A METHODOLOGICAL EXPLORATION OF SCENT PROCESSING IN VITRO IN HYPOGLYCEMIA ALERT DOGS

by

Catherine Reeve

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2017

Dedication

This thesis is dedicated to my cousin, the late James Walker Reeve and his father, my uncle, William Walker Reeve.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The olfactory acuity of dogs has resulted in them being trained for a wide range of applied tasks, including the detection and alert of medical conditions and states. Anecdotal reports of dogs signaling hypoglycemia in people with type 1 diabetes suggest that dogs can detect odour cues that signal physiological changes. I present a series of studies examining the training and testing of hypoglycemia detection and alert dogs. I first present a training program for training dogs with no previous detection training to detect human breath samples. The efficacy of the training program was then evaluated by testing four dogs' ability to 1) discriminate between breath samples from three different people and, 2) discriminate between breath samples from one person donated at three different times of the day. The results showed that all four dogs could discriminate the breath samples with a high degree of accuracy. Next, I present and test a method for maintaining breath sample integrity over time. Breath samples were prepared using silicone-coated cotton and uncoated cotton and four dogs' detectability of the samples were tested over time. The results showed that silicone-coated cotton balls did not improve dogs' detectability two hours after breath sample preparation, but greatly improved two dogs' detectability of breath samples stored over four weeks. Finally, I tested four dogs' ability to discriminate between breath samples donated by people with type 1 diabetes when their blood sugar was low, normal, and high. I then tested two dogs' ability to generalize the odour of hypoglycemia across multiple breath samples from one individual, and tested the ability of one dog to generalize the odour of hypoglycemia across breath samples from different individuals. The results showed that all four dogs could discriminate between the breath samples from different glycemic states (accuracy 93.3%-100%). One dog (of two tested) generalized the odour of hypoglycemia across multiple breath samples from the same person (Specificity 89%; Sensitivity 62%). More research is needed to determine whether the same dog could generalize the odour of hypoglycemia across breath samples from different people.

LIST OF ABBREVIATIONS USED

3AFC: Three alternative forced choice

CGM: Continuous glucose monitor

CSII: Continuous subcutaneous insulin infusion

DADs: Diabetic Alert Dogs

GC-MS: Gas chromatography – mass spectrometry

MTS: Matching-to-sample

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Simon Gadbois, for his invaluable instruction, guidance, patience, and continuous encouragement throughout my degree. I thank him not only for his supervision, but also for offering me innumerable opportunities to further my professional experience. Dr. Gadbois' passion and enthusiasm for animal behaviour is apparent in his love for field work, his teaching, and his supervision. I am thankful to have had a mentor that loves their profession as much as he does.

I would also like to thank my thesis committee members, Vincent Lolordo, Raymond Klein, and Elizabeth Cummings for their scientific contributions as well as their academic guidance. A further thank you to Raymond Klein and his wife, Marilyn Klein, for giving me the opportunity to spend time with one of my most favourite dogs, Marcus.

Furthermore, I would like to thank Dr. Peter Wentzell and Bjorn Wielens for their knowledgeable scientific contribution, and for their time and patience in helping me understand the chemistry of breath collection.

A special thank you to the lovely Karen Scott for keeping the project organized and for her consistent congeniality.

I would like to express a most sincere thank you to the dog owners (Spencer, Heather, Vanessa, Daniel, Bonnie, Eileen, Monty, and Sam) who reliably brought their beautiful dogs to Dalhousie each week. The completion of my thesis would literally not have been possible without their time and commitment. And, of course, I have to thank Charlee, Nutella, Bella, Mist, and Koda for showing up to school each week full of excitement and enthusiasm. These dogs' love and commitment to their work was contagious and I feel lucky to have had them be such a large part of my PhD experience. A special thank you goes to Heather for the dinners, hikes, and horse rides, and for supporting me in the process of finding my own dog.

Importantly, all of my work would not have been possible without the numerous honours students and lab volunteers who helped me collect data each week. A special thank you to Sonia Smith who contributed a lot of her time to helping me with data collection.

I must express my deepest appreciation to my parents for providing me with the skills and confidence to pursue a career I love, for always encouraging me, and supporting me whole-heartedly when it was tough. And a special thank you to my sisters for their love and support.

Finally, I must acknowledge my dog, Walker, for being the best hiking buddy I could ask for and for being an endless source of unconditional love when I need it most.

CHAPTER 1: INTRODUCTION

1.0 Overview

Because of our close evolutionary history, humans and dogs share a profound interspecies relationship. Owning a dog has been shown to reduce stress, and improve physical health and mental well-being (O'Haire, 2010). Moreover, dogs have contributed positively not only to our personal lives, but to our professional lives as well. Dogs' incredible olfactory acuity (Walker et al., 2006) combined with their unique ability to communicate with humans (Kaminski & Nitzschner, 2013; Udell & Wynne, 2008) has resulted in their training to assist humans in a wide range of olfactory processing tasks. For example, dogs have been successfully trained in olfactory tasks such as search and rescue, conservation and agricultural applications, and forensic identification (for a review, see Helton, 2009).

Further evidence suggests that dogs may be more attuned to their human companions than previously thought. A 1989 report by Williams and Pembroke documented the case of a dog that persistently sniffed at a mole on the owner's leg, which, after clinical examination, was discovered to be malignant melanoma. Since then, scientists have begun to examine whether dogs can recognize physiological changes in their owners, with promising results. Studies have shown that dogs may alert to changes in their owners' physiological state, such as migraines (Marcus & Bhowmick, 2013), seizures (Strong, Brown, & Walker, 1999), and diabetic hypoglycemia (Hardin et al., 2015; Wells et al., 2008). Further empirical studies report dogs' successful detection of infections and infectious agents (Bomers et al. 2012; Koivusalo et al., 2017), and cancers

(Moser & McCulloch, 2010). But additional studies report dogs showing less success detecting the same conditions (Amundsen et la., 2014; Dehlinger et al., 2013; Elliker et al., 2014; Gordon et al., 2008; Willis et al., 2004). However, as a growing field, discrepancies in training procedures, protocols for biological sample collection and storage, and formal testing designs may be responsible for the inconsistencies in the literature.

One area where biomedical detection and alert dogs have potential for greatly improving their owners' quality of life is the detection of hypoglycemia in people with type 1 diabetes. Although technologies for blood glucose monitoring continue to improve, research suggests that people may choose not to utilize these technologies because they can be uncomfortable to wear, entail an increased time burden, and can be expensive (Acerini, 2016). Furthermore, reports on the effect of current technologies on the incidence of hypoglycemia are contradictory (for reviews see Jeitler et al., 2008; Riemsma et al., 2016). Dogs have been reported to regularly detect hypoglycemia, often before the owner is even aware of the hypoglycemic episode (Chen et al., 2000; Tauveron et al., 2006; Wells et al., 2008).

In the scientific literature, however, studies that train and test dogs for hypoglycemia detection are few and their findings are inconsistent (Dehlinger et al., 2013; Hardin et al., 2015). Therefore, this thesis will contribute to the literature by presenting studies pertaining to the training and testing of hypoglycemia detection dogs.

1.1 Diabetes and Treatment Techniques

Type 1 diabetes (insulin-dependent diabetes) is one of the most common chronic diseases in children. In a worldwide study spanning 57 countries, De Beaufort (2006)

found that between 1990 and 1999, 40,000 children were diagnosed with type 1 diabetes, and that the average increase in prevalence was 2.8% per year.

Hypoglycemia, or low blood sugar, is the most common acute complication of insulin-dependent diabetes, occurring on average twice a week (Frier, 2008). For individuals with type 1 diabetes, hypoglycemia is difficult to avoid due to imperfect insulin replacement precipitated by an imbalance between factors that lower blood sugar, like exercise and insulin, and those that raise it, like food (Trang et al., 2014). Mild episodes of hypoglycemia may result in cognitive impairments (Becker & Ryan, 2000; Gschwend, Ryan, Atchison, Arslanian, & Becker, 1995), but if they are detected early, they do not compromise an individual's ability to administer necessary treatment to themselves (Unger & Parkin, 2011). But recurrent mild hypoglycemic events can increase the likelihood of asymptomatic hypoglycemia (Ovalle et al., 1998) and impair subsequent hormonal counter regulation mechanisms (Davis, Shavers, Mosqueda-Garcia, & Costa, 1997), thereby increasing the chances of future hypoglycemic episodes. If left untreated, a hypoglycemic event can result in neuroglycopenic symptoms (e.g., dizziness, lethargy, slurred speech and mental confusion, Gonder-Frederick, Nyer, Shepard, Vajda, & Clarke, 2011), at which point an individual requires the help of someone else to administer treatment (ADA, 2005). Without treatment, these severe hypoglycemic episodes can result in a loss of consciousness, a seizure, coma, or death (Frier, 2004) and repeated episodes can lead to permanent cognitive impairments (Bade-White & Obrzut, 2009; Becker & Ryan, 2000; Brands, Biessels, de Haan, Kappelle, & Kessels, 2005; Hershey et al., 2005). Although severe hypoglycemic events have been documented at a rate of only 5.8 out of 100 patient years (O'Connell, Cooper, Bulsara, Davis, & Jones,

2011), the potential severity of untreated hypoglycemia often leads people to develop a fear of hypoglycemia that can impede their blood-glucose maintenance and negatively affect their psychological well-being (Gonder-Frederick et al., 2011; Wild et al., 2007). Therefore, timely and accurate detection of hypoglycemia is imperative to maintain long term health in individuals with type 1 diabetes. But hypoglycemia can be difficult to detect (Gonder-Frederick et al., 2008), especially during sleep when hypoglycemic symptoms are less pronounced (Ly et al., 2014). Therefore, accurate and timely detection of hypoglycemia is a constant concern for people with type 1 diabetes.

Some patients self-monitor their blood sugar levels three to four times a day using a home glucose meter, and manual injection of insulin when needed (Riemsma et al., 2016). Alternatively, individuals may use continuous glucose monitors (CGM) which, through the use of sensors inserted under the skin, provide continuous real time measurements of interstitial glucose throughout the day and night (Howsman & Bequette, 2015). CGMs can be combined with continuous subcutaneous insulin infusion pumps (CSII) to maximize glycemic control (Jeitler et al., 2008), and the most recent technologies even allow for the infusion of insulin to be suspended if needed (Riemsma et al., 2016). An appealing feature of CGMs is that they can be programmed to sound an alarm when their users' blood sugar deviates into hyperglycemic or hypoglycemic levels (Klonoff, 2005), decreasing the chances of an individual delaying treatment due to asymptomatic hypoglycemic episodes, or nocturnal hypoglycemic events.

Although the use of these technologies can result in improved glycemic control, the detection of hypoglycemia is not necessarily improved with newer technologies (Acerini et al., 2016; Jeitler et al., 2008; Riemsma et al., 2016). Furthermore, CGMs can

often sound false alarms for hypoglycemia (Zijlstra et al., 2013). False alarms can be particularly dangerous since they increase the likelihood that an individual will develop "alarm fatigue"; when an individual is exposed to an alarm so frequently that his/her likelihood of responding to the alarm decreases (Shivers, Mackowiak, Anhalt, & Zisser, 2013). Alarm fatigue is particularly common in the case of nocturnal hypoglycemia (Buckingham et al., 2005), often leading to severe hypoglycemic episodes (Buckingham et al., 2008; Matkya, 2002).

Furthermore, despite the potential benefits, the use of CGMs tends to decline over time (Weinzimer et al., 2009), especially in children. Decreased adherence to the use of CGM technologies is likely because they can be uncomfortable to wear especially in young children with limited skin surface area, and the use of subcutaneous systems increases the likelihood of infection at insertion and tape sites (Aye, Block, & Buckingham, 2010). Moreover, these technologies can be quite costly with long term use (Acerini et al., 2016). And finally, individuals utilizing CGMs and CSII may see their monitors as a constant reminder of their disease, which may contribute to social impairments and a feeling of social stigmatization (Schabert, Browne, Mosely, & Speight, 2013).

