate humans by sniffing	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
10. Tendency to explore areas by sniffing	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	

Appendix

Assessment of Detection Task

Based on the dog's performance **in this test only**, please rate the following traits:								
Trait	NA	1 Very Low	2 Low	3 Interme-diate	4 High	5 Very High	6 Extremely High	***Comments***
11. Independence - ability to work without constant guidance	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
12. Consistency of behaviour from day to day	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
13. Ability to solve problems during task	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
14. Ability to learn from being rewarded	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
15. Willingness to try new behaviours even if they are wrong	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
16. Motivation to play with toys	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
17. Motivation to obtain food	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
18. Willingness to bring an object back to a person	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
19. Tendency to seek human attention	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
20. Tendency to be distracted when working	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	
21. General excitability - tendency to become highly aroused	☐ NA	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	

22. Please rate this dog's overall ability for medical detection tasks											
One of the worse dogs I have ever seen									**One of the best dogs I have ever seen**		
	1	2	3	4	5	6	7	8	9	10	

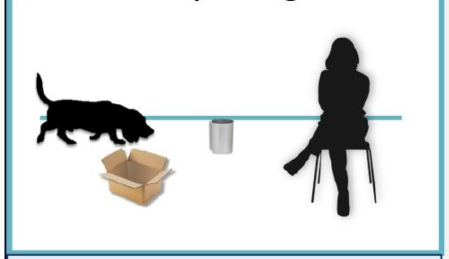
Any extra comments:

Appendix

Chapter 4

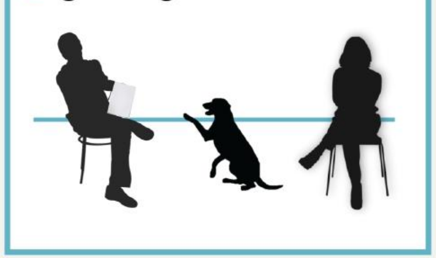
Appendix 8**.** Diagram with test battery's instructions

1. Room exploring



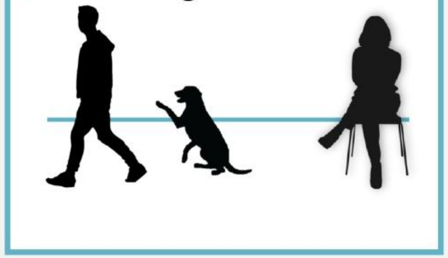
1. Please enter the room, take your dog's lead off and sit down.
2. Let them explore freely, unless they start to chew the objects in the room.
3. Please hold the folder located in your chair while we chat about your dog.

2. Ignoring



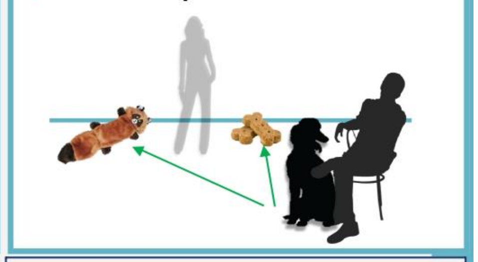
1. Ignore your dog for three minutes.
2. Meanwhile, please complete the questionnaire in the folder.
3. If you finish before the time is up, you may read the additional material inside the folder. Keep ignoring your dog until indicated.

3. Following



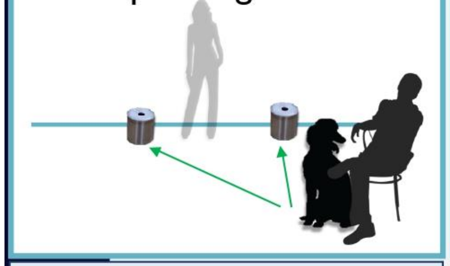
1. Please stand up and walk around the room in a relaxed way.
2. Ignore your dog.
3. Praise them when asked.

4. Reward preference



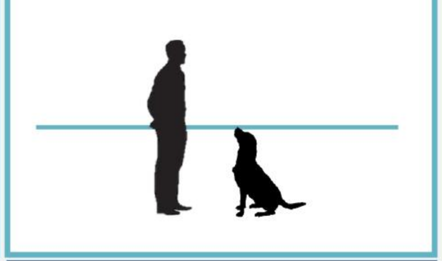
1. Please Sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to choose between a treat or a toy.
3. After your dog makes their first choice or if I say "STOP" call them back, if they do not come, please take them back to the start point.

5. Arm pointing



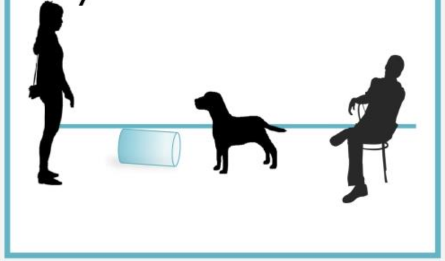
1. Please sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to choose between two cans on the floor.
3. After your dog makes their first choice or if I say "STOP" call them back, if they do not come, please take them back to the start point.

6. Obedience



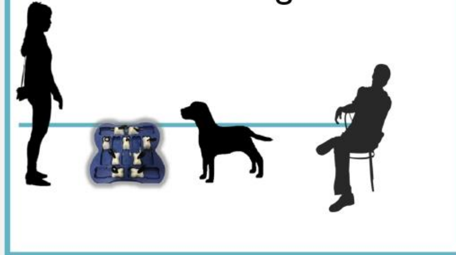
1. Ask your dog to sit, lay down and stay.
2. Walk five steps backwards and hold the stay for 30 sec.
3. When asked, call your dog back and praise them.
4. Please repeat the task when asked.

7. Cylinder



1. Please sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to take a treat out of an open ended cylinder.
3. Meanwhile, please remain in your seat quietly and ignore them if they try to get your attention.
4. This will be repeated a number of trials

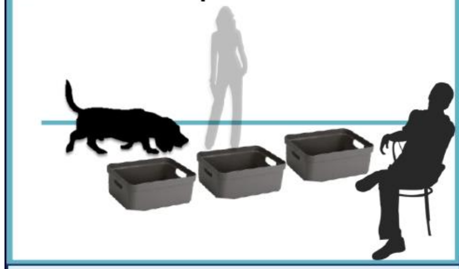
8. Problem solving



1. Please sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to explore the puzzle.
3. Meanwhile, please remain in your seat quietly and ignore them if they try to get your attention.

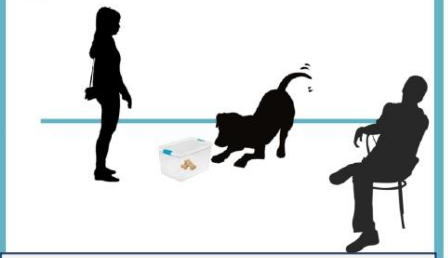
Appendix

9. Boxes exploration



1. Please sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to search for treats hidden in boxes.
3. After your dog finds the treat or if I say "STOP" call them back, if they do not come, please take them back to the start point.

11. Unsolvable task

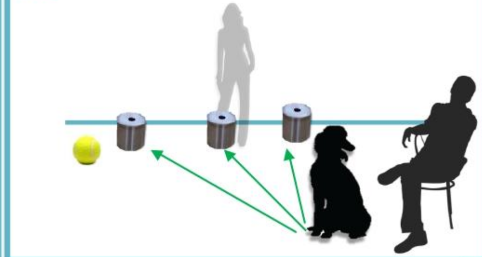


1. Please sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to try to take a treat out of a closed container.
3. Meanwhile, please remain in your seat quietly and ignore them if they try to get your attention.

12. Toy playing

1. Please play tug of war with your dog for one minute.
2. When I say "STOP", ask your dog to sit, lay down and stay.
3. Walk five steps backwards and hold the stay for 30 sec.
4. When asked, call your dog back and praise them.

13. Tennis ball



1. Please Sit and hold your dog behind the start line facing forward.
2. When I say "READY" release your dog, they are free to find a ball hidden under a can.
3. If your dog chooses the right can, I will throw them the ball, and you may release them. If they do not give up the ball, please take it back from them.

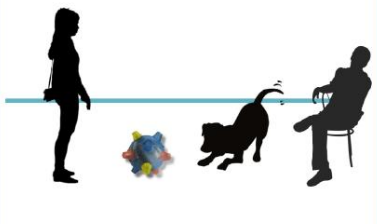
14. Coat wearing

1. Please put the coat on your dog while I hold them.
2. We will then let them go, and they are free to explore for 30 sec.
3. please take off the coat afterwards.

15. Vet check up

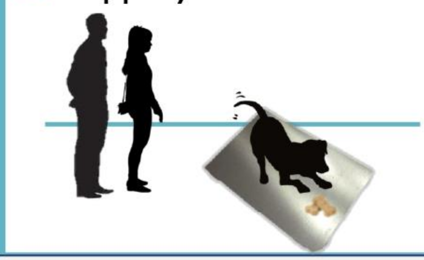
1. Please hold your dog while I simulate a vet check up.
2. Please let me know beforehand if your dog does not like being handled.

16. Moving ball



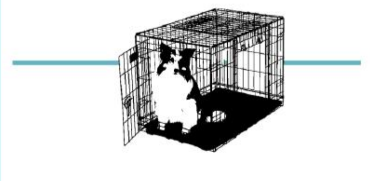
1. Please sit and hold your dog behind the start line facing forward.
2. I will show your dog a ball. Release them when I place the ball on the floor. The ball will move and make a noise for around 30 sec.
3. Meanwhile, please remain in your seat quietly and ignore your dog if they try to get your attention.

17. Slippery surface



1. Please follow me into the next room.
2. Hold your dog's leash while I toss some treats on a slippery surface.
3. When I said "READY" release them so they can get the treats

18. Adaptation to resting area



1. Please take your dog back to their crate (resting area).
2. Toss some treats in the crate and put your dog inside
3. We will leave them alone for five minutes while filming them.
4. Afterwards you may take them out if you wish.

Appendix

Appendix 9. Test battery ethogram

Test on Medical detection dog performance: part 1					
Date:		**Location:**		**Start time:**	
Dog's name:		**Dog's Breed:**		**Dog's Age:**	
Dog's sex:	**F** ☐ **M** ☐	**Neutered/ Spayed:**	**Y**☐ **N** ☐		
name:		**role:**			

S1. Exploring (Rooney, 2003a)		
Trait assessed	**Variable name**	
Tendency to explore areas by sniffing	Duration investigating environment within 1 min (*Nose <10 cm from object, floor or walls)*	Sec
Tendency to investigate humans by sniffing	Duration smelling experimenter throughout 1 min (*Nose <10 cm from person)*	Sec
Friendliness to new people	Latency to approach unfamiliar person: Latency to get close to experimenter for the first time, making contact or <50 cm within 1 min	Sec
	Human interaction duration: Time dog spends making contact or < 50 cm, from and paying attention human within 1 min.	Handler: Sec
		Exp: Sec

S2. Ignoring (Rooney, 2003a)		
Trait assessed	**Variable name**	
Level of attachment to human partner/ Tendency to seek human attention	Contact frequency: Number of times when the dog contacts person (within 3 min)	Handler: Freq
		Exp: Freq
	Interaction duration: Time dog spends making contact or < 50 cm, from and paying attention to handler within 3 min.	Handler: Sec
		Exp: Sec

S3. Following (Rooney, 2003a)		
Trait assessed	**Variable name**	
Level of attachment to human partner	Handler interaction duration: Time dog spends making contact or < 50 cm, from and paying attention to Handler (within 1 min)	Sec

S4. Reward preference (MacLean et al., 2017)		
Trait assessed	**Variable name**	
Motivation to obtain toys vs. food	**Number of trials in which the dog chooses food vs. toy amongst six trials. (frequency).** *(Dog makes contact or holds toy, or eats the food)*	
Trial	**Toy**	**Food**
1		
2		
3		
4		
5		
6		
Percentage		