The observed limitations of CGMs, both technological and psychological, have prompted research examining more user-friendly hypoglycemia detection systems. Anecdotal evidence of people with type 1 diabetes exhibiting fruity odour on the breath and in the urine during hyperglycemia (Phillips, 1997) and advancements in technology have spurred an interest in the utility of exhaled breath analysis for hypoglycemia detection.

1.2 Volatile Organic Compounds

As mentioned above, ancient physicians documented disease-specific odours emanating from their patients, including the fruity odour of acetone on the breath of diabetic people experiencing hyperglycemia (Phillips, 1997). With modern technology, we know that what these ancient physicians were smelling were volatile organic compounds being emitted from different channels in the human body.

Volatile organic compounds (VOCs) are molecular compounds that have high vapor pressure and are therefore in the gaseous state at room temperature (Angle et al., 2016; Schmidt & Podmore, 2015). Metabolic processes in human body cells release VOCs into the bloodstream and through pulmonary circulation, these VOCs enter the lungs and are exhaled in the breath. VOCs are also found in the headspace (the air above a sample) of feces, urine, sweat, milk, blood, and saliva (de Lacy Costello et al., 2014). Furthermore, infection, disease, or changes in metabolism, such as in the case of diabetic hypoglycemia, change cellular processes and consequently change the ratio and type of VOCs produced (Shirasu & Touhara, 2011). Therefore, VOCs indicate the metabolic state of a person and can signal physiological changes within the body, therefore having diagnostic potential.

Importantly, the collection of biological samples for analysis of VOCs can be very simple and non-invasive. For example, the collection of a breath sample typically involves having a patient simply breathe through a breath collection instrument containing an adsorbent material (adsorption, not absorption, because the molecules in the breath adhere to the surface of the material rather than dissolving into it). And sweat samples are obtained by wiping a person's skin with an adsorbent material (Shirasu &

Touhara, 2011). The sample collection materials can then be stored until analysis. Not only are these sampling techniques non-invasive, they are straightforward, do not require much training, and multiple samples can be collected without discomfort to the patient.

Chemical analysis of VOCs in biological samples is done with a number of analytical techniques. Most commonly, researchers first use solid-phase microextraction (SPME) to extract the volatiles of interest. Following extraction, gas chromatography (GC) is used to separate the compounds based on their volatility, and then mass spectrometry (MS) is used to identify the compounds based on their mass. But multiple different techniques (not discussed here), are required to accurately identify the collection of VOCs emitted by the different channels of the human body (de Lacy Costello et al., 2014).

Using these technologies, researchers have discovered specific VOCs associated with many physiological conditions (for a review, see Shirasu & Touhara, 2011) including fluctuations of blood glucose in people with diabetes (Minh et al., 2012). Novak et al. (2007) found a correlation between increased plasma glucose levels (hyperglycemia) and exhaled methyl nitrate and Neupane et al. (2016) found that exhaled breath isoprene rose significantly during a hypoglycemic event. Galassetti et al. (2005) measured chemicals in exhaled breath as a model for estimating glucose levels and found a strong correlation between exhaled breath measures and standard glucose measures. Despite the promising results of these studies, clinical application of exhaled breath VOC analysis is limited because of technological drawbacks. The concentrations of most VOCs in biological samples are very low [e.g., parts per billion ($nmol^{-1}$) to parts per trillion ($pmol^{-1}$) in breath; Schmidt & Podmore, 2015], and since current technologies are

only able to detect VOCs in concentrations of parts per billion (nmol/mol), a complicated preconcentration process is often required (Schmidt & Podmore, 2015). Moreover, conditions may be marked by multiple different VOCs or simply by changes in their concentrations, making analysis even with very sensitive technologies complicated. Finally, these techniques can be expensive and extensive training is required to perform and analyze their results (Shirasu & Touhara, 2011). Alternatively, researchers have begun to examine the efficacy of a cheaper, more user-friendly VOC analysis system: domestic dogs.

1.3 Physiology of Dog olfaction

Through the process of natural selection, the dog's nose has evolved as a sensitive and efficient olfactory processing organ. Settles et al. (2003) examined the aerodynamics of sniffing using a combination of high-speed videography of thermal air currents, light scattering by airborne particles, and direct imaging of the dogs' nostrils in motion. Their work showed how air currents carrying odorant molecules flow towards the dogs' nose during inhalation. Settles et al. (2003) also documented the physiological details of a dog's nose during sniffing. They showed that when a dog exhales, a flap of skin just inside the dogs' nostrils, called the alar fold, covers a large portion of the nostril, and air is forced out ventrally and laterally through the midlateral slits of the nose. Then, during the inspiration phase of sniffing, the alar fold retracts to allow the incoming air to be directed to a hole in the upper part of the nostril they call the "upper orifice". Further work by Craven et al. (2010) showed that each nostril samples air from spatially distinct regions, resulting in sampling from a large surface area. Craven et al. (2010) also measured the airflow rate of dogs during the inhalation and exhalation phases of a

sniffing bout and found that dogs sniffed at an average frequency of 4 to 7 Hz, with each sniffing bout being characterized as exhibiting a crescendo and decrescendo (Craven et al., 2010).

Following the flow of odour-laden air through the dogs' nostrils, into the nasal cavity and towards the olfactory tissues, it is clear that the dogs' nose is evolved to process olfactory stimuli efficiently and effectively. Using magnetic resonance imaging, Craven et al. (2010) modeled the internal structure of a dog's nasal cavity and revealed two distinct pathways that inspired air may travel. During respiration, airflow is directed below a bony plate (the lamina transvera) towards the nasopharynx and then out of the nasal cavity. During olfaction, however, air is directed above the lamina transvera and is transported to the olfactory recess at the rear of the nasal cavity. Here, airflow turns 180 degrees and encounters scroll-like ethmoturbinates lined with olfactory epithelium (Craven et al., 2010). A dog's olfactory epithelium measures around 150 to 170 cm^2, as compared to humans' olfactory epithelium that is only 5 cm^2 (Miklosi, 2007). Within the dog's olfactory epithelium are specialized olfactory receptor cells that take in molecules and convert them to neural signals. A dog's olfactory epithelium contains between 220 million to 2 billion olfactory receptor cells (compared to human's 12 to 40 million, Miklosi, 2007).

All this specialized olfactory hardware translates into dogs being able to detect odours at parts per trillion (Walker et al., 2006). Furthermore, a study conducted by Waggoner et al. (1998) showed that dogs could detect target odours in the presence of distracting odours, even when the concentration of the distracting odours increased to 100 times the concentration of the target odour. Therefore, the dog's nose is well designed to

process complex olfactory stimuli; a trait that lends itself well to the detection of disease-specific odours.

1.4 Dogs for Biomedical Detection and Alert

The field of canine biomedical detection was sparked by an intriguing report by Williams and Pembroke (1989), in which they presented the first report in a scientific journal (the Lancet) hypothesizing that a dog could smell a human pathology. They documented the case of a dog persistently sniffing a mole on woman's leg, and even attempting to bite the area on one occasion. After histopathological examination, the mole was found to be malignant melanoma. A similar report was presented by Williams and Church (2001) who documented two more cases where dogs correctly identified cancers on their owners' skin.

Likewise, Chen et al. (2008) presented anecdotal reports of dogs detecting hypoglycemia in their owners. In one case, a woman with type 2 diabetes who reported having good awareness of her hypoglycemic symptoms, reported that her dog would consistently get up and hide under a chair prior to a hypoglycemic event in her owner. Incredibly, the dog would hide before the owner experienced any symptoms of hypoglycemia. In two more cases, dogs were reported to wake up their owners during nocturnal hypoglycemic events. And Tauveron et al. (2006) reported a case of a dog that, while riding in the car with the owner, would suddenly sit up, stare and bark at the owner until the owner stopped the vehicle and checked his blood glucose level. The blood glucose meter would consistently confirm a hypoglycemic episode. Incredibly, these reports do not seem to be isolated. In a survey of 212 dog owners with type 1 diabetes, Wells et al. (2008) found that 36% of respondents believed that their dog showed

behavioural responses to most of their hypoglycemic episodes. Moreover, 33.6% of owners reported that their dogs responded to their hypoglycemic events before they were aware that their blood sugar was low.

Following such reports, organizations throughout North America and the U.K. have begun to train Diabetic Alert Dogs (DADs) for individuals with type 1 diabetes. DAD owners report a decrease in the frequency of both mild and severe hypoglycemic episodes, and consequently, less fear of hypoglycemia and a higher quality of life (Gonder-Frederick, Rice, Warren, Vajda, & Shepard, 2013). These promising reports suggest that olfactory alert dogs could serve as non-invasive, friendly hypoglycemia detection assistants.

As discussed above, it is believed that in these cases, dogs are detecting changes in the VOCs being expressed by their owners that signify physiological change. Empirical studies of canine biomedical detection, and specifically hypoglycemia, have begun to examine and attempt to confirm whether dogs can be trained to reliably detect and identify disease-specific odour signatures. Before a review of these studies, a discussion of the common procedures and statistical analyses used to test and evaluate olfactory processing dogs will be presented.

1.4.1 Common procedures for test of canine olfactory detection. The most commonly used procedure to test olfactory detection and discrimination by dogs is a stimulus line-up. The olfactory line-up is modeled on traditional forensic line-ups where an eye witness attempts to identify a perpetrator out of a line of suspects (Schoon & Haak, 2002). In the case of forensic identification biomedical detection dogs, however,

dogs are presented with olfactory stimuli that contain odour profiles of potential suspects (Curran et al., 2005; Schoon, 1996) or biological diseases (Shirasu & Touhara, 2011).

In forensic applications, testing a dog with a lineup of olfactory stimuli often includes a Matching-to-Sample (MTS) component. With a MTS line-up, a trained dog is first presented with an object recovered from a crime scene (that was left by, touched by, or simply exposed to a perpetrator, Vyplelova et al., 2014). This item serves as the "sample". Then the dog is directed to sniff a series of olfactory stimuli (usually five to seven pieces of clothing or sweat samples) gathered from suspects and arranged in a line on the floor. If one of the stimuli in the line-up is in fact from the perpetrator, the perpetrator's odour profile would be detected on the object and it would match that of the sample, and therefore, the dog should indicate this object as matching the sample.

In tests of biomedical detection, the lineup procedure is again, widely used, and is often recommended as the standard (Jezierski et al. 2015). In biomedical applications, however, a MTS component is not always used because biomedical detection dogs are often trained to detect a single odour signature (the disease signature, as discussed above). So, typically, a dog is presented with a series of five to seven biological stimuli, where one of the stimuli contains a disease-specific odour and the remaining stimuli are healthy, and the dog must indicate the diseased stimulus.

Empirical tests that evaluate the efficacy of scent detection training programs and testing procedures are lacking (Hall et al., 2013; Jezierski et al., 2010). The question of whether the traditional line-up procedure is the most valid method for testing dogs' detection and discrimination of stimuli is debated (Elliker et al., 2014; Gadbois & Reeve, 2014; Johnen et al., 2017). Experiments on perceptual learning show that when stimuli

are presented simultaneously, the perceiver can compare and identify distinctive features, thereby facilitating discrimination (Gibson, 1969). However, we must assume that when a dog completes an olfactory line-up, it is not sampling the odours simultaneously. In a MTS line-up, a dog will smell the sample odour, and then sniff the other stimuli in succession attempting to find the odour profile that matches the sample. As the dog is sniffing successive stimuli in rapid succession, it is accumulating both sensory/perceptual (odour) and mnemonic (sensory memory and working memory) interference. Therefore, a MTS line-up procedure with five to seven stimuli is more a test of a dog's perceptual and mnemonic abilities than it is of its ability to detect and discriminate a target odour from distracting odours (Gadbois & Reeve, 2014).

Therefore, as will be presented and discussed in Chapter 2, we advocate for a simpler version of the traditional line-up that utilizes only two or three stimuli. In presenting fewer stimuli, the detection of the target odour becomes less cognitively demanding and therefore learning is accelerated (as evidenced by the findings presented in Chapter 2) and the dogs' accuracy may increase as a result.