Reward motivation level	Subjective ratings relative to reach first selected reward with a scale of 1 to 5					
No apparent interest in toy/ food	**1** **Very Low** ☐ ☐	**2** ***Low*** ☐ ☐	**3** **Intermediate** ☐ ☐	**4.** **High** ☐ ☐	**5.** **Very High** ☐ ☐	**Most Intense response towards reward**

Appendix

S5. Arm pointing (Hare et al., 1998; Miklösi et al., 1998; MacLean et al., 2017)		
Trait assessed	**Variable name**	
Communication/ Cooperation with handler	**Occasions locating correct can: percentage of trials that a subject's first touch the pointed container with nose (or approaches nose <10 cm) or points container with trained indication within 6 trials (frequency)**	
Trial	**Yes**	**No**
1		
2		
3		
4		
5		
6		
percentage		

S6. Obedience to human commands (Rooney, 2003a)			
Trait assessed	**Variable name**	***1st***	***2nd***
Sit	Latency to obey: Latency to perform the task after saying the command within two trials.		
Lie down	Latency to obey: Latency to perform the task after saying the command within two trials.		
Stay	Latency to rise, --Duration to hold the "Stay" task without breaking it within 30 sec.		
total	Average		

S7. Cylinder (Bray et al., 2014; MacLean et al., 2017; Tiira et al. 2020)					
Trait assessed	Variable name				
Impulsivity/ response inhibition	Initial response performing a detour to take out the food from one of the cylinder's edges without nosing, pushing or pawing the transparent cylinder.				
Trial	**Yes**		**No**	**Latency**	
1					**sec**
2					**sec**
3					**sec**
4					**sec**
5					**sec**
6					**Sec**
total				**Mean**	**sec**
Behaviours presented	**Pawing/ scratching** ☐	**pushing with snooth** ☐	**Head fully inside** ☐	**biting** ☐	**Throwing** ☐
	Licking ☐	**Smelling** ☐	**kicking** ☐	**Holding** ☐	**Bark/grawl** ☐
Other behaviours					

S8. Puzzle (Bray et al., 2017b)		
Trait assessed	**Variable name**	
Ability to solve problems when working	Number of treats recovered (out of 20) within 3 minutes	Freq
	Time engaged with task within 3 minutes	Sec
Persistence when working/Independence	Latency to gaze human: latency to gaze handler	Sec
	Latency to gaze human: latency to gaze tester	Sec
	Total time (within 3 min) looking at experimenter or handler	Sec

Appendix

Test on Medical detection dog performance: part 2

S9. Boxes searching (Rooney, 2003a)

Trait assessed	Variable name	*Trial 1*	*Trial 2*	*Trial 3*	*AV*
Tendency to search by smell alone	Latency to find each treat				
	Duration in non searching activity				
	Number of treats located (one treat per trial) Freq				

S10. Noise distraction (Svartberg, 2002; Batt et al., 2008)

Trait assessed	Variable name	
Tendency to be distracted by novel noise	Time engaged in task while finding the treat (duration)	Sec
	Time attending to noise while finding the treat (duration)	Sec
	Latency to resume task	Sec

S11. Unsolvable task (Miklösi et al., 2003; MacLean et al., 2017)

Trait assessed	Variable name	Exp.	Handler
Persistence when working/Independence	duration of dog gazing the experimenter or handler within 1 min.	Sec	Sec
	Latency to gaze experimenter or handler wtihin 1 min.	Sec	Sec

Behaviours presented	Pawing/ scratching ☐	pushing with snooth ☐	Head fully inside ☐	biting ☐	Throwing ☐
	Licking ☐	Smelling ☐	kicking ☐	Holding ☐	Bark/grawl ☐
Other behaviours:					

S12. Ball searching

Trait assessed	Variable name
Acuity of sense of smell	Occasions locating correct can: number of occasions on which the dog chooses the bated pot (touching pot with snout or paw) where 0= fail, 1= found with intervention, 2= found (scoring 0-6)
Willingness to bring an object back to a person	retrieving ball: number of occasions where the ball is retrieved to experimenter, where 0= not retrieved, 1= retrieved with intervention, 2= retrieved to handler, 2= retrieved to experimenter (scoring 0-12)

Trial (3 total)	Ball located		Ball retrieved			
	Yes (1)	No (0)	Exp (2)	Hand (2)	Interv (1)	No (0)
1						
2						
3						
4						
5						
6						
Total						

S13. Toy playing (Rooney, 2003a)

Trait assessed	Variable name							
Motivation to play with toys	Playing duration: Play duration over 60 seconds (time)							Sec
Excitability: latency to obey	Leave	Sec	Sit	Sec	Down	Sec	Stay	Sec

Playing intensity level	Subjective ratings for intensity of involvement with tug of war game with a scale of 1 to 5					
No apparent interest in game	1 Very Low ☐	2 *Low* ☐	3 Intermediate ☐	4. High ☐	5. Very High ☐	Most Intense response towards game

Shake	Tug	Chew	Hold	Throw	Chase

Appendix

S14. Coat wearing						
Trait assessed	**Variable name**					
Body sensitivity: Coat tolerance	Duration wearing the coat without struggling within 30 seconds.					Sec
Level of restraint required	Subjective ratings for the level of restraint required when putting coat on of 1 to 5 ***Ear height, tail height, distress behaviours**					
Stays still, relaxed	**1 Very Low ☐ Staying still, relaxed**	**2 *Low* ☐ Staying still, rigid**	**3 Intermediate ☐ Moving around but tolerating**	**4. High ☐ Trying to escape**	**5. Very High ☐ Does not tolerate**	**Struggles continuously**

S15. Body condition check-up (Harvey et al., 2016)						
Level of tolerance to vet check up	Subjective ratings for the level of dog's tolerance to vet check up with a scale on of 1 to 5 ***Ear height, tail height, distress behaviours**					
Tries to escape	**1 Very Low ☐ Does not tolerate**	**2 *Low* ☐ Trying to escape**	**3 Intermediate ☐ Moving around but tolerating**	**4. High ☐ Staying still, rigid**	**5. Very High ☐ Staying still, relaxed**	**Stays still, relaxed**

S16. Novel object (Rooney, 2003a)						
Level of approach towards new object	Subjective ratings of dog's response to a novel object with a scale on of 1 to 5 ***Ear height, tail height, distress behaviours**					
Avoids object	**1 Very Low ☐ Retreat**	**2 *Low* ☐ Step back**	**3 Intermediate ☐ No movement**	**4. High ☐ Moving back and forth**	**5. Very High ☐ Approach**	**Approaches to object**

S17. Slippery surface (Rooney, 2003a)		
Trait assessed	**Variable name**	
Confidence, absence of fear	Latency to get on surface: mean latency to get on the surface with the 4 paws in it	Sec
	Duration on surface: Total time on Surface with the 4 paws in it (1 min max.)	Sec

Level of Approach to new surface	Subjective ratings of dogs response to a novel surface with a scale on of 1 to 5 (Score 0-5) ***Ear height, tail height, distress behaviours**					
Do not step on surface	**1 Very Low ☐ Do not step on surface**	**2 *Low* ☐ Hesitant to step on**	**3 Intermediate ☐ Tentative to step on**	**4. High ☐ Step on surface**	**5. Very High ☐ Jumps straight on**	**Jumps straight on surface**

S18. Holding pen test		
Trait assessed	**Variable name**	
Adaptation to crate or kennel	Latency to rest: Latency to settle, lay down, display relaxed body postures	Sec
	Duration resting: Total time laying down, relaxed within 5 minutes	Sec
	Duration moving: Total time walking in pen within 5 minutes	Sec
	Duration exploring: Total time investigating environment within 1 min (*Nose <10 cm from object, floor or walls/ fences)*	Sec
Tendency to vocalise	duration of vocalisations: duration vocalising over 5 minutes.	Sec

Appendix

Task N◦									
Behaviour	First part / Incidence of stress behaviours	S1 Exploring	S2 Ignoring	S3 Following	S4 Reward	S5 Arm Pointing	S6 Obedience	S7 Cylinder	S8 puzzle
	Body posture **(1 Crouched)**								
	Ear height (1 ears flattened back)								
	Tail height (1-tucked between legs or straight down								
	Tremble (0,1)								
	Yawn (0,1)								
	Lip lick (0,1)								
	Paw lift (0,1)								
	Body shake (0,1)								
	Step back/avoid (0,1)								
	Whine (0,1)								
	Bark (0,1)								
	Growl (0,1)								
	Bare teeth (0,1)								
	Mouth (0,1)								
	Snap/Bite (0,1)								
	Jump up (0,1)								
	Play bow (0,1)								
	Lick handler (0,1)								
	Lick exp. (0,1)								
	Drink (0,1)								

Appendix

	Task N◦										
Behaviour	Second part	**9 Box search**	**10 Distract ion**	**11 Unsolv able task**	**12 ball search**	**13 Toy playing**	**S14 Coat Wearing**	**15 Vet Check up**	**16 Novel object**	**17 Surfa ce**	**18 Holding pen**
	Body posture (0=low, 1=relaxed, neutral, 2=high and rigid)										
	Ear height (0=flattened back, 1= relaxed 2= rigid and forward)										
	Tail height (0= ducked between legs, 1=hanging down, 2= levelled, 3= upwright)										
	Tremble (0,1)										
	Yawn (0,1)										
	Lip lick (0,1)										
	Paw lift (0,1)										
	Body shake (0,1)										
	Whine (0,1)										
	Bark (0,1)										
	Growl (0,1)										
	Bare teeth (0,1)										
	Mouth (0,1)										
	Snap/Bite (0,1)										
	Jump up (0,1)										
	Play bow (0,1)										
	Lick handler (0,1)										
	Lick exp. (0,1)										
	Drink (0,1)										

Appendix 10. Questionnaire on dogs' details for trainers

Date:____________ Dogs name: _____________ Handler_______________

Questionnaire for trainers: Preliminary dog information

Thank you so much for participating in this study and taking time out of your day.

1. What breed is your dog?

2. Approx. How old are they?

3. What sex are they?

F _____ M_____

4. Have they been neutered/ spayed? If so, approx. When?

Y____ N____ when?_______

5. What role are they trained to do?

Bio_____ MAD_____ Dual _____ Other_______________

Details:

6. Are you their trainer, instructor, or handler?

Trainer_____ Instructor_____ Handler______

7. How long have you been working with them? And approximately how many times do you think you've met them?

Working Length _________ Times meeting per week ____________

8. Who do they currently live with?

Socializer ____ Client____ You ___ Another member of staff? ___ Other ______________

9. Who would you say they are most attached to?

You _____ Another trainer _____ Socializer _____ Other_____________________

1. How attached would you say they are to you on a scale of 1–10, with ten being the most?

1 2 3 4 5 6 7 8 9 **10**

References

2019. Personal communication with trainers from Medical Detection Dogs charity.

Adamkiewicz, E., Jezierski, T., Walczak, M., Gorecka-Bruzda, A., Sobczynska, M., Prokopczyk, M., Ensminger, J., 2013. Traits of drug and explosives detection in dogs of two breeds as evaluated by their handlers and trainers. Animal Science Papers and Reports 31, 205-217.

Arandjelovic, M., Bergl, R.A., Ikfuingei, R., Jameson, C., Parker, M., Vigilant, L., Arandjelovic et al., 2015. Detection dog efficacy for collecting faecal samples from the critically endangered Cross River gorilla (Gorilla gorilla diehli) for genetic censusing. R. Soc. Open Sci. 2, 14.

Arata, S., Momozawa, Y., Takeuchi, Y., Mori, Y., 2010. Important behavioral traits for predicting guide dog qualification. Journal of Veterinary Medical Science, 0912080094-0912080094.

Audrestch, H.M., Whelan, C.T., Grice, D., Asher, L., England, G.C.W., Freeman, S.L., 2015. Recognizing the value of assistance dogs in society. Disability and Health Journal 8, 469-474.

Bandura, A., 2000. Self-efficacy: The foundation of agency. Control of human behavior, mental processes, and consciousness: Essays in honor of the 60th birthday of August Flammer 16.

Banfield, C.M., Bartels, J.E., Hudson, J.A., Wright, J.C., Hathcock, J.T., Montgomery, R.D., 1996. A retrospective study of canine hip dysplasia in 116 military working dogs .1. Angle measurements and Orthopedic Foundation for Animals (OFA) grading. Journal of the American Animal Hospital Association 32, 413-422.

Banlaki, Z., Cimarelli, G., Viranyi, Z., Kubinyi, E., Sasvari-Szekely, M., Ronai, Z., 2017. DNA METHYLATION PATTERNS OF SEVERAL BEHAVIOUR-RELATED GENE PROMOTERS: A COMPARATIVE ANALYSIS BETWEEN DOMESTIC DOG BREEDS AND THE GREY WOLF. European Neuropsychopharmacology 27, S170-S171.

Barnard, S., Siracusa, C., Reisner, I., Valsecchi, P., Serpell, J.A., 2012. Validity of model devices used to assess canine temperament in behavioral tests. Applied Animal Behaviour Science 138, 79-87.

Barnard, S., Wells, D.L., Milligan, A.D., Arnott, G., Hepper, P.G., 2018. Personality traits affecting judgement bias task performance in dogs (Canis familiaris). Scientific reports 8, 1-8.

Batt, L.S., Batt, M.S., Baguley, J.A., McGreevy, P.D., 2008. Factors associated with success in guide dog training. Journal of Veterinary Behavior-Clinical Applications and Research 3, 143-151.

Batt, L.S., Batt, M.S., Baguley, J.A., McGreevy, P.D., 2009. The relationships between motor lateralization, salivary cortisol concentrations and behavior in dogs. Journal of Veterinary Behavior-Clinical Applications and Research 4, 216-222.

Beebe, S.C., Howell, T.J., Bennett, P.C., 2016. Using scent detection dogs in conservation settings: a review of scientific literature regarding their selection. Front. Vet. Sci. 3, 96.

Beerda, B., Schilder, M.B.H., Bernadina, W., Van Hooff, J., De Vries, H.W., Mol, J.A., 1999. Chronic stress in dogs subjected to social and spatial restriction. II. Hormonal and immunological responses. Physiology & Behavior 66, 243-254.

Beerda, B., Schilder, M.B.H., van Hooff, J., de Vries, H.W., Mol, J.A., 1998. Behavioural, saliva cortisol and heart rate responses to different types of stimuli in dogs. Applied Animal Behaviour Science 58, 365-381.

References

Beerda, B., Schilder, M.B.H., van Hooff, J., de Vries, H.W., Mol, J.A., 2000. Behavioural and hormonal indicators of enduring environmental stress in dogs. Animal Welfare 9, 49-62.

Berg, A.T., 2011. Epilepsy, cognition, and behavior: the clinical picture. Epilepsia 52, 7-12.

Bethell, E.J., Holmes, A., MacLarnon, A., Semple, S., 2012. Evidence that emotion mediates social attention in rhesus macaques.

Beyenburg, S., Mitchell, A.J., Schmidt, D., Elger, C.E., Reuber, M., 2005. Anxiety in patients with epilepsy: systematic review and suggestions for clinical management. Epilepsy & Behavior 7, 161-171.

Bijland, L.R., Bomers, M.K., Smulders, Y.M., 2013. Smelling the diagnosis A review on the use of scent in diagnosing disease. Neth. J. Med. 71, 300-307.

Bishop, S.J., 2007. Neurocognitive mechanisms of anxiety: an integrative account. Trends in cognitive sciences 11, 307-316.

Blackwell, E.J., Twells, C., Seawright, A., Casey, R.A., 2008. The relationship between training methods and the occurrence of behavior problems, as reported by owners, in a population of domestic dogs. Journal of Veterinary Behavior-Clinical Applications and Research 3, 207-217.

Boissy, A., 1995. FEAR AND FEARFULNESS IN ANIMALS. Quarterly Review of Biology 70, 165-191.

Bomers, M.K., Van Agtmael, M.A., Luik, H., Van Veen, M.C., Vandenbroucke-Grauls, C.M., Smulders, Y.M., 2012. Using a dog's superior olfactory sensitivity to identify Clostridium difficile in stools and patients: proof of principle study. Bmj 345.

Bonanno, G.A., 2005. Resilience in the face of potential trauma. Current directions in psychological science 14, 135-138.

Bonanno, G.A., Mancini, A.D., 2011. Toward a lifespan approach to resilience and potential trauma, Resilience and Mental Health: Challenges Across the Lifespan, pp. 120-134.

Bradshaw, J.W.S., Goodwin, D., Lea, A.M., Whitehead, S.L., 1996. A survey of the behavioural characteristics of pure-bred dogs in the United Kingdom. Veterinary Record 138, 465-468.

Brady, K., Cracknell, N., Zulch, H., Mills, D.S., 2018. Factors associated with long-term success in working police dogs. Applied Animal Behaviour Science 207, 67-72.

Braem, M., Asher, L., Furrer, S., Lechner, I., Würbel, H., Melotti, L., 2017. Development of the "Highly Sensitive Dog" questionnaire to evaluate the personality dimension "Sensory Processing Sensitivity" in dogs. PloS one 12, e0177616.

Bray, E.E., Gruen, M.E., Gnanadesikan, G.E., Horschler, D.J., Levy, K.M., Kennedy, B.S., Hare, B.A., MacLean, E.L., 2020. Cognitive characteristics of 8-to 10-week-old assistance dog puppies. Animal behaviour 166, 193-206.

Bray, E.E., Gruen, M.E., Gnanadesikan, G.E., Horschler, D.J., Levy, K.M., Kennedy, B.S., Hare, B.A., MacLean, E.L., 2021a. Dog cognitive development: a longitudinal study across the first 2 years of life. Animal cognition 24, 311-328.

Bray, E.E., Levy, K.M., Kennedy, B.S., Duffy, D.L., Serpell, J.A., MacLean, E.L., 2019. Predictive models of assistance dog training outcomes using the canine behavioral assessment and research questionnaire and a standardized temperament evaluation. Front. Vet. Sci. 6, 49.

Bray, E.E., MacLean, E.L., Hare, B.A., 2014. Context specificity of inhibitory control in dogs. Animal Cognition 17, 15-31.

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

References

Bray, E.E., Otto, C.M., Udell, M.A., Hall, N.J., Johnston, A.M., MacLean, E.L., 2021b. Enhancing the selection and performance of working dogs. Front. Vet. Sci. 8, 430.

Bray, E.E., Sammel, M.D., Cheney, D.L., Serpell, J.A., Seyfarth, R.M., 2017a. Effects of maternal investment, temperament, and cognition on guide dog success. Proceedings of the National Academy of Sciences of the United States of America 114, 9128-9133.

Bray, E.E., Sammel, M.D., Seyfarth, R.M., Serpell, J.A., Cheney, D.L., 2017b. Temperament and problem solving in a population of adolescent guide dogs. Animal Cognition 20, 923-939.

Bremhorst, A., Sutter, N.A., Würbel, H., Mills, D.S., Riemer, S., 2019. Differences in facial expressions during positive anticipation and frustration in dogs awaiting a reward. Scientific reports 9, 1-13.

Brewer, M.B., 1979. In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. Psychological bulletin 86, 307.

Broach, D., Dunham, A.E., 2016. Evaluation of a pheromone collar on canine behaviors during transition from foster homes to a training kennel in juvenile Military Working Dogs. Journal of Veterinary Behavior-Clinical Applications and Research 14, 41-51.

Brown, S.W., Goldstein, L.H., 2011. Can seizure-alert dogs predict seizures? Epilepsy Research 97, 236-242.

Browne, C.S., K. Fordham, R., 2006. The use of scent-detection dogs. Irish Vet. J. 59, 97-+.

Brownell, D., 2002. The Brownell-Marsolais scale: a proposal for the qualitative evaluation of SAR/disaster K9 candidates. Adv Rescue Technol 5, 57-67.

Brunner, D., Hen, R., 1997. Insights into the neurobiology of impulsive behavior from serotonin receptor knockout mice. Neurobiology of Suicide: from the Bench to the Clinic 836, 81-105.

Bryła, P., 2015. The impact of international student mobility on subsequent employment and professional career: A large-scale survey among Polish former Erasmus students. Procedia-Social and behavioral sciences 176, 633-641.

Burman, O., 2014. Do Dogs Show an Optimistic or Pessimistic Attitude to Life?: A Review of Studies Using the 'Cognitive Bias' Paradigm to Assess Dog Welfare. The Social Dog, 347-372.

Burman, O.H., Parker, R.M., Paul, E.S., Mendl, M.T., 2009. Anxiety-induced cognitive bias in non-human animals. Physiology & Behavior 98, 345-350.

Button, K.S., Ioannidis, J., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nature reviews neuroscience 14, 365-376.

Cablk, M.E., Heaton, J.S., 2006. Accuracy and reliability of dogs in surveying for desert tortoise (Gopherus agassizii). Ecol. Appl. 16, 1926-1935.

Caeiro, C., Guo, K., Mills, D., 2017. Dogs and humans respond to emotionally competent stimuli by producing different facial actions. Scientific reports 7, 1-11.

Campbell, W.E., 1972. A behavior test for puppy selection. Mod Vet Pract 12, 29-33.

Caron-Lormier, G., Harvey, N.D., England, G.C.W., Asher, L., 2016. Using the incidence and impact of behavioural conditions in guide dogs to investigate patterns in undesirable behaviour in dogs. Scientific Reports 6, 9.

Catala, A., Grandgeorge, M., Schaff, J.-L., Cousillas, H., Hausberger, M., Cattet, J., 2019. Dogs demonstrate the existence of an epileptic seizure odour in humans. Scientific reports 9, 1-7.

References

Catala, A., Latour, P., Martinez-Caja, A.M., Cousillas, H., Hausberger, M., Grandgeorge, M., 2020. Is there a Profile of Spontaneous Seizure-Alert Pet Dogs? A Survey of French People with Epilepsy. Animals 10, 254.

Chen, M., Daly, M., Natt, Susie, Candy, Williams, G., 2000. Non-invasive detection of hypoglycaemia using a novel, fully biocompatible and patient friendly alarm system. Br. Med. J. 321, 1565-1566.

Clark, C.C., Sibbald, N.J., Rooney, N.J., 2020a. Search dog handlers show positive bias when scoring their own dog's performance. Front. Vet. Sci., 612.

Clark, C.C.A., Rooney, N.J., 2021. Does Benchmarking of Rating Scales Improve Ratings of Search Performance Given by Specialist Search Dog Handlers? Front. Vet. Sci. 8, 11.

Clark, C.C.A., Sibbald, N.J., Rooney, N.J., 2020b. Search Dog Handlers Show Positive Bias When Scoring Their Own Dog's Performance. Front. Vet. Sci. 7, 9.

Cobb, M., Branson, N., McGreevy, P., Lill, A., Bennett, P., 2015. The advent of canine performance science: Offering a sustainable future for working dogs. Behavioural Processes 110, 96-104.

Cobb, M.L., Iskandarani, K., Chinchilli, V.M., Dreschel, N.A., 2016. A systematic review and meta-analysis of salivary cortisol measurement in domestic canines. Domest. Anim. Endocrinol. 57, 31-42.