Evidence from eyewitness identification lineups show that people often use "relative judgements" to identify perpetrators in a simultaneous presentation of lineup members. That is, in a lineup that does not contain the perpetrator, eye witnesses may choose the person that looks most like the perpetrator relative to the others in the lineup (Lindsay & Wells, 1985). Consequently, the rate of false identification errors is high with sequential presentation of stimuli (Lindsay & Wells, 1985; Steblay et al., 2001). Therefore, Lindsay and Wells (1985) advocated the use of a sequential presentation of lineup members, one at a time. Their 1985 study showed that sequential presentation of

stimuli resulted in eye witnesses committing fewer false identifications than when stimuli were presented simultaneously, without compromising correct indications. Applying this line of evidence to olfactory detection tasks, presenting dogs with one stimulus at a time and requiring a "yes" or "no" response may result in more accurate representations of dogs' abilities to detect target odours.

Sequential presentation of olfactory stimuli can be achieved using a Go/No-Go procedure. Contrary to the simultaneous presentation of stimuli in a line-up, single stimuli are presented sequentially with the Go/No-Go procedure, and a dog is required to identify the stimulus as the target odour (the "Go" component) or not (the "No-Go" component). Using this procedure, the dog has less opportunity to compare amongst stimuli. While this may seem counterproductive to learning, a Go/No-Go procedure may be more ecologically valid for certain applications of alert dogs. For example, if we consider the deployment of dogs for biomedical detection, it would be much more practical for a clinician to present a dog with a single sample and have the dog identify it as possessing the target condition or not, as opposed to having to gather a series of samples and present them to a dog in a line-up procedure. A second example is the case of a trained and deployed biomedical alert dog. An alert dog must be able to detect indicators of a condition (e.g., seizure, hypoglycemia) without ever comparing its owner to another person, or explicitly between different states of the same person. Furthermore, the use of the Go/No-Go procedure allows for the calculation of informative statistics that are not so easily calculated with a line-up procedure such as the d' and C statistics used in Signal Detection Theory that quantify observers' stimulus detectability and bias in decision making.

One factor that must be carefully considered when training and testing dogs for biomedical detection and alert is the incidence of the target condition. Although no empirical studies have examined how the prevalence of target-positive trials influences dogs' response rates (Lit, 2009), human research shows that when there is a low prevalence of trials where a positive indication is required, detection rates are low. Evans, Birdwell, and Wolfe (2013) examined how the prevalence of breast cancer-positive mammography scans influenced radiologists' detection of breast cancer. In the low-prevalence condition where positive scans made up roughly 1% of the total scans, radiologists missed 30% of positive mammograms, whereas in the high prevalence condition where positive mammograms made up 50% of cases presented, radiologists missed only 12% of positive cases.

Wolfe et al. (2013) replicated this prevalence effect in airport checkpoint screener's detection of safety threats in baggage. However, they then showed that if screeners were presented with a series of high prevalence trials, their performance then improved on subsequent low prevalence trials. Given that biomedical detection and alert dogs would likely encounter low prevalence rates of target stimuli in real world applied settings, particularly in the case of hypoglycemia detection where a dogs' owner may go days or weeks without a hypoglycemic event, this low prevalence effect should be examined in biomedical detection and alert dogs (Lit, 2009).

1.4.2 Statistical evaluation of performance. In tests of canine biomedical detection, a dog's performance is assessed by recording its response on a binary classification test; whether a target odour is present or absent. If the target odour is present the dog can either identify it as present; termed a true positive or a hit, or the dog

may incorrectly identify the odour as being absent; a false negative, otherwise known as a miss (equivalent to a Type II error). Conversely, if the target odour is not present, the dog may identify it as such; a true negative, or a correct rejection, or the dog may incorrectly identify the target odour as being present; a false positive, or a false alarm (equivalent to a Type I error) (see Figure 1). Using the line-up procedure, researchers interpret the lack of an indication in response to control stimuli as a correct rejection.

Given a total number of trials in which a target odour may be present or absent, the proportion of the types of responses a dog commits can be used to calculate different statistics to evaluate the dog's performance (see Gadbois and Reeve, 2016). Which statistics are calculated, however, varies depending on convention in a given field. Since biomedical detection dogs may be studied from both a medical perspective and a psychophysical perspective, I will present the most commonly used statistics for both fields, thereby allowing the reader to understand the dogs' performance regardless of their domain of study. A short description of these statistics follows.

The most commonly reported statistics in the literature of biomedical detection dogs are those from medical diagnostics: sensitivity and specificity. Sensitivity (probability of detection, true positive rate) is the proportion of positives correctly identified as such, usually expressed as a percent. Applied to canine biomedical detection, sensitivity would be a measure of a dog's ability to identify a target odour when the target odour is present. Sensitivity is calculated as follows:

$$\text{Sensitivity} = \frac{\text{hits}}{\text{hits+misses}} \text{ X } 100$$

Specificity (true negative rate) is the proportion of negatives correctly identified as such, typically expressed as a percent. Applied to canine biomedical detection, specificity

is a measure of a dog's ability to identify what is *not* the target odour (what odours are not associated with the disease or condition of interest). Specificity is calculated as follows:

$$\text{Specificity} = \frac{\text{correct rejections}}{\text{correct rejections+false alarms}} \text{ X } 100$$

Other measures of interest commonly reported in medical diagnostics include precision (positive predictive value) and accuracy. In assessing the performance of a biomedical detection dog, precision is a measure of how consistent the dog is; how consistently they present the same response pattern (regardless of whether it is the correct response given the stimulus). Precision is calculated as:

$$\text{Precision} = \frac{\text{hits}}{\text{hits+false alarms}} \text{ X } 100$$

Accuracy is akin to percent correct in that it provides a measure of how many correct decisions are made in all trials (hits and correct rejections). In canine biomedical detection, accuracy provides a measure of both how well the dog indicates when the target odour is present (hits) and how well the dog can indicate that the target odour is absent (correct rejections). Accuracy is calculated as:

$$\text{Accuracy} = \frac{\text{hits+correct rejections}}{\text{hits+misses+correct rejections+false alarms}} \text{ X } 100$$

In psychophysics, Signal Detection Theory (SDT) is used to measure an observer's ability to distinguish between a target stimulus (termed the "signal" in SDT) and distracting stimuli (termed "noise"). SDT assumes that, due to an observer's own internal noise or uncertainty as to whether they have perceived the stimulus or not, the detection of stimuli can be represented stochastically within an observer's brain; the probability that a target stimulus will be present (or absent) is distributed normally (Kingdom &

Prins, 2016). Therefore, as seen in Figure 1, the probability distribution of the signal can be presented alongside the probability distribution of noise.

If differentiation between the signal and noise is difficult, the probability distributions will overlap, as depicted in Figure 1. In this case, an observer may decide that the stimulus is present when it is not, thereby committing a false alarm, or they may decide the signal is absent, when it is actually present, thereby committing a miss. If discrimination between the signal and noise is clearer, however, the two curves will not overlap as much (or at all) and an observer will be less likely to commit misses and false alarms.

The statistic d' ("dee prime"), termed the sensitivity index, provides a measure of the distance between the means of the signal and the noise distributions. It is therefore a measure of detectability. d' is defined in terms of a standardized distribution, the details of which will not be discussed here (but see Green & Swets; 1966; Macmillan & Creelman, 2005). Essentially, the variables used to calculate d' (hits and false alarms) are converted to *z* scores (standard deviation units) and d' is calculated as:

$$d' = z_{\text{hits}} - z_{\text{false alarms}}$$

where a small d' value would indicate overlap between the signal and noise distributions, and a large d' would indicate little overlap between the distributions. In considering the types of responses an observer may commit, an "ideal observer" is one "who make(s) optimal use of the information in the stimulus in making their decisions" (Macmillan & Creelman, 2005 pp. 267); they minimize both false alarms and misses. SDT recognizes, however, that most observers have a response bias; one observer may be more likely to commit false alarms and another observer may be more likely to commit

misses, regardless of their actual sensitivity to the signal. The most common measure of bias, the Criterion, C, is calculated as:

$$C = -\frac{(z_{hits} + z_{false\ alarms})}{2}$$

In Figure 1, the Criterion is represented by the "neutral criterion" vertical line in-between the two distributions. As discussed above, an ideal observer would have no bias towards committing false alarms or misses, and as a result, his/her Criterion value would be close to zero. Therefore, the areas under the distributions that represent the false alarms and misses would be equal (as seen in Figure 1). If an observer was a liberal decision maker, however, his/her Criterion value would be negative, and the line would be shifted to the left, resulting in the area under the distribution representing false alarms to be larger than the area representing the misses. Conversely, a conservative decision maker would commit more misses, his/her Criterion value would be positive, and the Criterion line would be shifted to the right resulting in the area under the distribution that represents misses being larger than that of the area representing false alarms.

The Criterion is a valuable calculation when assessing a biomedical detection dog because it tells you about a dog's internal bias in decision making. A dog's internal biases in decision making may be a product of a few different factors. First, a dog may simply have a rigid decision making process as a result of previous training or cognitive heuristics. Second, motivational factors may influence a dog's decision making process. For instance, a hungry dog may become a more liberal decision maker if they are rewarded with food for correct identifications. Finally, a dog's decision making bias may be a product of the testing procedure. For example, although not empirically examined in dogs, evidence from human studies of Go/No-Go tests suggest that if more "Go" trials

are presented than "No-Go" trials, a dog may develop a liberal decision making bias (Chikazoe et al., 2009). Knowledge of a dogs' tendency for bias is important information with respect to biomedical detection and alert dogs. In the case of hypoglycemia detection, it would be advantageous to have a detection dog with a liberal bias. Although a liberal dog may commit false alarms, signaling hypoglycemic events when they are not occurring, it will be less likely to commit misses; neglecting to signal hypoglycemic events when they occur.

In comparing medical diagnostic statistics and SDT, one can see that both calculate measures of sensitivity. But they differ in an important way; the error term used in diagnostic sensitivity uses the proportion of misses whereas d' uses false alarms. In other words, diagnostic sensitivity uses both the "yes" and "no" responses and the d' calculation utilizes only the "yes" responses. This difference may represent the emphasis diagnostic analysis puts on misses. With diagnostic tests, there is an obvious need for a balance between false alarms and misses. It could be said, however, that misses are costlier to the patient than false alarms. In cases where early detection of disease is imperative for positive prognosis, a missed diagnosis is potentially life threatening. Moreover, in the case of hypoglycemia detection, missed detection of a hypoglycemic event could ultimately result in death. But, as discussed above, the most important difference between diagnostic statistics and SDT is that SDT allows researchers to identify the bias of a detection dog, independent of their sensitivity. With this detailed knowledge of a dog's response profile handlers may alter the dog's distribution of responses through procedural manipulations and reward contingencies, and ultimately allows handlers to assess a dog's utility in specific applied tasks.