Cockburn, A., Smith, M., Rusbridge, C., Fowler, C., Paul, E.S., Murrell, J.C., Blackwell, E.J., Casey, R.A., Whay, H.R., Mendl, M., 2018. Evidence of negative affective state in Cavalier King Charles Spaniels with syringomyelia. Applied animal behaviour science 201, 77-84.

Concha, A., Mills, D.S., Feugier, A., Zulch, H., Guest, C., Harris, R., Pike, T.W., 2014. Using Sniffing Behavior to Differentiate True Negative from False Negative Responses in Trained Scent-Detection Dogs. Chemical Senses 39, 749-754.

Concha, A.R., Guest, C.M., Harris, R., Pike, T.W., Feugier, A., Zulch, H., Mills, D.S., 2019. Canine Olfactory Thresholds to Amyl Acetate in a Biomedical Detection Scenario. Front. Vet. Sci. 5.

Conti, G., Pudney, S., 2011. Survey design and the analysis of satisfaction. Review of Economics and Statistics 93, 1087-1093.

Cornu, J.N., Cancel-Tassin, G., Ondet, V., Girardet, C., Cussenot, O., 2011. Olfactory detection of prostate cancer by dogs sniffing urine: A step forward in early diagnosis. European Urology 59, 197-201.

Dalla Villa, P., Barnard, S., Di Fede, E., Podaliri, M., Candeloro, L., Di Nardo, A., Siracusa, C., Serpell, J.A., 2013. Behavioural and physiological responses of shelter dogs to long-term confinement. Veterinaria Italiana 49, 231-241.

Dalziel, D.J., Uthman, B.M., Mcgorray, S.P., Reep, R.L., 2003. Seizure-alert dogs: a review and preliminary study. Seizure 12, 115-120.

Davies, J.C., Alton, E., Simbo, A., Murphy, R., Seth, I., Williams, K., Somerville, M., Jolly, L., Morant, S., Guest, C., 2019. Training dogs to differentiate Pseudomonas aeruginosa from other cystic fibrosis bacterial pathogens: not to be sniffed at? European Respiratory Journal 54.

Davis, P., 2017. The Investigation of Human Scent from Epileptic Patients for the Identification of a Biomarker for Epileptic Seizures.

De Palma, C., Viggiano, E., Barillari, E., Palme, R., Dufour, A.B., Fantini, C., Natoli, E., 2005. Evaluating the temperament in shelter dogs. Behaviour 142, 1307-1328.

de Ridder, D.T., de Boer, B.J., Lugtig, P., Bakker, A.B., van Hooft, E.A., 2011. Not doing bad things is not equivalent to doing the right thing: Distinguishing between inhibitory and initiatory self-control. Personality and Individual Differences 50, 1006-1011.

References

DeMatteo, K.E., Blake, L.W., Young, J.K., Davenport, B., 2018. How Behavior of Nontarget Species Affects Perceived Accuracy of Scat Detection Dog Surveys. Scientific Reports 8, 11.

Descovich, K.A., Wathan, J., Leach, M.C., Buchanan-Smith, H.M., Flecknell, P., Farningham, D., Vick, S.J., 2017. Facial Expression: An Under-Utilized Tool for the Assessment of Welfare in Mammals. Altex-Alternatives to Animal Experimentation 34, 409-429.

Diederich, C., Giffroy, J.M., 2006. Behavioural testing in dogs: A review of methodology in search for standardisation. Applied Animal Behaviour Science 97, 51-72.

Diverio, S., Barbato, O., Cavallina, R., Guelfi, G., Iaboni, M., Zasso, R., Di Mari, W., Santoro, M.M., Knowles, T.G., 2016. A simulated avalanche search and rescue mission induces temporary physiological and behavioural changes in military dogs. Physiology & Behavior 163, 193-202.

Diverio, S., Menchetti, L., Riggio, G., Azzari, C., Iaboni, M., Zasso, R., Di Mari, W., Santoro, M.M., 2017. Dogs' coping styles and dog-handler relationships influence avalanche search team performance. Applied Animal Behaviour Science 191, 67-77.

Dogs, M.D., 2020. Medical Detection Dogs. About us. Retrieved from https://www.medicaldetectiondogs.org.uk/about-us/.

Doherty, M.J., Haltiner, A.M., 2007. Wag the dog: skepticism on seizure alert canines. Neurology 68, 309-309.

Doyle, R.E., Fisher, A.D., Hinch, G.N., Boissy, A., Lee, C., 2010. Release from restraint generates a positive judgement bias in sheep. Applied Animal Behaviour Science 122, 28-34.

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

Duffy, D.L., Serpell, J.A., 2012. Predictive validity of a method for evaluating temperament in young guide and service dogs. Applied Animal Behaviour Science 138, 99-109.

Duranton, C., Bedossa, T., Gaunet, F., 2016. When facing an unfamiliar person, pet dogs present social referencing based on their owners' direction of movement alone. Animal Behaviour 113, 147-156.

Duranton, C., Horowitz, A., 2019. Let me sniff! Nosework induces positive judgment bias in pet dogs. Applied Animal Behaviour Science 211, 61-66.

Eckel, L.A., 2011. The ovarian hormone estradiol plays a crucial role in the control of food intake in females. Physiology & behavior 104, 517-524.

Edney, A., 1993. Dogs and human epilepsy. The Veterinary Record 132, 337-338.

Edwards, T.L., Browne, C.M., Schoon, A., Cox, C., Poling, A., 2017. Animal olfactory detection of human diseases: Guidelines and systematic review. Journal of veterinary behavior 20, 59-73.

Elliker, K.R., Sommerville, B.A., Broom, D.M., Neal, D.E., Armstrong, S., Williams, H.C., 2014. Key considerations for the experimental training and evaluation of cancer odour detection dogs: lessons learnt from a double-blind, controlled trial of prostate cancer detection. BMC Urol. 14, 9.

Eysenck, M.W., Mogg, K., May, J., Richards, A., Mathews, A., 1991. Bias in interpretation of ambiguous sentences related to threat in anxiety. Journal of abnormal psychology 100, 144.

Fadel, F.R., Driscoll, P., Pilot, M., Wright, H., Zulch, H., Mills, D., 2016. Differences in Trait Impulsivity Indicate Diversification of Dog Breeds into Working and Show Lines. Scientific Reports 6, 10.

Fallani, G., Previde, E.P., Valsecchi, P., 2007. Behavioral and physiological responses of guide dogs to a situation of emotional distress. Physiology & Behavior 90, 648-655.

References

Fanelli, C.G., Porcellati, F., Pampanelli, S., Bolli, G.B., 2004. Insulin therapy and hypoglycaemia: the size of the problem. Diabetes-Metab. Res. Rev. 20, S32-S42.

Fatjó, J., 2001. Impulsivity in Dogs—Assessment and Treatment World Small Animal Veterinary Association World Congress Proceedings, 2001.

Fatjo, J., Amat, M., Mariotti, V., de la Torre, J.R., Manteca, X., 2005. Aggression in dogs: analysis of 761 cases. MILLS, D. et al. Current issues and research in veterinary behavioral medicine. Indiana: Purdue University, 251-254.

Field, A., 2013. Discovering statistics using IBM SPSS statistics. sage.

Fischer-Tenhagen, C., Wetterholm, L., Tenhagen, B.A., Heuwieser, W., 2011. Training dogs on a scent platform for oestrus detection in cows. Applied Animal Behaviour Science 131, 63-70.

Flint, H.E., Coe, J.B., Serpell, J.A., Pearl, D.L., Niel, L., 2017. Risk factors associated with stranger-directed aggression in domestic dogs. Applied Animal Behaviour Science 197, 45-54.

Foyer, P., Svedberg, A.M., Nilsson, E., Wilsson, E., Faresjo, A., Jensen, P., 2016. Behavior and cortisol responses of dogs evaluated in a standardized temperament test for military working dogs. Journal of Veterinary Behavior-Clinical Applications and Research 11, 7-12.

Fratkin, J.L., Sinn, D.L., Patall, E.A., Gosling, S.D., 2013. Personality Consistency in Dogs: A Meta-Analysis. Plos One 8, 19.

Fratkin, J.L., Sinn, D.L., Thomas, S., Hilliard, S., Olson, Z., Gosling, S.D., 2015. Do you see what I see? Can non-experts with minimal training reproduce expert ratings in behavioral assessments of working dogs? Behavioural Processes 110, 105-116.

Fugazza, C., Mongillo, P., Marinelli, L., 2017. Sex differences in dogs' social learning of spatial information. Animal Cognition 20, 789-794.

Furton, K.G., Myers, L.J., 2001. The scientific foundation and efficacy of the use of canines as chemical detectors for explosives. Talanta 54, 487-500.

Gadbois, S., Reeve, C., 2016. The semiotic canine: scent processing dogs as research assistants in biomedical and environmental research. Dog Behavior 2, 26-32.

Gaines, S.A., Rooney, N.J., Bradshaw, J.W.S., 2008. The Effect of Feeding Enrichment upon Reported Working Ability and Behavior of Kenneled Working Dogs. J. Forensic Sci. 53, 1400-1404.

Ganitskaya, Y.V., Feoktistova, N.Y., Vasukov, D., Surov, A., 2020. Some behavioral features required for the selection of detection dogs. Biology Bulletin 47, 501-505.

Gazit, I., Goldblatt, A., Terkel, J., 2005. The role of context specificity in learning: the effects of training context on explosives detection in dogs. Animal Cognition 8, 143-150.

Gazit, I., Lavner, Y., Bloch, G., Azulai, O., Goldblatt, A., Terkel, J., 2003. A simple system for the remote detection and analysis of sniffing in explosives detection dogs. Behav. Res. Methods Instr. Comput. 35, 82-89.

Geary, N., Lovejoy, L., 2008. Sex differences in energy metabolism, obesity and eating behavior. Sex on the brain: from genes to behavior. New York: Oxford Publishers, 253-274.

Gerencsér, L., Bunford, N., Moesta, A., Miklósi, Á., 2018. Development and validation of the Canine Reward Responsiveness Scale–Examining individual differences in reward responsiveness of the domestic dog. Scientific reports 8, 1-14.

Goddard, M.E., Beilharz, R.G., 1982. GENETIC AND ENVIRONMENTAL-FACTORS AFFECTING THE SUITABILITY OF DOGS AS GUIDE DOGS FOR THE BLIND. Theoretical and Applied Genetics 62, 97-102.

References

Goddard, M.E., Beilharz, R.G., 1983. GENETICS OF TRAITS WHICH DETERMINE THE SUITABILITY OF DOGS AS GUIDE-DOGS FOR THE BLIND. Applied Animal Ethology 9, 299-315.

Goddard, M.E., Beilharz, R.G., 1984. A FACTOR-ANALYSIS OF FEARFULNESS IN POTENTIAL GUIDE DOGS. Applied Animal Behaviour Science 12, 253-265.

Goldsmith, H.H., Buss, A.H., Plomin, R., Rothbart, M.K., Thomas, A., Chess, S., Hinde, R.A., McCall, R.B., 1987. Roundtable: What is temperament? Four approaches. Child development, 505-529.

Gonder-Frederick, L., Rice, P., Warren, D., Vajda, K., Shepard, J., 2013. Diabetic alert dogs: a preliminary survey of current users. Diabetes Care 36, e47-e47.

Gonder-Frederick, L.A., Ducar, D., Grabman, J.H., Shepard, J., 2014. Diabetes Alert Dogs: A Review of the Industry. Diabetes 63, A223-A223.

Gonder-Frederick, L.A., Grabman, J.H., Shepard, J.A., Tripathi, A.V., Ducar, D.M., McElgunn, Z.R., 2017. Variability of diabetes alert dog accuracy in a real-world setting. Journal of diabetes science and technology 11, 714-719.

Goodloe, L.P., Borchelt, P.L., 1998. Companion dog temperament traits. J. Appl. Anim. Welf. Sci. 1, 303-338.