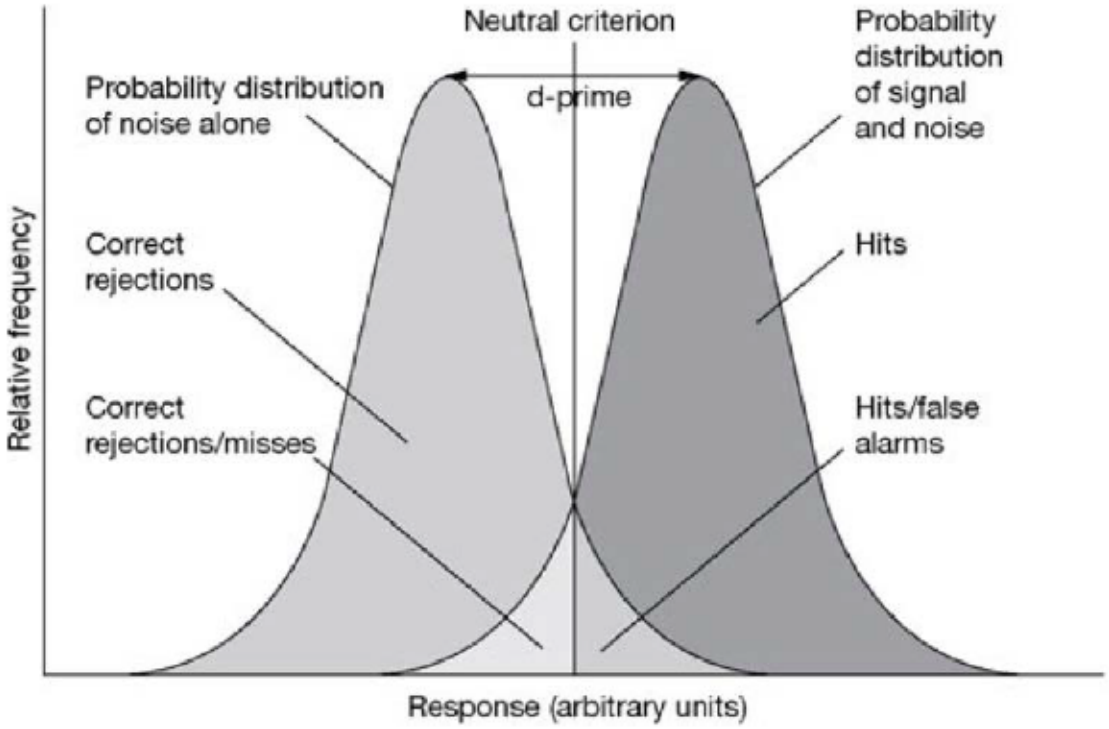

Figure 1.1. The distribution of yes and no responses as modeled by Signal Detection Theory. From: https://www.nature.com/nrneurol/journal/v4/n6/images/ncpneuro0794-f1.jpg

1.4.3 Review of canine biomedical detection literature. The study of canine biomedical detection is a burgeoning field of research. Beginning in the early 2000s, empirical tests have been conducted to examine dogs' ability to detect a wide range of conditions including infections and infectious agents, cancers, and diabetic hypoglycemia. What follows here is a review of these empirical studies of canine biomedical detection.

1.4.3.1 Infections and infectious agents. In 2012, Bomers et al. trained a beagle to detect the bacteria *C. diffficile*, which commonly infects people in healthcare facilities, causing severe gastrointestinal ailments (Bomers et al., 2012). After successfully training the dog to discriminate between odour imprints of stool samples positive for *C. difficile* and odour imprints of stool samples negative for *C. difficile,* Bomers et al. (2012) then tested whether the dog could detect people infected with *C. difficile* in a hospital setting. They reported that the dog accurately identified 25 out of 30 people infected (83% sensitivity) and correctly ignored 265 out of 270 non-infected people (98% specificity).

Koivusalo et al. (2017) trained a golden retriever to detect cultured Methicillin-resistant *S. aureus* (MRSA); an infection-causing bacteria that is resistant to methicillin treatment. Koivusalo et al. (2017) reported that the dog could discriminate between MRSA and a similar, yet methicillin sensitive strain, *S. aureus* (MSSA), as well as methicillin resistant *S. epidermidis* (MRSE) and MSSA with a SCC*mec* remnant (MSSAr) at a high level of sensitivity and specificity.

1.4.3.2 Cancer. Early detection of cancer is imperative to increase survivability, but research is needed to develop more accurate early detection tests (Szulejko et al., 2010). Given that many potential VOCs have been implicated as potential biomarkers of

different cancers, canine biomedical detection appears a promising alternative to current diagnostic tests. Therefore, the majority of empirical studies of canine biomedical detection have examined dogs' ability to detect cancer. A review of these studies follows, including canine detection of melanoma, colorectal, bladder, ovarian, breast, prostate, and lung cancer.

1.4.3.2.1 Melanoma. Following the intriguing report by Williams and Pembroke of a dog identifying a cancerous mole on its owner, Pickel et al. (2004) examined whether two dogs could be trained to detect melanoma. After initial training and testing procedures showed promising results, Pickel et al. (2004) had the dogs sniff people with skin areas suspected by a dermatologist to be cancerous. The dogs' indications were later compared with biopsy results. Results revealed that the first dog correctly indicated cancerous tissue on five people, and failed to indicate the skin on one person as cancerous that was later confirmed as such. On a seventh person, the first dog indicated one sample as cancerous that was initially found to be non-cancerous, but upon further histopathological examination, was in fact found to contain cancerous cells. The second dog searched four of the seven people and provided the same responses as the first dog. The findings of this study provide very promising evidence that dogs can be trained to detect melanoma with a high degree of accuracy.

1.4.3.2.2 Colorectal Cancer. Sonoda et al. (2011) report the only empirical study examining dogs' ability to detect colorectal cancer, with equally promising results. After training, Sonoda et al. (2011) tested the ability of one dog to detect colorectal cancer in breath and watery stool samples presented in a five-station line-up. In the breath sample test, the dog detected the colorectal sample with 91% sensitivity and 99% specificity, and

in the watery stool sample test, the dog detected the cancerous sample with 97% sensitivity and 99% specificity.

1.4.3.2.3 Bladder Cancer. Willis et al. (2004) examined dogs' ability to detect bladder cancer with less promising results than the aforementioned studies. Here, researchers presented the dogs with a target urine sample obtained from a person with bladder cancer, and six control samples from healthy people and people with non-bladder malignancies. Willis et al. (2004) reported that, overall, the dogs could detect the cancerous samples with a mean accuracy rate of 41%. They conclude that the dogs' performance was greater than the chance level of one in seven (14%) and therefore, the dogs could detect bladder cancer in urine at above chance levels.

1.4.3.2.4 Ovarian Cancer. Horvath et al. (2008, 2010, 2013) have conducted a series of studies of studies examining dogs' ability to detect ovarian cancer in both tumor tissue and blood. Beginning in 2008, Horvath et al. examined dogs' ability to detect ovarian cancer. In their first study, Horvath et al. (2008) trained a dog to detect cancerous ovarian tissue samples and then tested its ability to locate it amongst non-cancerous tissue samples (fat, intra-abdominal muscle). In this test, the dog detected the cancerous ovarian samples with sensitivity and specificity values of 100%. In a second test, Horvath et al. (2008) examined whether the dog was simply detecting an odour associated with gynecological tissues by presenting the cancerous ovarian samples with other gynecological tumor samples, such as cervical, vulvar, and endometrial carcinomas as the control samples. The use of other gynecological tissues as controls had very little impact on the dog's performance; the dog successfully detected the ovarian cancer samples with 100% sensitivity and 91% specificity.

Then, in 2010, Horvath et al. furthered their study of canine detection of ovarian cancer by training the same dog to detect blood samples from people with ovarian cancer, while a second dog was trained to detect cancerous ovarian tissue (as was done in their 2008 study). Then the dogs were tested for their ability to detect both cancerous tissue and blood samples against control samples from healthy people and from people with other gynecological carcinomas. The first dogs' sensitivity and specificity detecting blood samples from people with ovarian cancer were both 100%. And, even though the second dog had never been tested with blood samples, her detection sensitivity was 100%, and specificity was 96%. In the tests with tissues samples, the first dog's detection sensitivity and specificity were 100% and 94% respectively, and the second dog's sensitivity were 100% and specificity 94%.

Finally, in a follow-up study, Horvath et al. (2013) tested the same two dogs' ability to detect ovarian cancer in the blood of two groups of people: people undergoing chemotherapy (blood was taken between the fifth and sixth courses) and people who had completed their last round of chemotherapy three to six months prior. In the first test, the dogs showed the ability to detect ovarian cancer odours with sensitivity and specificity of 97% and 99% respectively. In the second test, the dogs positively indicated three out of ten samples at both the three and sixth month follow-up tests. Incredibly, all three of individuals whose samples the dogs indicated as cancerous had recurrences of cancer.

1.4.3.2.5 Breast Cancer. McCulloch et al. (2006) conducted a study examining dogs' ability to detect breast cancer in exhaled breath samples. After training with a set of cancerous and healthy control samples, they tested the dogs by presenting new samples in a five-station line-up where one station contained a cancerous breath sample, and the

remaining four contained breath samples from healthy controls. The results showed that, overall, the dogs could detect the cancerous samples with 88% sensitivity and 98% specificity.

A second study examining dog's detection of breast cancer was conducted by Gordon et al. (2008), using urine samples. In this study, the researchers recruited professional dog trainers to train their dogs on their own time and in their own homes. Four dogs were trained in this manner. Gordon et al. (2008) believe that the lack of standardized and controlled training procedures was responsible for the overall poor performance of the dogs during the testing trials; none of the dogs achieved detection sensitivity at levels above chance, and only two of the dogs had detection specificity at levels just above chance.

1.4.3.2.6 Prostate Cancer. The most common screening tool for prostate cancer is to test for Prostate Specific Antigen (PSA) in the blood. This test, however, has been shown to have poor diagnostic utility (Stamey et al., 2004) and efforts have been made to develop better diagnostic tests including testing dogs' detectability of VOCs associated with prostate cancer.

Cornu et al. (2011) tested one dogs' detectability of prostate cancer in urine samples. The urine samples were obtained from men who showed elevated PSA levels or had abnormal digital rectal findings. After prostate biopsies revealed the presence or absence of prostate cancer, the urine samples were classified as cancerous or controls accordingly. After training with a series of cancerous and control samples was complete, Cornu et al. (2011) tested the dogs with samples they had not previously smelled. The samples were presented in a six-station line-up with one station containing a cancerous

urine sample and the remaining five containing control urine samples. The results showed the dogs' detection sensitivity and specificity were both 91%. Interestingly, the dog committed three false alarms, and after these men had new biopsies conducted, one of the men was diagnosed with prostate cancer. These results suggest that canine detection of prostate cancer may be more sensitive than traditional blood and rectal exams, offering a promising non-invasive alternative. However, additional studies do not show the same promising results.

Elliker et al. (2014) began by training ten dogs for the detection of prostate cancer in urine samples. Cancerous urine samples were obtained from men with biopsy-confirmed prostate cancer and control samples were obtained from men with a normal PSA test. After two training stages, however, only two dogs were formally tested with new samples. The results showed that the first dog could only detect the prostate cancer samples with detection sensitivity of 13% and specificity of 71%. The second dog's detection sensitivity was 25% and specificity was 75%. In interpreting their results, Elliker et al. (2014) make the important point that it is possible that the dogs memorized the odour of the training samples, therefore identifying individual odours as opposed to generalizing a common odour of prostate cancer across samples.

The third study to examine dogs' abilty to detect prostate cancer was conducted by Gordon et al. (2008) in the same series of studies where they tested dogs' ability to detect breast cancer. Using the same training procedure, four dogs were trained by their owners in their homes. During testing, the dogs were presented with urine samples from people with biopsy-confirmed prostate cancer alongside urine samples from people with normal PSA test results. The samples were presented in a seven-station line-up with one

cancerous sample and six healthy control samples. The results of the prostate cancer test were similar to those of the breast cancer test; none of the dogs achieved detection sensitivity above chance levels, and only two of the four dogs reached above chance levels of specificity.

1.4.3.2.7 Lung Cancer. One of the most common causes of cancer death is lung cancer, due, in part to late diagnosis (Siegel et al., 2016). Therefore, there appears to be a focus in canine biomedical detection to examine dogs' potential as early detectors of lung cancer.

McCulloch et al. (2006) tested five dogs' ability to detect breath samples obtained from people with biopsy-confirmed lung cancer against breath samples from healthy people. The samples were presented in a five-station line-up with one cancerous sample and four healthy samples. The dogs' overall detection sensitivity and specificity was 99%.

In a second study examining canine detection of lung cancer, Buszewski et al. (2012b) obtained two sets breath samples from people with lung cancer and healthy controls. The first set of samples was analyzed for VOC content using GC-MS and the second set was presented to dogs (the number of dogs is not mentioned). The dog tests used a five-station line-up where one station contained a cancerous sample and the other four contained healthy samples. The dogs' performance showed detection sensitivity of 82.2% and specificity of 82.4%. Interestingly, a comparison of the dogs' positive indications with the results of the GC-MS analysis revealed that the dogs' indications were positively correlated with specific VOCs in the breath: ethyl acetate and 2-pentanone.

In a further two studies, researchers again presented dogs with breath samples from people with confirmed lung cancer, but the control samples included other lung pathologies. In a series of tests, Ehmann et al. (2012) presented four dogs with lung cancer samples against healthy control and against samples from people with Chronic Obstructive Pulmonary Disorder (COPD). Overall, the dogs' detection sensitivity was 72% and specificity was 94%.

In 2014, Rudnicka et al. presented two dogs with breath samples from people with lung cancer against breath samples from people with asthma, Chronic Obstructive Pulmonary Disease, or synthetic samples, as well as from healthy people. Samples were presented to the dogs in a five-station line-up with one station containing a cancerous breath sample and the other four containing control samples. The dogs' overall detection sensitivity and specificity were 86% and 72% respectively. Furthermore, Rudnicka et al. (2014) used CG-MS to examine the VOCs present in breath samples from healthy people, people with lung cancer, and people with other lung diseases. Their analyses revealed that the concentration of acetone, isoprene, ethanol, 1-propanol, 2-propanol, hexanal, and dimethyl sulfide were higher in patients with lung cancer than in healthy volunteers and people with other lung diseases.

Finally, Amundsen et al. (2014) examined whether dogs could detect lung cancer in a series of heterogeneous samples (malignant and benign); thereby testing their performance in a task more comparable to a real world applied test. To begin, Amundsen et al. (2014) trained the dogs to detect lung cancer in breath and urine samples against healthy controls in a six-station roundel (a method like a line-up, except samples are presented on an apparatus in a circle rather than a line). After successful training, three

dogs were tested with the heterogeneous breath samples from people with either malignant or benign conditions. The results showed that the dogs' overall detection sensitivity was 56% and specificity was 33%. Therefore, despite evidence that dogs can detect VOCs associated with lung cancer, empirical tests report differing levels of success. But, taken together, studies of canine detection of cancer show that dogs may serve as a promising alternative to current diagnostic tests, and in some cases, show a higher degree of sensitivity than traditional measures (Cornu et al., 2011). Further research is required before dogs can be deployed in real-world clinical applications.

1.4.3.3 Diabetic Hypoglycemia. To the best of my knowledge, only two empirical studies of canine detection of hypoglycemia have been conducted. In 2013, Dehlinger et al. tested whether three dogs that were serving as DADs in the homes of individuals with Type 1 Diabetes could identify hypoglycemia in sweat samples from three individuals unknown to the dogs. Dehlinger et al. (2013) obtained sweat samples using the same protocol used to collect the samples for the initial training of the dogs; individuals rubbed two cotton balls on the skin of their arms during two episodes of hypoglycemia, and then an additional two cotton balls were rubbed on the skin of their arms during each of two normoglycemic episodes. To indicate the detection of hypoglycemia, the three dogs had previously been trained to push a bell after smelling a hypoglycemic sample. When Dehlinger et al. (2013) presented the dogs with the sweat samples from the people unknown to the dogs, however, none of the three dogs could indicate the hypoglycemic samples at above chance levels.

In a second study, Hardin et al. (2015) tested the ability of six dogs to detect hypoglycemic samples. The biological samples used by Hardin et al. (2015) were

combined breath and sweat samples on gauze, donated by four people during normoglycemic and hypoglycemic states. In this study, researchers trained the dogs to signal the hypoglycemic samples in three phases. In the first phase, the dogs were trained to smell the hypoglycemic sample and sit. In the second phase, the dogs were presented with hypoglycemic and normoglycemic samples donated by the same person, as well as blank stimuli, and were required to locate the hypoglycemic sample and sit. In the third phase, the dogs were trained to smell a hypoglycemic sample held on a person's body and poke the person. Once a dog successfully completed all three phases of training, they were introduced to the hypoglycemic and normal glycemic samples from the other donors and training continued until the dogs could discriminate between the samples. Then, to test the dogs' ability to detect the hypoglycemic samples, Hardin et al. (2014) presented the samples in a seven-station line-up where one station contained a hypoglycemic sample, two stations contained normoglycemic samples, and the remaining four stations contained blank gauze. Hardin et al. (2015) reported that the dogs could detect the hypoglycemic samples with sensitivity ranging from 50% to 87.5% (two of the dogs demonstrated lower sensitivity than the other four), and specificity ranging from 89.6 to 97.9%. While these results may seem very impressive, it is important to note that during testing, the dogs were only ever presented with four hypoglycemic samples that they had been reinforced for identifying many times during the initial training phases. It is possible that the dogs tested by Hardin et al. (2015) memorized the odour of the individual samples, rather than identifying an odour signature of hypoglycemia that is consistent across samples. Therefore, Hardin et al. (2015)'s conclusion that dogs can be trained to identify hypoglycemia is misleading.

Taken together, the empirical studies of canine detection of hypoglycemia provide little evidence that dogs are detecting an odour signature of hypoglycemia. Therefore, the experiments presented here in Chapter 4 will seek to contribute to the field of canine detection of hypoglycemia by testing whether dogs can be trained to detect an odour associated with hypoglycemia that generalizes across hypoglycemic events.

1.5 Outline of Dissertation Papers

The goal of my dissertation was to address current gaps in the literature of canine biomedical detection. As a new and developing field, little empirical evaluation of training and testing procedures exists in the literature. Furthermore, concerns exist surrounding the collection and storage of biological samples. Finally, empirical studies of hypoglycemia are few and their results are inconsistent. Therefore, the studies presented here seek to contribute valuable procedures and findings to the field of canine biomedical detection, specifically canine detection of hypoglycemia for people with type 1 diabetes.

1.5.1 Outline of chapter 2. My first series of studies, as presented in Chapter 2, describe and evaluate a novel program to train dogs to detect human breath. Given that there is a notable gap in the literature examining the efficacy of current training and testing procedures (Hall et al., 2013; Jezierski et al., 2010), this manuscript contributes to the literature by presenting a training and testing program that utilizes a procedure not commonly used in current tests of canine olfactory detection: a cued, 3 Alternative Forced Choice procedure. In this manuscript, the authors and I present our "low saliency training" procedure, where dogs are trained to detect progressively lower concentrations of orange pekoe tea, then tea breath, and finally, clean breath. We then present two experiments to evaluate the utility of the training program. In the first experiment, we

examined dogs' ability to detect and discriminate between breath samples from three different people. In the second experiment, we tested dogs' ability to discriminate between breath samples from one person at three different times of the day. The results showed that the dogs transitioned from the training phase to the testing phases with ease. The results of Experiment 1 showed that the dogs could discriminate between breath samples from different people at above chance levels, and the results of Experiment 2 showed that all the dogs further discriminated between breath samples from one person at three different times of the day at above chance levels.

1.5.2 Outline of chapter 3. The manuscript in Chapter 3 presents studies concerning the maintenance of breath sample integrity. In studies of canine detection of breath samples, breath samples are sometimes stored for up to six months (Ehmann et al., 2012). However, no studies have ever examined whether the detectability of these samples decreases with storage time. With the knowledge that our future studies would require breath samples to be stored for up to a month, we therefore sought to examine how to increase sample integrity over time. With the help of Dr. Peter Wentzell from Dalhousie's Department of Chemistry, we developed a procedure for coating cotton balls in silicone oil. The liquid phase coating the cotton ball allows more VOCs to be absorbed onto the cotton ball compared to uncoated cotton. As a result, diffusion of the compounds into the surrounding air will persist for longer than the diffusion of VOCs off uncoated cotton balls. We then tested whether the use of silicone-coated cotton balls results in breath samples being detectable longer than breath samples obtained using uncoated cotton balls in two experiments. In the first experiment, we tested four dogs' detectability of breath samples on coated and uncoated cotton up to two hours after the breath samples

were donated. In the second experiment, we tested the same four dogs' detectability of breath samples that were collected using coated and uncoated cotton and then stored for one, two, three, and four weeks. The results showed that, in Experiment 1, the silicone coated cotton balls did not affect the dogs' detection of the breath samples at up to two hours after sample donation. The results of Experiment 2 showed that, for two of the four dogs, the silicone coated cotton balls improved the detectability of the breath samples from one week onward. For the remaining two dogs, the silicone coated cotton balls did not affect their detectability of the breath samples up to three and four weeks after sample donation.

1.5.3 Outline of chapter 4. In Chapter 4, I present three studies in which we tested dogs' ability to detect and discriminate between breath samples from people with type 1 diabetes. As discussed above, currently only two empirical studies of canine detection of hypoglycemia have been conducted, and their results are inconsistent. Therefore, the series of studies presented here serve to further the study of canine detection of hypoglycemia. In the first experiment, we used a cued, 3AFC procedure to test four dogs' ability to discriminate between breath samples obtained from people with type 1 diabetes during normoglycemia, hypoglycemia, and hyperglycemia. Then in the second experiment we used a Go/No-Go procedure to test two dogs' ability to generalize the odour of hypoglycemia across multiple hypoglycemic events from one person. In the third experiment, we tested one dog's ability to generalize the odour of hypoglycemia across multiple hypoglycemic events from two different people. The results of the first experiment showed that all four dogs could discriminate between the breath samples from the different glycemic states at above chance levels. The results of the second experiment

showed that one dog generalized the odour of hypoglycemia across multiple samples from one person. The results of the third experiment were difficult to interpret and would require further study before concluding whether the dog could generalize the odour of hypoglycemia across different people.

1.6 References

Acerini, C. (2016). The rise of technology in diabetes care. Not all that is new is necessarily better. *Pediatric diabetes, 17,* 168-173. doi: 10.1111/pedi.12366

Amundsen, T., Sundstrøm, S., Buvik, T., Gederaas, O. A., & Haaverstad, R. (2014). Can dogs smell lung cancer? First study using exhaled breath and urine screening in unselected patients with suspected lung cancer. *Acta Oncologica*, *53*, 307-315. doi: 10.3109/0284186X.2013.819996

Angle, C., Waggoner, L. P., Ferrando, A., Haney, P., & Passler, T. (2016). Canine detection of the Volatilome: a review of implications for pathogen and disease detection. *Frontiers in Veterinary Science*, *3,* 1-7. doi: 10.3389/fvets.2016.00047

Aye, T., Block, J., & Buckingham, B. (2010). Toward closing the loop: an update on insulin pumps and continuous glucose monitoring systems. *Endocrinology and Metabolism Clinics of North America*, *39*, 609-624. doi:10.1016/j.ecl.2010.05.005

Bade-White, P. A., & Obrzut, J. E. (2009). The neurocognitive effects of type 1 diabetes mellitus in children and young adults with and without hypoglycemia. *Journal of Developmental and Physical Disabilities*, *21*, 425-440. doi: 10.1007/s10882-009-9151-y

Becker, D. J., & Ryan, C. M. (2000). Hypoglycemia: a complication of diabetes therapy in children. *Trends in Endocrinology & Metabolism*, *11*, 198-202.

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*, e7396. doi: 10.1136/bmj.e7396

Brands, A. M., Biessels, G. J., De Haan, E. H., Kappelle, L. J., & Kessels, R. P. (2005). The effects of type 1 diabetes on cognitive performance. *Diabetes Care*, *28*, 726-735.

Buckingham, B., Block, J., Burdick, J., Kalajian, A., Kollman, C., Choy, M., ... & Chase, P. (2005). Response to nocturnal alarms using a real-time glucose sensor. *Diabetes Technology & Therapeutics*, *7*, 440-447.