Gosling, S.D., 2008. Personality in non-human animals. Social and Personality Psychology Compass 2, 985-1001.

Grandjean, D., Al Marzooqi, D.H., Lecoq-Julien, C., Muzzin, Q., Al Hammadi, H.K., Alvergnat, G., Al Blooshi, K.M., khalifa Al Mazrouei, S., Alhmoudi, M.S., Al Ahbabi, F.M., 2021. Use Of Canine Olfactory Detection For COVID-19 Testing Study On UAE Trained Detection Dog Sensitivity. bioRxiv.

Grant, A.M., Nurmohamed, S., Ashford, S.J., Dekas, K., 2011. The performance implications of ambivalent initiative: The interplay of autonomous and controlled motivations. Organizational Behavior and Human Decision Processes 116, 241-251.

Gray, J.A., 1987. The psychology of fear and stress. CUP Archive.

Gruen, M.E., Foster, M.L., Lazarowski, L., Jeffries, A., Sherman, B.L., Dorman, D.C., 2019. Does the cognitive bias test in dogs depend on spatial learning? Journal of Veterinary Behavior 33, 1-6.

Guennif, S., 2002. From knowledge to individual action. Confidence, the hidden face of uncertainty. A rereading of the works of Knight and Keynes. Mind & Society 3, 13-28.

Guest, C., Pinder, M., Doggett, M., Squires, C., Affara, M., Kandeh, B., Dewhirst, S., Morant, S.V., D'Alessandro, U., Logan, J.G., 2019. Trained dogs identify people with malaria parasites by their odour. The Lancet Infectious Diseases 19, 578-580.

Hale, H.E., 2021. Clinical Trial: Investigating the efficacy of behavioural treatment and indicators of emotional state in the fearful domestic dog (Canis familiaris), University of Bristol.

Hall, N.J., 2017. Persistence and resistance to extinction in the domestic dog: Basic research and applications to canine training. Behavioural Processes 141, 67-74.

Hardin, D.S., Anderson, W., Cattet, J., 2015. Dogs can be successfully trained to alert to hypoglycemia samples from patients with type 1 diabetes. Diabetes Therapy 6, 509-517.

Hardin, D.S., Cattet, J., Anderson, W., Skrivanek, Z., 2013. Can diabetes alert dogs truly detect hypoglycemia?, Diabetes, AMER DIABETES ASSOC 1701 N BEAUREGARD ST, ALEXANDRIA, VA 22311-1717 USA, pp. A104-A104.

Hardin, D.S., Hillman, D., Cattet, J., 2012. Hypoglycemia Alert Dogs-Innovative Assistance for People With Type 1 Diabetes. Diabetes 61, A99-A100.

Harding, E.J., Paul, E.S., Mendl, M., 2004. Cognitive bias and affective state. Nature 427, 312-312.

References

Hare, B., Call, J., Tomasello, M., 1998. Communication of food location between human and dog (Canis familiaris). Evolution of communication 2, 137-159.

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

Hare, E., Kelsey, K.M., Serpell, J.A., Otto, C.M., 2018. Behavior Differences Between Search-and-Rescue and Pet Dogs. Front. Vet. Sci. 5, 5.

Hargrave, C., 2015. Anxiety, fear, frustration and stress in cats and dogs—Implications for the welfare of companion animals and practice finances. Companion Animal 20, 136-141.

Harvey, N.D., Craigon, P.J., Blythe, S.A., England, G.C.W., Asher, L., 2017. An evidence-based decision assistance model for predicting training outcome in juvenile guide dogs. Plos One 12, 26.

Harvey, N.D., Craigon, P.J., Sommerville, R., McMillan, C., Green, M., England, G.C.W., Asher, L., 2016. Test-retest reliability and predictive validity of a juvenile guide dog behavior test. Journal of Veterinary Behavior-Clinical Applications and Research 11, 65-76.

Haverbeke, A., De Smet, A., Depiereux, E., Giffroy, J.M., Diederich, C., 2009. Assessing undesired aggression in military working dogs. Applied Animal Behaviour Science 117, 55-62.

Haverbeke, A., Laporte, B., Depiereux, E., Giffroy, J.M., Diederich, C., 2008. Training methods of military dog handlers and their effects on the team's performances. Applied Animal Behaviour Science 113, 110-122.

Haverbeke, A., Messaoudi, F., Depiereux, E., Stevens, M., Giffroy, J.M., Diederich, C., 2010a. Efficiency of working dogs undergoing a new Human Familiarization and Training Program. Journal of Veterinary Behavior-Clinical Applications and Research 5, 112-119.

Haverbeke, A., Rzepa, C., Depiereux, E., Deroo, J., Giffroy, J.M., Diederich, C., 2010b. Assessing efficiency of a Human Familiarisation and Training Programme on fearfulness and aggressiveness of military dogs. Applied Animal Behaviour Science 123, 143-149.

Hayes, A.F., Krippendorff, K., 2007. Answering the call for a standard reliability measure for coding data. Communication methods and measures 1, 77-89.

Hayes, J.E., McGreevy, P.D., Forbes, S.L., Laing, G., Stuetz, R.M., 2018. Critical review of dog detection and the influences of physiology, training, and analytical methodologies. Talanta 185, 499-512.

Heeger, D., Landy, M., 1997. Signal detection theory. Dept. Psych., Stanford Univ., Stanford, CA, Teaching Handout.

Helton, W.S., 2010. Does perceived trainability of dog (Canis lupus familiaris) breeds reflect differences in learning or differences in physical ability? Behavioural Processes 83, 315-323.

Hennessy, M.B., Voith, V.L., Mazzei, S.J., Buttram, J., Miller, D.D., Linden, F., 2001. Behavior and cortisol levels of dogs in a public animal shelter, and an exploration of the ability of these measures to predict problem behavior after adoption. Applied animal behaviour science 73, 217-233.

Henriksen, R., Höglund, A., Fogelholm, J., Abbey-Lee, R., Johnsson, M., Dingemanse, N.J., Wright, D., 2020. Intra-individual behavioural variability: A trait under genetic control. International journal of molecular sciences 21, 8069.

Hiby, E.F., Rooney, N.J., Bradshaw, J.W.S., 2004. Dog training methods: their use, effectiveness and interaction with behaviour and welfare. Animal Welfare 13, 63-69.

Hiby, E.F., Rooney, N.J., Bradshaw, J.W.S., 2006. Behavioural and physiological responses of dogs entering re-homing kennels. Physiology & Behavior 89, 385-391.

References

Hielm-Björkman, A., Niuro, H., 2021. Finska coronahundar i samhällets tjänst. i Norden, 24.

Hobbs, S.L., Law, T.H., Volk, H.A., Younis, C., Casey, R.A., Packer, R., 2020. Impact of canine epilepsy on judgement and attention biases. Scientific reports 10, 1-11.

Horvath, G., Järverud, G.a.K., Järverud, S., Horváth, I., 2008. Human ovarian carcinomas detected by specific odor. Integrative cancer therapies 7, 76-80.

Houpt, K.A., Smith, S.L., 1981. Taste preferences and their relation to obesity in dogs and cats. The Canadian Veterinary Journal 22, 77.

Hournmady, S., Peron, F., Grandjean, D., Clero, D., Bernard, B., Titeux, E., Desquilbet, L., Gilbert, C., 2016. Relationships between personality of human-dog dyads and performances in working tasks. Applied Animal Behaviour Science 177, 42-51.

Hoyenga, K.B., Hoyenga, K.T., 1982. Gender and energy balance: sex differences in adaptations for feast and famine. Physiology & Behavior 28, 545-563.

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-+.

Humphrey, E., Warner, L.H., 1934. Working dogs: An attempt to produce a strain of German Shepherds which combines working ability and beauty of conformation. Dogwise Pub.

Hurt, A., Smith, D.A., 2009. Conservation dogs. Canine ergonomics: the science of working dogs. CRC Press, Boca Raton, 175-194.

Isen, A.M., 2001. An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. Journal of consumer psychology 11, 75-85.

Jakovcevic, A., Elgier, A.M., Mustaca, A.E., Bentosela, M., 2013. Frustration behaviors in domestic dogs. Journal of applied animal welfare science 16, 19-34.

Jamieson, L.T.J., Baxter, G.S., Murray, P.J., 2017. Identifying suitable detection dogs. Applied Animal Behaviour Science 195, 1-7.

Jendrny, P., Schulz, C., Twele, F., Meller, S., von Köckritz-Blickwede, M., Osterhaus, A.D.M.E., Ebbers, J., Pilchová, V., Pink, I., Welte, T., 2020. Scent dog identification of samples from COVID-19 patients– a pilot study. BMC infectious diseases 20, 1-7.

Jenkins, E.K., Lee-Fowler, T.M., Angle, T.C., Behrend, E.N., Moore, G.E., 2016. Effects of oral administration of metronidazole and doxycycline on olfactory capabilities of explosives detection dogs. American Journal of Veterinary Research 77, 906-912.

Jezierski, T., Adamkiewicz, E., Walczak, M., Sobczynska, M., Gorecka-Bruzda, A., Ensminger, J., Papet, E., 2014. Efficacy of drug detection by fully-trained police dogs varies by breed, training level, type of drug and search environment. Forensic Sci.Int. 237, 112-118.

Jezierski, T., Walczak, M., Ligor, T., Rudnicka, J., Buszewski, B., 2015. Study of the art: canine olfaction used for cancer detection on the basis of breath odour. Perspectives and limitations. Journal of Breath Research 9.

Johnen, D., Heuwieser, W., Fischer-Tenhagen, C., 2017. An approach to identify bias in scent detection dog testing. Applied Animal Behaviour Science 189, 1-12.

Jones, A.C., Gosling, S.D., 2005. Temperament and personality in dogs (Canis familiaris): A review and evaluation of past research. Applied Animal Behaviour Science 95, 1-53.

Jones, K.E., Dashfield, K., Downend, A.B., Otto, C.M., 2004. Search-and-rescue dogs: an overview for veterinarians. Javma-Journal of the American Veterinary Medical Association 225, 854-860.

References

Jones, R.T., Guest, C., Lindsay, S.W., Kleinschmidt, I., Bradley, J., Dewhirst, S., Last, A., Logan, J.G., 2020. Could bio-detection dogs be used to limit the spread of COVID-19 by travellers? Journal of Travel Medicine 27, taaa131.

Kaiser, H.F., 1974. An index of factorial simplicity. psychometrika 39, 31-36.

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 (Reconcile™) and a behaviour modification plan. BMC veterinary research 11, 1-10.

King, T., Hemsworth, P.H., Coleman, G.J., 2003. Fear of novel and startling stimuli in domestic dogs. Applied Animal Behaviour Science 82, 45-64.

King, T., Marston, L.C., Bennett, P.C., 2012. Breeding dogs for beauty and behaviour: Why scientists need to do more to develop valid and reliable behaviour assessments for dogs kept as companions. Applied Animal Behaviour Science 137, 1-12.

Kirton, A., Winter, A., Wirrell, E., Snead, O., 2008. Seizure response dogs: evaluation of a formal training program. Epilepsy & Behavior 13, 499-504.

Kirton, A., Wirrell, E., Zhang, J., Hamiwka, L., 2004. Seizure-alerting and-response behaviors in dogs living with epileptic children. Neurology 62, 2303-2305.

Kis, A., Ciobica, A., Topal, J., 2017. The effect of oxytocin on human-directed social behaviour in dogs (Canis familiaris). Horm. Behav. 94, 40-52.

Kis, A., Hernadi, A., Kanizsar, O., Gacsi, M., Topal, J., 2015. Oxytocin induces positive expectations about ambivalent stimuli (cognitive bias) in dogs. Horm. Behav. 69, 1-7.

Knol, B.W., Roozendaal, C., Vandenbogaard, L., Bouw, J., 1988. THE SUITABILITY OF DOGS AS GUIDE DOGS FOR THE BLIND - CRITERIA AND TESTING PROCEDURES. Vet. Q. 10, 198-204.

Knol, B.W., Roozendaal, C., Vandenbogaard, L., Bouw, J., 1989. THE SUITABILITY OF DOGS AS GUIDE DOGS FOR THE BLIND - CRITERIA AND TESTING PROCEDURES. Tijdschr. Diergeneeskd. 114, 839-845.

Knowles, T.G., 2020.

Konno, A., Inoue-Murayama, M., Yabuta, S., Tonoike, A., Nagasawa, M., Mogi, K., Kikusui, T., 2018. Effect of Canine Oxytocin Receptor Gene Polymorphism on the Successful Training of Drug Detection Dogs. Journal of Heredity 109, 566-572.

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.

Kovacs, K., Kis, A., Pogany, A., Koller, D., Topal, J., 2016. Differential effects of oxytocin on social sensitivity in two distinct breeds of dogs (Canis familiaris). Psychoneuroendocrinology 74, 212-220.

Krishnan-Sarin, S., Reynolds, B., Duhig, A.M., Smith, A., Liss, T., McFetridge, A., Cavallo, D.A., Carroll, K.M., Potenza, M.N., 2007. Behavioral impulsivity predicts treatment outcome in a smoking cessation program for adolescent smokers. Drug and alcohol dependence 88, 79-82.

Kubinyi, E., Turcsan, B., Miklosi, A., 2009a. Dog and owner demographic characteristics and dog personality trait associations. Behavioural Processes 81, 392-401.

Kubinyi, E., Turcsán, B., Miklósi, Á., 2009b. Dog and owner demographic characteristics and dog personality trait associations. Behavioural processes 81, 392-401.

Lalkhen, A.G., McCluskey, A., 2008. Clinical tests: sensitivity and specificity. Continuing education in anaesthesia critical care & pain 8, 221-223.

References

Lazarowski, L., Foster, M.L., Gruen, M.E., Sherman, B.L., Fish, R.E., Milgram, N.W., Dorman, D.C., 2015. Olfactory discrimination and generalization of ammonium nitrate and structurally related odorants in Labrador retrievers. Animal Cognition 18, 1255-1265.

Lazarowski, L., Haney, P.S., Brock, J., Fischer, T., Rogers, B., Angle, C., Katz, J.S., Waggoner, L.P., 2018. Investigation of the behavioral characteristics of dogs purpose-bred and prepared to perform Vapor Wake® detection of person-borne explosives. Front. Vet. Sci. 5, 50.

Lazarowski, L., Krichbaum, S., DeGreeff, L.E., Simon, A., Singletary, M., Angle, C., Waggoner, L.P., 2020a. Methodological considerations in canine olfactory detection research. Front. Vet. Sci. 7, 408.

Lazarowski, L., Krichbaum, S., Waggoner, L.P., Katz, J.S., 2020b. The development of problem-solving abilities in a population of candidate detection dogs (Canis familiaris). Animal Cognition 23, 755-768.

Lazarowski, L., Rogers, B., Waggoner, L.P., Katz, J.S., 2019a. When the nose knows: ontogenetic changes in detection dogs'(Canis familiaris) responsiveness to social and olfactory cues. Animal Behaviour 153, 61-68.

Lazarowski, L., Strassberg, L.R., Waggoner, L.P., Katz, J.S., 2019b. Persistence and human-directed behavior in detection dogs: ontogenetic development and relationships to working dog success. Applied Animal Behaviour Science 220, 104860.

Lazarowski, L., Waggoner, L.P., Krichbaum, S., Singletary, M., Haney, P.S., Rogers, B., Angle, C., 2020c. Selecting dogs for explosives detection: behavioral characteristics. Front. Vet. Sci., 597.

Lefebvre, D., Diederich, C., Delcourt, M., Giffroy, J.M., 2007. The quality of the relation between handler and military dogs influences efficiency and welfare of dogs. Applied Animal Behaviour Science 104, 49-60.

Lesniak, A., Walczak, M., Jezierski, T., Sacharczuk, M., Gawkowski, M., Jaszczak, K., 2008. Canine olfactory receptor gene polymorphism and its relation to odor detection performance by sniffer dogs. Journal of Heredity 99, 518-527.

Lit, L., Schweitzer, J.B., Oberbauer, A.M., 2011. Handler beliefs affect scent detection dog outcomes. Animal Cognition 14, 387-394.

Los, E.A., Ramsey, K.L., Guttmann-Bauman, I., Ahmann, A.J., 2017. Reliability of trained dogs to alert to hypoglycemia in patients with type 1 diabetes. Journal of diabetes science and technology 11, 506-512.

Lund, J.D., Jørgensen, M.C., 1999. Behaviour patterns and time course of activity in dogs with separation problems. Applied Animal Behaviour Science 63, 219-236.

Lynn, S.K., Barrett, L.F., 2014. "Utilizing" signal detection theory. Psychological science 25, 1663-1673.

Maa, E., Arnold, J., Ninedorf, K., Olsen, H., 2021. Canine detection of volatile organic compounds unique to human epileptic seizure. Epilepsy & Behavior 115, 107690.

MacLean, E.L., Hare, B., 2018. Enhanced Selection of Assistance and Explosive Detection Dogs Using Cognitive Measures. Front. Vet. Sci. 5, 14.

MacLean, E.L., Herrmann, E., Suchindran, S., Hare, B., 2017. Individual differences in cooperative communicative skills are more similar between dogs and humans than chimpanzees. Animal Behaviour 126, 41-51.

Maejima, M., Inoue-Murayama, M., Tonosaki, K., Matsuura, N., Kato, S., Saito, Y., Weiss, A., Murayama, Y., Ito, S., 2007. Traits and genotypes may predict the successful training of drug detection dogs. Applied Animal Behaviour Science 107, 287-298.

References

Mahoney, A.M., 2012. Mine detection rats: effects of repeated extinction on detection rates. Western Michigan University.

Maier, S.F., Seligman, M.E., 1976. Learned helplessness: theory and evidence. Journal of experimental psychology: general 105, 3.

Martin, C.S., Earleywine, M., Blackson, T.C., Vanyukov, M.M., Moss, H.B., Tarter, R.E., 1994. Aggressivity, inattention, hyperactivity, and impulsivity in boys at high and low risk for substance abuse. Journal of Abnormal Child Psychology 22, 177-203.

Masini, P., Zampetti, S., Cimici, C.A., Bianchi, L., Stingeni, L., Llera, G.M., Biancolini, F., 2017. Bed bug dermatitis: detection dog as a useful survey tool for environmental research of Cimex lectularius. Int. J. Dermatol. 56, E204-E206.

Maurer, M., McCulloch, M., Willey, A.M., Hirsch, W., Dewey, D., 2016. Detection of bacteriuria by canine olfaction, Open Forum Infectious Diseases, Oxford University Press.

McAulay, V., Deary, I.J., Frier, B.M., 2001. Symptoms of hypoglycaemia in people with diabetes. Diabetic Med. 18, 690-705.

McConnell, I., Marker, L., Rooney, N., 2022. Preliminary investigation into personality and effectiveness of livestock guarding dogs in Namibia. Journal of Veterinary Behavior-Clinical Applications and Research 48, 11-19.

McCrae, R.R., Costa Jr, P.T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M.D., Sanz, J., Sanchez-Bernardos, M.L., Kusdil, M.E., Woodfield, R., 2000. Nature over nurture: temperament, personality, and life span development. Journal of personality and social psychology 78, 173.

McCulloch, M., Jezierski, T., Broffman, M., Hubbard, A., Turner, K., Janecki, T., 2006. Diagnostic accuracy of canine scent detection in early-and late-stage lung and breast cancers. Integrative cancer therapies 5, 30-39.

McGarrity, M.E., Sinn, D.L., Thomas, S.G., Marti, C.N., Gosling, S.D., 2016. Comparing the predictive validity of behavioral codings and behavioral ratings in a working-dog breeding program. Applied Animal Behaviour Science 179, 82-94.

McGreevy, P.D., Georgevsky, D., Carrasco, J., Valenzuela, M., Duffy, D.L., Serpell, J.A., 2013. Dog Behavior Co-Varies with Height, Bodyweight and Skull Shape. Plos One 8, 7.

McMillan, F.D.S., J. A. Duffy, D. L. Masaoud, E. Dohoo, I. R., 2013. Differences in behavioval characteristics between dogs obtained as puppies from pet stores and those obtained from noncommercial breeders. Javma-Journal of the American Veterinary Medical Association 242, 1359-1363.

McNicol, D., 2005. A primer of signal detection theory. Psychology Press.

McPeake, K.J., Collins, L.M., Zulch, H., Mills, D.S., 2019. The canine frustration questionnaire—development of a new psychometric tool for measuring frustration in domestic dogs (Canis familiaris). Front. Vet. Sci., 152.

Meagher, R.K., 2009. Observer ratings: Validity and value as a tool for animal welfare research. Applied Animal Behaviour Science 119, 1-14.

Mehrkam, L.R., Wynne, G.D.L., 2014. Behavioral differences among breeds of domestic dogs (Canis lupus familiaris): Current status of the science. Applied Animal Behaviour Science 155, 12-27.

Mendel, J., Frank, K., Edlin, L., Hall, K., Webb, D., Mills, J., Holness, H.K., Furton, K.G., Mills, D., 2021. Preliminary accuracy of COVID-19 odor detection by canines and HS-SPME-GC-MS using exhaled breath samples. Forensic Science International: Synergy 3, 100155.

References

Mendl, M., Brooks, J., Basse, C., Burman, O., Paul, E., Blackwell, E., Casey, R., 2010a. Dogs showing separation-related behaviour exhibit a 'pessimistic' cognitive bias. Current Biology 20, R839-R840.

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, 161-181.

Mendl, M., Burman, O.H.P., Paul, E.S., 2010b. An integrative and functional framework for the study of animal emotion and mood. Proceedings of the Royal Society B-Biological Sciences 277, 2895-2904.

Meyer, I., Ladewig, J., 2008. The relationship between number of training sessions per week and learning in dogs. Applied animal behaviour science 111, 311-320.

Miklósi, Á., 2014. Dog behaviour, evolution, and cognition. oUp Oxford.

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.

Miklösi, Á., Polgárdi, R., Topál, J., Csányi, V., 1998. Use of experimenter-given cues in dogs. Animal cognition 1, 113-121.

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

Mills, D.S., Demontigny-Bédard, I., Gruen, M., Klinck, M.P., McPeake, K.J., Barcelos, A.M., Hewison, L., Van Haevermaet, H., Denenberg, S., Hauser, H., 2020. Pain and problem behavior in cats and dogs. Animals 10, 318.

Mirko, E., Kubinyi, E., Gacsi, M., Miklosi, A., 2012. Preliminary analysis of an adjective-based dog personality questionnaire developed to measure some aspects of personality in the domestic dog (Canis familiaris). Applied Animal Behaviour Science 138, 88-98.

Moeller, F.G., Barratt, E.S., Dougherty, D.M., Schmitz, J.M., Swann, A.C., 2001. Psychiatric aspects of impulsivity. American journal of psychiatry 158, 1783-1793.

Mogg, K., Bradley, B.P., 1998. A cognitive-motivational analysis of anxiety. Behaviour research and therapy 36, 809-848.

Moser, E., McCulloch, M., 2010. Canine scent detection of human cancers: A review of methods and accuracy. Journal of Veterinary Behavior-Clinical Applications and Research 5, 145-152.