Buckingham, B., Wilson, D. M., Lecher, T., Hanas, R., Kaiserman, K., & Cameron, F. (2008). Duration of nocturnal hypoglycemia before seizures. *Diabetes Care*, *31*, 2110-2112. doi: 10.2337/dc08-0863

Buszewski, B., Ligor, T., Jezierski, T., Wenda-Piesik, A., Walczak, M., & Rudnicka, J. (2012). Identification of volatile lung cancer markers by gas chromatography–mass spectrometry: comparison with discrimination by canines. *Analytical and Bioanalytical Chemistry*, *404*, 141-146. doi: 10.1007/s00216-012-6102-8

Chen, M., Daly, M., Williams, N., Williams, S., Williams, C., & Williams, G. (2000). Non-invasive detection of hypoglycaemia using a novel, fully biocompatible and patient friendly alarm system. *BMJ*, *321*, 1565-1566.

Chikazoe, J., Jimura, K., Asari, T., Yamashita, K. I., Morimoto, H., Hirose, S., ... & Konishi, S. (2009). Functional dissociation in right inferior frontal cortex during performance of go/no-go task. *Cerebral Cortex*, *19*, 146-152. doi:10.1093/cercor/bhn065

Church, J., & Williams, H. (2001). Another sniffer dog for the clinic?. *The Lancet*, *358*, 930.

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. doi: 10.1016/j.eururo.2010.10.006

Craven, B. A., Paterson, E. G., & Settles, G. S. (2009). The fluid dynamics of canine olfaction: unique nasal airflow patterns as an explanation of macrosmia. *Journal of The Royal Society Interface*, *7,* 933-943. doi: 10.1098/rsif.2009.0490

Curran, A. M., Rabin, S. I., Prada, P. A., & Furton, K. G. (2005). Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *Journal of Chemical Ecology*, *31*, 1607-1619. doi: 10.1007/s10886-005-5801-4

Davis, S. N., Shavers, C., Mosqueda-Garcia, R., & Costa, F. (1997). Effects of differing antecedent hypoglycemia on subsequent counterregulation in normal humans. *Diabetes*, *46*, 1328-1335.

De Beaufort, C. (2006). Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. *Diabetic Medicine: a Journal of the British Diabetic Association*, *23*, 857-866. doi: 10.1111/j.1464-5491.2006.01925.x

Dehlinger, K., Tarnowski, K., House, J. L., Los, E., Hanavan, K., Bustamante, B., ... & Ward, W. K. (2013). Can trained dogs detect a hypoglycemic scent in patients with type 1 diabetes?. *Diabetes Care*, *36*, e98-e99. doi: 10.2337/dc12-2342

de Lacy Costello, B., Amann, A., Al-Kateb, H., Flynn, C., Filipiak, W., Khalid, T., ... & Ratcliffe, N. M. (2014). A review of the volatiles from the healthy human body. *Journal of Breath Research*, *8*, 014001. doi:10.1088/1752-7155/8/1/014001

Ehmann, R., Boedeker, E., Friedrich, U., Sagert, J., Dippon, J., Friedel, G., & Walles, T. (2012). Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *European Respiratory Journal*, *39*, 669-676.

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 urology*, *14*, 22. doi: http://www.biomedcentral.com/1471-2490/14/22

Evans, K. K., Birdwell, R. L., & Wolfe, J. M. (2013). If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PLoS One*, *8*, e64366. doi:10.1371/journal.pone.0064366

Frier, B. M. (2004). Morbidity of hypoglycemia in type 1 diabetes. *Diabetes Research and Clinical Practice*, *65*, S47-S52. doi:10.1016/j.diabres.2004.07.008

Frier, B.M (2008). How hypoglycaemia can affect the life of a person with diabetes. *Diabetes/Metabolism Research and Reviews, 24,* 87-92. doi: 10.1002/dmrr.796

Gadbois S., Reeve C. (2014). Canine olfaction: Scent, sign, and situation. In: *Horowitz, A. (Ed.), Domestic dog cognition and behavior (*pp 3-29). New York, NY: Springer.

Gadbois, S., & Reeve, C. (2016). The semiotic canine: scent processing dogs as research assistants in biomedical and environmental research. *DOG BEHAVIOR*, *2*, 26-32. doi 10.4454/db.v2i3.43

Galassetti, P. R., Novak, B., Nemet, D., Rose-Gottron, C., Cooper, D. M., Meinardi, S., ... & Blake, D. R. (2005). Breath ethanol and acetone as indicators of serum glucose levels: an initial report. *Diabetes Technology & Therapeutics*, *7*, 115-123. doi: https://doi.org/10.1089/dia.2005.7.115

Gibson, E.J. (1969). *Principles of perceptual learning and development.* New York, NY: Appleton-Century-Crofts.

Gonder-Frederick, L., Nyer, M., Shepard, J. A., Vajda, K., & Clarke, W. (2011). Assessing fear of hypoglycemia in children with type 1 diabetes and their parents. *Diabetes Management*, *1*, 627-639. doi:10.2217/DMT.11.60

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. doi: 10.2337/dc12-1998

Gonder-Frederick, L., Zrebiec, J., Bauchowitz, A., Lee, J., Cox, D., Ritterband, L., ... & Clarke, W. (2008). Detection of hypoglycemia by children with type 1 diabetes 6 to 11 years of age and their parents: a field study. *Pediatrics*, *121*, e489-e495. doi: 10.1542/peds.2007-0808

Gordon, R. T., Schatz, C. B., Myers, L. J., Kosty, M., Gonczy, C., Kroener, J., ... & Arthur, N. (2008). The use of canines in the detection of human cancers. *The Journal of Alternative and Complementary Medicine*, *14*, 61-67. doi: 10.1089/acm.2006.6408

Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Gschwend, S., Ryan, C., Atchison, J., Arslanian, S., & Becker, D. (1995). Effects of acute hyperglycemia on mental efficiency and counterregulatory hormones in adolescents with insulin-dependent diabetes mellitus. *The Journal of Pediatrics*, *126*, 178-184.

Hall, N. J., Smith, D. W., & Wynne, C. D. (2013). Training domestic dogs (Canis lupus familiaris) on a novel discrete trials odor-detection task. *Learning and Motivation*, *44*, 218-228. http://dx.doi.org/10.1016/j.lmot.2013.02.004

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*(4), 509-517. doi: 10.1007/s13300-015-0135-x

Hershey, T., Perantie, D. C., Warren, S. L., Zimmerman, E. C., Sadler, M., & White, N. H. (2005). Frequency and timing of severe hypoglycemia affects spatial memory in children with type 1 diabetes. *Diabetes Care*, *28*, 2372-2377.

Helton, W. S. (2009). Canine ergonomics: Introduction to the new science of working dogs. In W. S. Helton (Ed.), *Canine ergonomics. The science of working dogs* (pp. 1–16). Boca Raton, FL: Taylor and Francis Group.

Horvath, G., Andersson, H., & Nemes, S. (2013). Cancer odor in the blood of ovarian cancer patients: a retrospective study of detection by dogs during treatment, 3 and 6 months afterward. *BMC cancer*, *13*, 396. doi: http://www.biomedcentral.com/1471-2407/13/396

Horvath, G., Andersson, H., & Paulsson, G. (2010). Characteristic odour in the blood reveals ovarian carcinoma. *BMC cancer*, *10*, 643. doi: http://www.biomedcentral.com/1471-2407/10/643

Horvath, G., Järverud, G. A. K., Järverud, S., & Horváth, I. (2008). Human ovarian carcinomas detected by specific odor. *Integrative cancer therapies*, *7*, 76-80. doi: 10.1177/1534735408319058

Howsmon, D., & Bequette, B. W. (2015). Hypo-and hyperglycemic alarms: devices and algorithms. *Journal of diabetes science and technology*, *9*, 1126-1137. doi: 10.1177/1932296815583507

Jeitler, K., Horvath, K., Berghold, A., Gratzer, T. W., Neeser, K., Pieber, T. R., & Siebenhofer, A. (2008). Continuous subcutaneous insulin infusion versus multiple daily insulin injections in patients with diabetes mellitus: systematic review and meta-analysis. *Diabetologia, 51,* 941-951. doi: 10.1007/s00125-008-0974-3

Jezierski, T., Gorecka-Bruzda, A., Walczak, M., Swiergiel, A. H., Chruszczewski, M. H., & Pearson, B. L. (2010). Operant conditioning of dogs (Canis familiaris) for identification of humans using scent lineup. *Animal Science Papers and Reports*, *1*, 81-93.

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*, 027001. doi:10.1088/1752-7155/9/2/027001

Kaminski, J., & Nitzschner, M. (2013). Do dogs get the point? A review of dog–human communication ability. *Learning and Motivation*, *44*, 294-302. http://dx.doi.org/10.1016/j.lmot.2013.05.001

Kingdom F.A.A. & Prins N. (2016). *Psychophysics. A practical introduction*. London, England: Elsevier.

Klonoff, D. C. (2005). Continuous glucose monitoring. *Diabetes care*, *28*, 1231-1239.

Koivusalo, M., Vermeiren, C., Yuen, J., Reeve, C., Gadbois, S., & Katz, K. (2017). Canine scent detection as a tool to distinguish meticillin-resistant Staphylococcus aureus. *Journal of Hospital Infection*, *96*, 93-95. doi: http://dx.doi.org/10.1016/j.jhin.2017.03.005

Lit, L. (2009). Evaluation learning tasks commonly applied in detection dog training. In W. S. Helton (Ed.), *Canine ergonomics. The science of working dogs* (pp. 99–114). Boca Raton, FL: Taylor and Francis Group.

Ly, T.T., Maahs, D.M., Rewers, A., Dunger, D., Oduwole, A., & Jones, T.W. (2014). Assessment and management of hypoglycemia in children and adolescents with diabetes. *Pediatric Diabetes, 15,* 180-192.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. New Jersey: Lawrence Erlbaum Associates Inc.

Marcus, D.A., & Bhowmick, A. (2013). Survey of migraine sufferers with dogs to evaluate for canine migraine-alerting behaviors. *The Journal of Alternative and Complementary Medicine, 19,* 501-508.

Matkya, K.A. (2002) Sweet dreams? – nocturnal hypoglycemia in children with type 1 diabetes. *Pediatric Diabetes, 3*, 74-81. doi: 10.1034/j.1399-5448.2002.30203.x

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. doi: 10.1177/1534735405285096

Miklosi, A. (2007). *Dog Behaviour, Evolution, and Cognition*. Oxford University Press: Oxford, UK.

Minh, T. D. C., Blake, D. R., & Galassetti, P. R. (2012). The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes Research and Clinical Practice*, *97*, 195-205. doi: 10.1016/j.diabres.2012.02.006

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. doi:10.1016/j.jveb.2010.01.002

Novak, B. J., Blake, D. R., Meinardi, S., Rowland, F. S., Pontello, A., Cooper, D. M., & Galassetti, P. R. (2007). Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *Proceedings of the National Academy of Sciences*, *104*, 15613-15618. doi: www.pnas.org cgi doi 10.1073 pnas.0706533104

Neupane, S., Peverall, R., Richmond, G., Blaikie, T. P., Taylor, D., Hancock, G., & Evans, M. L. (2016). Exhaled breath isoprene rises during hypoglycemia in type 1 diabetes. *Diabetes Care*, *39*, e97-e98. doi: 10.2337/dc16-0461

O'Connell, S. M., Cooper, M. N., Bulsara, M. K., Davis, E. A., & Jones, T. W. (2011). Reducing rates of severe hypoglycemia in a population-based cohort of children and adolescents with type 1 diabetes over the decade 2000–2009. *Diabetes Care*, *34*, 2379-2380. doi: 10.2337/dc11-0748

O'Haire, M. (2010). Companion animals and human health: Benefits, challenges, and the road ahead. *Journal of Veterinary Behavior: Clinical Applications and Research*, *5*, 226-234. doi:10.1016/j.jveb.2010.02.002

Ovalle, F., Fanelli, C. G., Paramore, D. S., Hershey, T., Craft, S., & Cryer, P. E. (1998). Brief twice-weekly episodes of hypoglycemia reduce detection of clinical hypoglycemia in type 1 diabetes mellitus. *Diabetes*, *47*, 1472-1479.

Phillips, M. (1992). Breath tests in medicine. *Scientific American*, *267*, 74-79.