Moulton, D.G., Ashton, E.H., Eayrs, J.T., 1960. Studies in olfactory acuity. 4. Relative detectability of n-aliphatic acids by the dog. Animal Behaviour 8, 117-128.

Muller, C.A., Riemer, S., Rosam, C.M., Schosswender, J., Range, F., Huber, L., 2012. Brief owner absence does not induce negative judgement bias in pet dogs. Animal Cognition 15, 1031-1035.

Müller, C.A., Riemer, S., Virányi, Z., Huber, L., Range, F., 2016. Inhibitory control, but not prolonged object-related experience appears to affect physical problem-solving performance of pet dogs. PloS one 11, e0147753.

Naranjo, C., Del Reguero, L., Moratalla, G., Hercberg, M., Valenzuela, M., Failde, I., 2019. Anxiety, depression and sleep disorders in patients with diabetic neuropathic pain: a systematic review. Expert review of neurotherapeutics 19, 1201-1209.

Nesse, R.M., 2000. Is depression an adaptation? Archives of general psychiatry 57, 14-20.

Netto, W.J., Planta, D.J.U., 1997. Behavioural testing for aggression in the domestic dog. Applied Animal Behaviour Science 52, 243-263.

References

Nevin, J.A., 1969. Signal detection theory and operant behavior: A review of David M. Green and John A. Swets' Signal detection theory and psychophysics1. Journal of the Experimental Analysis of Behavior 12, 475.

O'Connor, M.B., O'Connor, C., Walsh, C.H., 2008. A dog's detection of low blood sugar: a case report. Irish J. Med. Sci. 177, 155-157.

Oesterhelweg, L., Krober, S., Rottmann, K., Willhoft, J., Braun, C., Thies, N., Puchel, K., Silkenath, J., Gehl, A., 2008. Cadaver dogs - A study on detection of contaminated carpet squares. Forensic Sci.Int. 174, 35-39.

Oldenburg, C., Schoon, A., Heitkonig, I.M.A., 2016. Wildlife detection dog training: A case study on achieving generalization between target odor variations while retaining specificity. Journal of Veterinary Behavior-Clinical Applications and Research 13, 34-38.

Ong, A.D., Bergeman, C.S., Bisconti, T.L., Wallace, K.A., 2006. Psychological resilience, positive emotions, and successful adaptation to stress in later life. Journal of personality and social psychology 91, 730.

Ortiz, R., Liporace, J., 2005. “Seizure-alert dogs”: Observations from an inpatient video/EEG unit. Epilepsy & Behavior 6, 620-622.

Overall, K.L., Dunham, A.E., Juarbe-Diaz, S.V., 2016. Phenotypic determination of noise reactivity in 3 breeds of working dogs: A cautionary tale of age, breed, behavioral assessment, and genetics. Journal of Veterinary Behavior-Clinical Applications and Research 16, 113-125.

Panksepp, J., 1998. The periconscious substrates of consciousness: Affective states and the evolutionary origins of the self. Journal of consciousness studies 5, 566-582.

Panksepp, J., 2004. Affective neuroscience: The foundations of human and animal emotions. Oxford university press.

Panksepp, J., 2005. Affective consciousness: Core emotional feelings in animals and humans. Consciousness and Cognition 14, 30-80.

Papini, M.R., Ludvigson, H.W., 1994. FRUSTRATION THEORY - AN ANALYSIS OF DISPOSITIONAL LEARNING AND MEMORY - ANSEL,A. American Journal of Psychology 107, 604-612.

Pastore, C., Pirrone, F., Balzarotti, F., Faustini, M., Pierantoni, L., Albertini, M., 2011. Evaluation of physiological and behavioral stress-dependent parameters in agility dogs. Journal of Veterinary Behavior-Clinical Applications and Research 6, 188-194.

Pastore, R., Scheirer, C., 1974. Signal detection theory: Considerations for general application. Psychological Bulletin 81, 945.

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, 469-491.

Peremans, K., Audenaert, K., Coopman, F., Hoybergs, Y., Slegers, G., van Bree, H., Verschooten, F., Dierckx, R., 2002. Functional brain imaging of serotonin-2A receptors in impulsive dogs: a pilot study. Vlaams Diergeneeskundig Tijdschrift 71, 340-347.

Perry, E., Gulson, N., Cross, T.W.L., Swanson, K.S., 2017. Physiological effects of stress related to helicopter travel in Federal Emergency Management Agency search-and-rescue canines. Journal of Nutritional Science 6, 5.

Petry, N.M., Wagner, J.A., Rash, C.J., Hood, K.K., 2015. Perceptions about professionally and non-professionally trained hypoglycemia detection dogs. Diabetes Res. Clin. Pract. 109, 389-396.

Pflaumer, S., 1992. Seizure-alert dogs. Dog World 77, 42-43.

References

Pfungst, O., 1911. Clever Hans:(the horse of Mr. Von Osten.) a contribution to experimental animal and human psychology. Holt, Rinehart and Winston.

Phelan, J.M., Barnett, J.L., 2002. Chemical sensing thresholds for mine detection dogs, in: Broach, J.T., Harmon, R.S., Dobeck, G.J. (Eds.), Detection and Remediation Technologies for Mines and Minelike Targets Vii, Pts 1 and 2, Spie-Int Soc Optical Engineering, Bellingham, pp. 532-543.

Pickel, D., Manucy, G.P., Walker, D.B., Hall, S.B., Walker, J.C., 2004. Evidence for canine olfactory detection of melanoma. Applied Animal Behaviour Science 89, 107-116.

Pinikahana, J., Dono, J., 2009. The lived experience of initial symptoms of and factors triggering epileptic seizures. Epilepsy & Behavior 15, 513-520.

Piotti, P., Szabó, D., Bognár, Z., 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 46, 537-553.

Podberscek, A.L., Serpell, J.A., 1996. The English Cocker Spaniel: Preliminary findings on aggressive behaviour. Applied Animal Behaviour Science 47, 75-89.

Polgar, Z., Blackwell, E.J., Rooney, N.J., 2019. Assessing the welfare of kennelled dogs-A review of animal-based measures. Applied Animal Behaviour Science 213, 1-13.

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

Porritt, F., Shapiro, M., Waggoner, P., Mitchell, E., Thomson, T., Nicklin, S., Kacelnik, A., 2015. Performance decline by search dogs in repetitive tasks, and mitigation strategies. Applied Animal Behaviour Science 166, 112-122.

Pronin, E., Lin, D.Y., Ross, L., 2002. The bias blind spot: Perceptions of bias in self versus others. Personality and Social Psychology Bulletin 28, 369-381.

Proyer, R.T., 2011. Being playful and smart? The relations of adult playfulness with psychometric and self-estimated intelligence and academic performance. Learning and Individual Differences 21, 463-467.

Proyer, R.T., Ruch, W., Müller, L., 2010. Sense of humor among the elderly. Zeitschrift für Gerontologie und Geriatrie 43, 19-24.

Qian, X.L., Yarnal, C., 2011. The role of playfulness in the leisure stress-coping process among emerging adults: An SEM analysis. Leisure/Loisir 35, 191-209.

Raffan, E., Dennis, R.J., O'Donovan, C.J., Becker, J.M., Scott, R.A., Smith, S.P., Withers, D.J., Wood, C.J., Conci, E., Clements, D.N., 2016. A deletion in the canine POMC gene is associated with weight and appetite in obesity-prone labrador retriever dogs. Cell metabolism 23, 893-900.

Rasmussen, H.N., Scheier, M.F., Greenhouse, J.B., 2009. Optimism and physical health: A meta-analytic review. Annals of behavioral medicine 37, 239-256.

Reale, D., Reader, S.M., Sol, D., McDougall, P.T., Dingemanse, N.J., 2007. Integrating animal temperament within ecology and evolution. Biological Reviews 82, 291-318.

Reeve, C., Koivusalo, M., 2018. Biomedical scent detection dogs: would they pass as a health technology? Pet behaviour science, 1-7.

Riemer, S., Mills, D.S., Wright, H., 2014a. Impulsive for life? The nature of long-term impulsivity in domestic dogs. Animal Cognition 17, 815-819.

References

Riemer, S., Muller, C., Viranyi, Z., Huber, L., Range, F., 2014b. The Predictive Value of Early Behavioural Assessments in Pet Dogs - A Longitudinal Study from Neonates to Adults. Plos One 9, 13.

Riemer, S., Müller, C., Virányi, Z., Huber, L., Range, F., 2014c. The predictive value of early behavioural assessments in pet dogs–a longitudinal study from neonates to adults. PloS one 9, e101237.

Riezzo, I., Neri, M., Rendine, M., Bellifemina, A., Cantatore, S., Fiore, C., Turillazzi, E., 2014. Cadaver dogs: Unscientific myth or reliable biological devices? Forensic Sci.Int. 244, 213-221.

Rocznik, D., Sinn, D.L., Thomas, S., Gosling, S.D., 2015. Criterion Analysis and Content Validity for Standardized Behavioral Tests in a Detector-Dog Breeding Program. J. Forensic Sci. 60, S213-S221.

Rooney, N., Gaines, S., Hiby, E., 2009. A practitioner's guide to working dog welfare. Journal of Veterinary Behavior-Clinical Applications and Research 4, 127-134.

Rooney, N.J., Bradshaw, J.W.S., 2002. An experimental study of the effects of play upon the dog-human relationships. Applied Animal Behaviour Science 75, 161-176.

Rooney, N.J., Bradshaw, J.W.S., 2004. Breed and sex differences in the behavioural attributes of specialist search dogs - a questionnaire survey of trainers and handlers. Applied Animal Behaviour Science 86, 123-135.

Rooney, N.J., Bradshaw, J.W.S., Almey, H., 2004. Attributes of specialist search dogs - A questionnaire survey of UK dog handlers and trainers. J. Forensic Sci. 49, 300-306.

Rooney, N.J., Bradshaw, J.W.S., Gaines, S.A., 2003a. UK Detection Dog Rearing Project - an investigation into factors affecting search ability Dstl, UK.

Rooney, N.J., Bradshaw, J.W.S., Gaines, S.A., 2003b. UK Detection Dog Rearing Project - an investigation into factors affecting search

ability Dstl.

Rooney, N.J., Bradshaw, J.W.S., Robinson, I.H., 2000. A comparison of dog-dog and dog-human play behaviour. Applied Animal Behaviour Science 66, 235-248.

Rooney, N.J., Clark, C.C.A., 2021. Development of a Performance Monitoring Instrument for Rating Explosives Search Dog Performance. Front. Vet. Sci. 8, 15.

Rooney, N.J., Clark, C.C.A., Casey, R.A., 2016. Minimizing fear and anxiety in working dogs: A review. Journal of Veterinary Behavior-Clinical Applications and Research 16, 53-64.

Rooney, N.J., Gaines, S.A., Bradshaw, J.W.S., 2007a. Behavioural and glucocorticoid responses of dogs (Canis familiaris) to kennelling: Investigating mitigation of stress by prior habituation. Physiology & Behavior 92, 847-854.

Rooney, N.J., Gaines, S.A., Bradshaw, J.W.S., Penman, S., 2007b. Validation of a method for assessing the ability of trainee specialist search dogs. Applied Animal Behaviour Science 103, 90-104.

Rooney, N.J., Guest, C.M., Swanson, L.C.M., Morant, S.V., 2019. How effective are trained dogs at alerting their owners to changes in blood glycaemic levels?: Variations in performance of glycaemia alert dogs. Plos One 14, 16.

Rooney, N.J., Morant, S., Guest, C., 2013. Investigation into the Value of Trained Glycaemia Alert Dogs to Clients with Type I Diabetes. Plos One 8, 12.

Saetre, P., Strandberg, E., Sundgren, P.E., Pettersson, U., Jazin, E., Bergström, T.F., 2006. The genetic contribution to canine personality. Genes, Brain and Behavior 5, 240-248.