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. doi:10.1016/j.applanim.2004.04.008

Riemsma, R., Ramos, I. C., Birnie, R., Büyükkaramikli, N., Armstrong, N., Ryder, S., ... & Kleijnen, J. (2016). Integrated sensor-augmented pump therapy systems [the MiniMed® Paradigm™ Veo system and the Vibe™ and G4® PLATINUM CGM (continuous glucose monitoring) system] for managing blood glucose levels in type 1 diabetes: a systematic review and economic evaluation. *Health Technology Assessment, 20*, 1 - 288. doi: 10.3310/hta20170

Rudnicka, J., Walczak, M., Kowalkowski, T., Jezierski, T., & Buszewski, B. (2014). Determination of volatile organic compounds as potential markers of lung cancer by gas chromatography–mass spectrometry versus trained dogs. *Sensors and Actuators B: Chemical*, *202*, 615-621. doi: 10.1183/09031936.00051711

Schabert, J., Browne, J. L., Mosely, K., & Speight, J. (2013). Social stigma in diabetes. *The Patient-Patient-Centered Outcomes Research*, *6*, 1-10. doi: 10.1007/s40271-012-0001-0

Schmidt, K., & Podmore, I. (2015). Current challenges in volatile organic compounds analysis as potential biomarkers of cancer. *Journal of biomarkers*, *2015,* 1-16. http://dx.doi.org/10.1155/2015/981458

Schoon, G. A. A. (1996). Scent identification lineups by dogs (Canis familiaris): experimental design and forensic application. *Applied Animal Behaviour Science*, *49*, 257-267.

Schoon, G.A. & Haak, R. (2002). *K9 suspect discrimination: Training and practicing scent identification line- ups*. Alberta, Canada: Detselig Enterprises.

Settles, G. S., Kester, D. A., & Dodson-Dreibelbis, L. J. (2003). The external aerodynamics of canine olfaction. In *Sensors and sensing in biology and engineering* (pp. 323-335). Vienna: Springer.

Shirasu, M., & Touhara, K. (2011). The scent of disease: volatile organic compounds of the human body related to disease and disorder. *Journal of Biochemistry*, *150*, 257-266. doi: 10.1093/jb/mvr090

Shivers, J. P., Mackowiak, L., Anhalt, H., & Zisser, H. (2013). "Turn it off!": diabetes device alarm fatigue considerations for the present and the future. *Journal of Diabetes Science and Technology*, *7*, 789-794.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: a cancer journal for clinicians*, *66*, 7-30. doi: 10.3322/caac.21332

Sonoda, H., Kohnoe, S., Yamazato, T., Satoh, Y., Morizono, G., Shikata, K., ... & Inoue, F. (2011). Colorectal cancer screening with odour material by canine scent detection. *Gut, 60,* 814-819. doi:10.1136/gut.2010.218305

Stamey, T. A., Caldwell, M., McNEAL, J. E., Nolley, R., Hemenez, M., & Downs, J. (2004). The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? *The Journal of Urology*, *172*, 1297-1301. doi: 10.1097/01.ju.0000139993.51181.5d

Strong, V., Brown, S.W., & Walker, R. (1999). Seizure alert dogs – fact or fiction? *Seizure, 8,* 62-65.

Szulejko, J. E., McCulloch, M., Jackson, J., McKee, D. L., Walker, J. C., & Solouki, T. (2010). Evidence for cancer biomarkers in exhaled breath. *IEEE Sensors Journal*, *10*, 185-210.

Tauveron, I., Delcourt, I., Desbiez, F., Somda, F., & Thiéblot, P. (2006). Canine detection of hypoglycaemic episodes whilst driving. *Diabetic Medicine*, *23*, 335-335.

Unger, J., & Parkin, C. (2011). Recognition, prevention, and proactive management of hypoglycemia in patients with type 1 diabetes mellitus. *Postgraduate medicine*, *123*, 71-80.

Udell, M. A., & Wynne, C. D. (2008). A review of domestic dogs' (*Canis Familiaris*) human-like behaviors: or why behavior analysts should stop worrying and love their dogs. *Journal of the experimental analysis of behavior*, *89*, 247-261. doi: 10.1901/jeab.2008.89-247

Vyplelová, P., Vokálek, V., Pinc, L., Pacáková, Z., Bartoš, L., Santariová, M., & Čapková, Z. (2014). Individual human odor fallout as detected by trained canines. *Forensic Science International*, *234*, 13-15. doi: http://dx.doi.org/10.1016/j.forsciint.2013.10.018

Waggoner, L. P., Jones, M. H., Williams, M., Johnston, J. M., Edge, C. C., & Petrousky, J. A. (1998, December). Effects of extraneous odors on canine detection. In *Enabling Technologies for Law Enforcement and Security* (pp. 355-362). International Society for Optics and Photonics. doi:10.1117/12.335008

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. doi:10.1016/j.applanim.2005.07.009

Weinzimer, S., Xing, D., Tansey, M., Fiallo-Scharer, R., Mauras, N., Wysocki, T., ... & Ruedy, K. (2009). Prolonged use of continuous glucose monitors in children with type 1 diabetes on continuous subcutaneous insulin infusion or intensive multiple-daily injection therapy. *Pediatric Diabetes*, *10*, 91-96. doi: 10.1111/j.1399-5448.2008.00476.x

Wells, D. L., Lawson, S. W., & Siriwardena, A. N. (2008). Canine responses to hypoglycemia in patients with type 1 diabetes. *The Journal of Alternative and Complementary Medicine*, *14*, 1235-1241. doi: 10.1089/acm.2008.0288

Wild, D., von Maltzahn, R., Brohan, E., Christensen, T., Clauson, P., & Gonder-Frederick, L. (2007). A critical review of the literature on fear of hypoglycemia in diabetes: Implications for diabetes management and patient education. *Patient Education and Counseling*, *68*, 10-15. doi:10.1016/j.pec.2007.05.003

Williams, H., & Pembroke, A. (1989). Sniffer dogs in the melanoma clinic? *The Lancet*, *333*, 734.

Willis, C. M., Church, S. M., Guest, C. M., Cook, W. A., McCarthy, N., Bransbury, A. J., ... & Church, J. C. (2004). Olfactory detection of human bladder cancer by dogs: proof of principle study. *BMJ*, *329*, 712 – 714.

Wolfe, J.M., Brunelli, D.N., Rubinstein, J., Horowitz, T.S. (2013). Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *Journal of Vision, 13,* 1-9. doi: 10.1167/13.3.33

Zijlstra, E., Heise, T., Nosek, L., Heinemann, L., & Heckermann, S. (2013). Continuous glucose monitoring: quality of hypoglycaemia detection. *Diabetes, Obesity and Metabolism*, *15*, 130-135.

CHAPTER 2: A NOVEL METHOD FOR TRAINING DOGS TO DETECT AND DISCRIMINATE HUMAN BREATH SAMPLES

The manuscript prepared for this study is presented below. Catherine Reeve, under the supervision of Dr. Simon Gadbois, was responsible for devising the research questions, training protocol, and plotting and interpreting the dogs' performance. She was the lead on dog training and testing, with the support of her co-authors. Catherine wrote the initial draft of the manuscript, and received and incorporated feedback from her co-authors and committee members. The manuscript will be submitted for publication in the near future. The full reference for this manuscript is:

Reeve, C., Wallace, K., & Gadbois, S. (2017). A novel method for training dogs to detect and discriminate human breath samples.

2.1 Abstract

Despite the growing interest in the use of dogs for disease diagnosis and alert, little research has examined the best training protocols and testing procedures for biomedical detection and alert dogs. The current study presents a novel three-phase training program to train dogs to detect breath samples. All three phases used a cued, three-alternative forced choice procedure. In the first phase, five dogs were trained to detect decreasing concentrations of orange pekoe tea as a target odour. In the second phase, four dogs were trained to detect tea breath, and in the third phase, the dogs detected clean breath. Subsequently, the training program was evaluated in two discrimination experiments. In the first experiment, the dogs were required to discriminate between three breath samples from three different people. In the second experiment, the dogs were tested on their ability to discriminate between three breath samples from the same person donated at three different times of the day. The results of both experiments showed that all four dogs successfully discriminated between the breath samples at above chance levels, suggesting that the training program was effective in training dogs to detect and discriminate between human breath samples with a high degree of accuracy.

2.2 Introduction

Domestic dogs' (*Canis lupus familiaris*) incredible sense of smell (Craven et al., 2010) and proclivity for training has resulted in their successful application in a wide variety of olfactory detection tasks including bomb and narcotics detection, conservation applications, search and rescue, forensic applications, and biomedical detection (for reviews, see Helton, 2009; Gadbois and Reeve, 2014). In the case of forensic and biomedical detection, the dogs are faced with similar tasks because they are required to detect and discriminate human odours that are of very low saliency. When dogs detect and discriminate human odours, they are smelling volatile organic compounds (VOCs); carbon based compounds produced during cell metabolism that are gases at room temperature (Schmidt & Podmore, 2015). VOCs are exhaled in the breath during pulmonary circulation, and are present in the headspace of breath, sweat, urine, and feces samples (Kusano et al., 2011; Schmidt & Podmore, 2015). Sweat and breath samples, which are most commonly used in forensic and biomedical tests of canine detection, show variation in their VOC content between individuals (Curran et al., 2005; Phillips et al., 1999), and can therefore serve as person-specific odour profiles. Furthermore, when cells experience metabolic change or disease, the concentration and type of VOCs emitted change. Therefore, VOCs in breath and sweat may further serve as diagnostic olfactory biomarkers (Shirasu & Touhara, 2011).

Given the incredible sensitivity of a dog's nose, however, VOCs are odours that dogs have likely been exposed to, without reward, for their entire lives. Therefore, in training dogs for forensic and biomedical applications, a good training protocol would

increase the incentive salience of previously ignored VOCs by teaching dogs that such odours are relevant and rewarding. In the current literature, a lot of variety exists in the training protocols used for canine detection of human odours. Empirical evaluations of the efficacy of different training and testing protocols for scent detection dogs, however, are lacking.

With respect to forensic applications, dogs are usually required to identify criminals based on sweat samples. Traditionally, canine identification of a perpetrator is done using a matching to sample (MTS) line-up procedure with anywhere from five to 12 stimuli (Schoon & Haak, 2002). Using this procedure, law enforcement would collect items left or touched by a perpetrator at a crime scene, and would also collect sweat samples from a series of potential suspects. Then in the lab, the dog would first sniff the crime scene stimulus, the "sample", and then be directed to sniff a series (typically five to seven) of the suspect sweat stimuli arranged beside one another in a "line-up". If one of the suspects was in fact the perpetrator, its sweat stimulus will match the sample, and that will be indicated by the dog.

Empirical studies that evaluate dogs' ability to match human odours using a MTS line-up procedure are few, but those that do exist show inconsistencies and methodological concerns. For example, Brisbin and Austad (1991) found that three dogs previously trained on human scent discrimination could successfully discriminate between their handler's hand scent and blank controls in 93.1% of trials, but could only discriminate between their handler's hand scent and a stranger's hand scent in 75.7% of trials. Brisbin and Austad (1991) also found that dogs could not detect and discriminate a target individual's odour from strangers when the target individual's sample was taken

from somewhere other than their hand, but Settle et al. (1994) found that dogs could match individual odour samples in a six-stimulus line-up, regardless of where on the body the sample was obtained from. Furthermore, Schoon (1996) found that methodological details in scent identification tests such as the number of stimuli presented to the dog, whether positive control trials are completed, and the strength of the sample odour can influence the performance of specially trained human scent detection dogs. Therefore, evaluation of a dog's performance is not straightforward. However, despite the documented variability, dogs' performance in such tests are admitted as evidence in some European countries (Schoon, 1996, 2005). Given the findings highlighted above, the potential for inaccurate dog identification of suspects is real and it is therefore imperative that the validity of training programs and dogs' corresponding accuracy be evaluated carefully.