References

Sakr, R., Ghsoub, C., Rbeiz, C., Lattouf, V., Riachy, R., Haddad, C., Zoghbi, M., 2022. COVID-19 detection by dogs: from physiology to field application—a review article. Postgraduate Medical Journal 98, 212-218.

Scott, S.M., James, D., Ali, Z., 2006. Data analysis for electronic nose systems. Microchimica Acta 156, 183-207.

Seo, I.S., Lee, H.G., Koo, B., Koh, C.S., Park, H.Y., Im, C., Shin, H.C., 2018. Cross detection for odor of metabolic waste between breast and colorectal cancer using canine olfaction. Plos One 13.

Serpell, J.A., 1996. Evidence for an association between pet behavior and owner attachment levels. Applied Animal Behaviour Science 47, 49-60.

Serpell, J.A., Duffy, D.L., 2014. Dog breeds and their behavior, Domestic dog cognition and behavior, Springer, pp. 31-57.

Serpell, J.A., Hsu, Y.Y., 2001. Development and validation of a novel method for evaluating behavior and temperament in guide dogs. Applied Animal Behaviour Science 72, 347-364.

Serpell, J.A., Hsu, Y.Y., 2005. Effects of breed, sex, and neuter status on trainability in dogs. Anthrozoos 18, 196-207.

Sharpe, J.P., Martin, N.R., Roth, K.A., 2011. Optimism and the Big Five factors of personality: Beyond Neuroticism and Extraversion. Personality and Individual Differences 51, 946-951.

Sheppard, G., Mills, D.S., 2002. The development of a psychometric scale for the evaluation of the emotional predispositions of pet dogs. International journal of comparative psychology 15.

Shettleworth, S.J., 2009. Cognition, evolution, and behavior. Oxford university press.

Shiverdecker, M.D., Schiml, P.A., Hennessy, M.B., 2013. Human interaction moderates plasma cortisol and behavioral responses of dogs to shelter housing. Physiology & behavior 109, 75-79.

Sinn, D.L., Gosling, S.D., Hilliard, S., 2010. Personality and performance in military working dogs: Reliability and predictive validity of behavioral tests. Applied Animal Behaviour Science 127, 51-65.

Slabbert, J.M., Odendaal, J.S.J., 1999. Early prediction of adult police dog efficiency - a longitudinal study. Applied Animal Behaviour Science 64, 269-288.

Sonoda, H., Kohnoe, S., Yamazato, T., Satoh, Y., Morizono, G., Shikata, K., Morita, M., Watanabe, A., Kakeji, Y., Inoue, F., Maehara, Y., 2011. Colorectal cancer screening with odour material by canine scent detection. Gut 60, 814-819.

Soproni, K.M., A. Topal, J. Csanyi, V., 2001. Comprehension of human communicative signs in pet dogs (Canis familiaris). Journal of Comparative Psychology 115, 122-126.

Stankov, L., Lee, J., Paek, I., 2009. Realism of confidence judgments. European Journal of Psychological Assessment 25, 123.

Statistics, L., 2015a. Multiple regression using SPSS Statistics. Statistical tutorials and software guides. Retrieved from https://statistics.laerd.com/, p. Retrieved from https://statistics.laerd.com/.

Statistics, L., 2015b. Principal components analysis (PCA) using SPSS Statistics. Statistical tutorials and software guides. Retrieved from https://statistics.laerd.com/

Statistics, L., 2017. Chi-square test of homogeneity (2 x C) using SPSS Statistics. Statistical tutorials and software guides. Retrieved from https://statistics.laerd.com/

References

Strong, V., Brown, S., Huyton, M., Coyle, H., 2002. Effect of trained Seizure Alert Dogs® on frequency of tonic–clonic seizures. Seizure 11, 402-405.

Strong, V., Brown, S.W., 2000. Should people with epilepsy have untrained dogs as pets? Seizure 9, 427-430.

Strong, V., Brown, S.W., Walker, R., 1999a. Seizure-alert dogs—fact or fiction? Seizure 8, 62-65.

Strong, V., Brown, S.W., Walker, R., 1999b. Seizure-alert dogs - fact or fiction? Seizure 8, 62-65.

Stur, I., 1987. Genetic aspects of temperament and behaviour in dogs. J. Small Anim. Pract. 28, 957-964.

Sumegi, Z., Kis, A., Miklosi, A., Topal, J., 2014. Why Do Adult Dogs (Canis familiaris) Commit the A-not-B Search Error? Journal of Comparative Psychology 128, 21-30.

Svartberg, K., 2002. Shyness-boldness predicts performance in working dogs. Applied Animal Behaviour Science 79, 157-174.

Svartberg, K., Forkman, B., 2002. Personality traits in the domestic dog (Canis familiaris). Applied Animal Behaviour Science 79, 133-155.

Svartberg, K., Tapper, I., Temrin, H., Radesater, T., Thorman, S., 2005. Consistency of personality traits in dogs. Animal Behaviour 69, 283-291.

Swift, A., Heale, R., Twycross, A., 2020. What are sensitivity and specificity? Evidence-Based Nursing 23, 2-4.

Tauveron, I., Delcourt, I., Desbiez, F., Somda, F., Thieblot, P., 2006. Canine detection of hypoglycaemic episodes whilst driving. Diabetic Med. 23, 335-335.

Taylor, K.D., Mills, D.S., 2006. The development and assessment of temperament tests for adult companion dogs. Journal of Veterinary Behavior-Clinical Applications and Research 1, 94-108.

Tiira, K., Tikkanen, A., Vainio, O., 2020. Inhibitory control–Important trait for explosive detection performance in police dogs? Applied Animal Behaviour Science 224, 104942.

Titulaer, M., Blackwell, E.J., Mendl, M., Casey, R.A., 2013. Cross sectional study comparing behavioural, cognitive and physiological indicators of welfare between short and long term kennelled domestic dogs. Applied Animal Behaviour Science 147, 149-158.

Tomkins, L.M., Thomson, P.C., McGreevy, P.D., 2011. Behavioral and physiological predictors of guide dog success. Journal of Veterinary Behavior-Clinical Applications and Research 6, 178-187.

Topal, J., Miklosi, A., Csanyi, V., 1997. Dog-human relationship affects problem solving behavior in the dog. Anthrozoos 10, 214-224.

Uccheddu, S., Mariti, C., Sannen, A., Vervaecke, H., Arnout, H., Gutienrrez Rufo, J., Gazzano, A., Haverbeke, A., 2018. Behavioural and cortisol responses of shelter dogs to a cognitive bias test after olfactory enrichment with essential oils. Dog Behaviour 4.

UK, A.d., 2022a. Assistance dogs UK. Retrieved from https://www.assistancedogs.org.uk/.

UK, S.d., 2022b. Support dogs UK. Retrieved from https://www.supportdogs.org.uk/.

Van den Berg, L., Schilder, M., Knol, B., 2003. Behavior genetics of canine aggression: behavioral phenotyping of golden retrievers by means of an aggression test. Behavior genetics 33, 469-483.

Van den Bergh, F., Spronk, M., Ferreira, L., Bloemarts, E., Groenink, L., Olivier, B., Oosting, R., 2006. Relationship of delay aversion and response inhibition to extinction learning, aggression, and sexual behaviour. Behavioural brain research 175, 75-81.

References

van der Borg, J.A.M., Graat, E.A.M., Beerda, B., 2017. Behavioural testing based breeding policy reduces the prevalence of fear and aggression related behaviour in Rottweilers. Applied Animal Behaviour Science 195, 80-86.

Vas, J., Topal, J., Pech, E., Miklosi, A., 2007. Measuring attention deficit and activity in dogs: A new application and validation of a human ADHD questionnaire. Applied Animal Behaviour Science 103, 105-117.

Walczak, M., Jezierski, T., Gorecka-Bruzda, A., Sobczynska, M., Ensminger, J., 2012. Impact of individual training parameters and manner of taking breath odor samples on the reliability of canines as cancer screeners. Journal of Veterinary Behavior-Clinical Applications and Research 7, 283-294.

Walker, D.B., Walker, J.C., Cavnar, P.J., Taylor, J.L., Pickel, D.H., Hall, S.B., Suarez, J.C., 2006. Naturalistic quantification of canine olfactory sensitivity. Applied Animal Behaviour Science 97, 241-254.

Waller, B.M., Micheletta, J., 2013. Facial expression in nonhuman animals. Emotion Review 5, 54-59.

Wallis, L.J., Szabó, D., Kubinyi, E., 2020. Cross-sectional age differences in canine personality traits; influence of breed, sex, previous trauma, and dog obedience tasks. Front. Vet. Sci. 6, 493.

Wasser, S.K., Davenport, B., Ramage, E.R., Hunt, K.E., Parker, M., Clarke, C., Stenhouse, G., 2004. Scat detection dogs in wildlife research and management: application to grizzly and black bears in the Yellowhead Ecosystem, Alberta, Canada. Can. J. Zool.-Rev. Can. Zool. 82, 475-492.

Wathan, J., McComb, K., 2014. The eyes and ears are visual indicators of attention in domestic horses. Current Biology 24, R677-R679.

Wells, D.L., Hepper, P.G., Milligan, A.D., Barnard, S., 2017a. Cognitive bias and paw preference in the domestic dog (Canis familiaris). Journal of Comparative Psychology 131, 317.

Wells, D.L., Hepper, P.G., Milligan, A.D.S., Barnard, S., 2017b. Cognitive Bias and Paw Preference in the Domestic Dog (Canis familiaris). Journal of Comparative Psychology 131, 317-325.

Wells, D.L., Lawson, S.W., Siriwardena, A.N., 2008. Canine Responses to Hypoglycemia in Patients with Type 1 Diabetes. J. Altern. Complement Med. 14, 1235-1241.

Williams, H., Pembroke, A., 1989. SNIFFER DOGS IN THE MELANOMA CLINIC. Lancet 1, 734-734.

Williams, M., Johnston, J.M., 2002. Training and maintaining the performance of dogs (Canis familiaris) on an increasing number of odor discriminations in a controlled setting. Applied Animal Behaviour Science 78, 55-65.

Willis, C.M., Church, S.M., Guest, C.M., Cook, W.A., McCarthy, N., Bransbury, A.J., Church, M.R.T., Church, J.C.T., 2004. Olfactory detection of human bladder cancer by dogs: proof of principle study. Br. Med. J. 329, 712-714A.

Wilson, C., Morant, S., Kane, S., Pesterfield, C., Guest, C., Rooney, N.J., 2019. An Owner-Independent Investigation of Diabetes Alert Dog Performance. Front. Vet. Sci. 6, 11.

Wilson, C.H., Kane, S.A., Morant, S.V., Guest, C.M., Rooney, N.J., 2020. Diabetes alert dogs: Objective behaviours shown during periods of owner glucose fluctuation and stability. Applied Animal Behaviour Science 223, 10.

Wilsson, E., Sinn, D.L., 2012. Are there differences between behavioral measurement methods? A comparison of the predictive validity of two ratings methods in a working dog program. Applied Animal Behaviour Science 141, 158-172.

References

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

Wilsson, E., Sundgren, P.E., 1997b. 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.

Wiseman-Orr, M., Nolan, A., Reid, J., Scott, E., 2001. Preliminary study on owner-reported behaviour changes associated with chronic pain in dogs. Veterinary Record 149, 423-424.

Wright, H.F., Mills, D.S., Pollux, P.M., 2011. Development and validation of a psychometric tool forassessing impulsivity in the domestic dog (Canis familiaris). International Journal of Comparative Psychology 24.

Wright, H.F., Mills, D.S., Pollux, P.M.J., 2012. Behavioural and physiological correlates of impulsivity in the domestic dog (Canis familiaris). Physiology & Behavior 105, 676-682.