Similar issues arise in tests of canine biomedical detection. Unlike forensic applications where the dogs must identify person-specific odours, the goal for canine biomedical detection is for dogs to identify disease-specific odour profiles (Shirasu & Touhara, 2011) that generalize across multiple different people. In tests of canine biomedical detection, biological samples such as breath, sweat, saliva, urine, feces, or tissues may be presented to the dog. These samples are typically presented using the same line-up procedure discussed above (Jezierski et al., 2015), but, beyond initial training, there is no MTS component (no "sample" is presented) because, in most cases, the dogs are looking for the same target odour profile in every trial. In this case, some researchers may use a "cue" sample to remind the dog of the particular odour they are searching for.

Empirical studies of canine detection of disease have reported promising results in the detection of cancers such as melanoma (Pickel et al., 2004), lung cancer (McCulloch et al., 2006; Buszewski et al., 2012a; Ehmann et al., 2012; Rudnicka et al., 2014), breast cancer (McCulloch et al., 2006), colorectal cancer (Sonoda et al., 2011), prostate cancer (Cornu et al., 2011), and ovarian cancer (Horvath et al., 2008, 2010, 2013). There is also evidence that dogs can successfully detect *C. difficile* (Bomers et al., 2012), MRSA (Koivusalo et al., 2017), and diabetic hypoglycemia (Hardin et al., 2015). But further studies show that dogs may have difficulty detecting some of the same conditions including lung cancer (Amundsen et al., 2014), breast and prostate cancer (Gordon et al., 2008; Elliker et al., 2014), bladder cancer (Willis et al., 2004), and diabetic hypoglycemia (Dehlinger et al., 2013). Such inconsistent findings prevent the clinical application of dogs as detectors of disease. As a new field of research, inconsistencies across studies of canine biomedical detection could be the result of many factors (for a discussion of these variables, see Elliker et al., 2014; Jezierksi et al., 2015; Johnen et al., 2017), one of which could be the training program and sample presentation. Elliker et al. (2014) suggest that the traditional line-up presentation of samples may not be optimal for the level of detection and discrimination required by biomedical detection dogs. As discussed in Gadbois and Reeve (2014), we agree with the proposition put forth by Elliker et al., (2014). We argue that line-ups of five to seven stimuli actually test dogs' working memory more than they test dogs' ability to detect a target stimulus and discriminate it from other stimuli. Furthermore, a dog completes quick sniffing bouts in succession as they approach each stimulus in a line-up. In doing so they experience sensory memory and perceptual interference as the odours from each new stimulus accumulate (Gadbois &

Reeve, 2014). As a result, traditional line-up procedures make olfactory detection tasks more difficult than necessary and therefore may misrepresent dogs' abilities.

Therefore, despite the potential for scent detection dogs to make important contributions in forensic and biomedical applications, their implementation has been limited because of inconsistencies in the results of empirical studies and procedural limitations. A closer examination of the design of the training and testing protocols may benefit these fields.

Jezierski et al., (2010) presented the first empirical evaluation of a multi-phase program designed to train dogs to detect target human odours and discriminate between different individuals. Here, researchers trained six dogs using a MTS, five-station line-up procedure. After preliminary training phases that encouraged the dogs to sniff and indicate at scent stations (see Jezierski et al., 2010 for details of the full training program), the dogs were trained on palm samples obtained by having people hold cotton swabs in their hands for 15 minutes. In the first phase with palm samples, one station contained a target palm sample, and the remaining four stations contained blank control samples. After this phase, one station contained the target palm sample while the other four stations contained decoy palm samples from different individuals. In the final "working phase", the dogs were presented with both samples from training, plus palm samples obtained from police forensic situations. Jezierski et al. (2010) found that in the initial training phases the dogs' performance was very high, but decreased over successive training phases, with the percent of correct responses on the working phase being only 58.1%. Jezierski et al. (2010) concluded that the dogs could be easily trained on a line-up procedure, but that their performance during training did not necessarily

predict their performance with new stimuli. Taken together, the results of Jezierski et al. (2010) combined with those of the canine forensic and biomedical detection studies underscore the need for effective training programs and valid procedures for canine detection.

Moreover, as discussed above, some biomedical alert and forensic identification tasks require dogs to smell odours that they have smelled throughout their lives without reward, resulting in the dogs learning to ignore, habituate to, or simply not attend these "irrelevant" odours. In the Canid Behaviour Research Lab at Dalhousie University, we train dogs for biomedical detection. Previous work in our lab showed that after dogs had been trained to successfully perform Matching-to-Sample tasks with cooking herbs, they subsequently could not match human breath samples (pilot study, unpublished data). We assumed that the dogs were unable to detect human breath samples because of learned irrelevance. Therefore, we developed the training program presented here to teach dogs to focus on breath samples.

After assessing highly motivated dogs, five dogs were trained. Our training program consisted of three phases. In the first phase, dogs were presented with decreasing concentrations of orange pekoe tea. The goal of this first phase was to first assess the motivation and persistence of the dogs to complete a repetitive task, and second, to train the dogs to detect odours of decreasing saliency. If a dog could not successfully complete the first phase of training, they were unlikely to be successful with subsequent training phases. In the second phase of the training program, we attempted to bridge the gap between the tea stimulus and a human breath stimulus by presenting the dogs with a "tea breath" stimulus. In the third and final phase of the training program, the

dogs were presented with "clean breath samples". After the training program was complete, we tested dogs' ability to discriminate between human breath samples in two experiments. In the first experiment, we tested dogs' ability to discriminate between three breath samples from three different individuals. In the second experiment we tested the dogs' ability to discriminate between three breath samples from one individual donated at three different times of the day.

2.3 Training Program

2.3.1 Method

2.3.1.1 Subjects. Dogs were recruited from the surrounding city by word of mouth, and were assessed for basic obedience, food or play drive, and motivation to work. Five dogs were selected and trained for the current study: Nutella (3 years old at the time of training, intact), Bella (2 years old), Mist (3 years old), and Charlee (2 years old) were female purebred border collies and Koda (1 year old) was a male border collie mix.

The dogs were brought to the Canid Behaviour Research Lab at Dalhousie University once or twice a week, as this training schedule has been found to result in optimal task acquisition (Demant et al., 2011). Work days lasted two to three hours depending on the owner's availability. All dogs were trained individually, with no other dogs present at the time of training. All procedures were approved by the University Committee on Laboratory Animals before the study was conducted.

2.3.1.2 Preliminary training. All training procedures used reward-based positive reinforcement methods, with the use of a clicker as a secondary reinforcer and kibble or the toss of a ball as a primary reinforcer. If a dog was not previously clicker trained, its first day working in the lab was spent clicker training.

2.3.1.2.1 ***Shaping an indication response.*** In shaping the desired indication response, we first trained the dogs to approach the stimuli apparatus and to indicate samples using a "poke-and-hold" behaviour. First, a piece of kibble was placed inside a stainless-steel jar with holes in the lid. The jar was then placed on the floor of the testing room and when a dog approached and sniffed the jar a clicker was activated and a food reward was given. Once a dog consistently oriented towards and sniffed the jar, a cotton ball was placed inside the jar alongside the kibble, and 3mL of tea that had steeped for 15 minutes was dispensed onto the cotton (see procedure below). Then, when the dog approached and sniffed the jar, researchers required that the dog hold its nose against the jar progressively longer before the clicker was activated. Once a dog could hold its nose against the jar for five full seconds, the kibble was removed and further trials were conducted until the dog could reliable hold its nose against the jar containing only tea for five seconds. Then, a second jar containing only cotton was presented simultaneously with the jar containing the tea stimulus. The dog was encouraged to sniff both jars and indicate the jar containing the tea using the poke-and-hold. Once a dog could reliably indicate the jar containing the tea, it began the first phase of the training program.

2.3.1.3 Training program. The training program was composed of three phases. The general procedure was the same for all three phases, but different stimuli were presented in each phase. A description of the general procedure will be presented first, followed by the details of each individual phase.

2.3.1.3.1 General procedure. The dogs were tested at Dalhousie's Canid Behaviour Lab. The lab contains three adjoining rooms: Room 1 where the dogs stayed between work sessions, Room 2, a small interior room connecting the other two rooms,

and Room 3, where the testing was done (see Figure 2.1). Dogs began a work session by waiting with a handler inside Room 2 while the test stimuli were set up in Room 3 by a researcher. When a test trial began, the door to Room 3 was opened and the dog was led in.

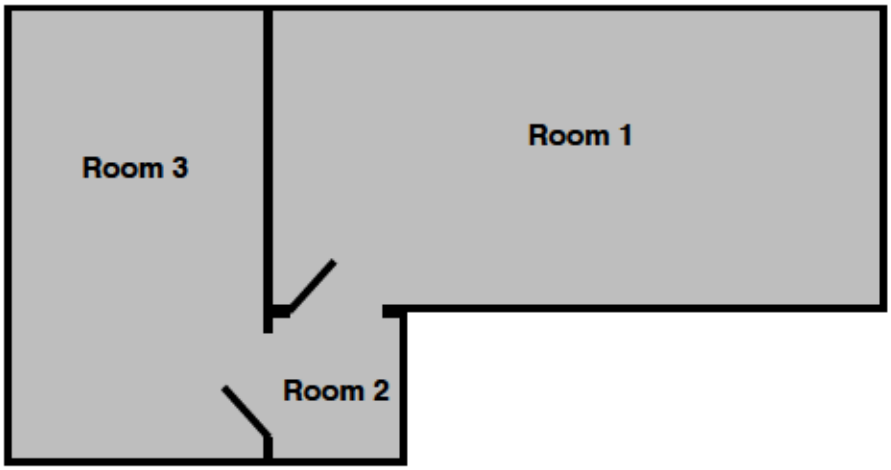

Figure 2.1. The layout of the Canid Behaviour Research Lab at Dalhousie University. Room 1 is where dogs spend time when they are not being tested. Room 2 is where dogs wait between test trials, and Room 3 is where testing takes place.

The samples were presented to the dogs using a cued, 3 Alternative Forced Choice (3AFC) procedure. When the door between rooms two and three was opened, a jar containing the odour sample was presented to the dog to sniff, serving as a scent "cue". Then the dog was directed towards the apparatus holding the other three stimuli; one containing a cotton ball with an odour sample identical to the cue, and two control stimuli. The dog was required to indicate which stimulus contained the odour sample that matched the cue by using a previously trained "nose hold" behaviour where the dog held its nose against the correct jar or tube for 4 seconds. A successful indication was rewarded with a few pieces of kibble or a few throws of a ball. If the dog indicated incorrectly, a researcher uttered "nope" using a gentle tone, and the trial was complete. After completing a trial, the dog was brought into a small adjoining room and the door was closed, allowing researchers to randomize the position of the jars for the next trial.

Each session consisted of 10 trials. The position of the target odour stimulus relative to the control stimuli (left, middle, or right) was determined for each trial using a semi-randomization procedure of rolling a die (if the die rolled 1 or 2 the tea stimulus was in the left position, 3 or 4, in the middle position, and 5 or 6, in the right position). Constraints were applied to the randomization procedure such that the tea stimulus had to be in each position (left, right, or middle) at least three times within the 10 trials, the tea stimulus could not be in the same position for more than three consecutive trials, and the position of the tea could not follow a pattern (such as left, right, left, right) for more than two iterations. Dogs completed between three and five sessions during a work period, depending on the amount of time they could spend in the lab.

2.3.1.3.2 Phase 1: Tea. The odour used for the first phase of training was varying concentrations of Orange Pekoe tea. The required concentration of tea was prepared at the beginning of every work session. Tea was prepared by placing one tea bag in a 530mL teapot and adding boiled water on top of the tea bag until the water reached the teapot's fill line. Decreasing concentrations of tea were prepared by allowing the tea to steep for: 15 minutes, ten minutes, seven minutes, five minutes, three minutes, one minute, 45 seconds, 30 seconds, 15 seconds, and five seconds. Once the required steep time had been reached, the tea was poured into a glass jar. Two lower concentrations were created by diluting the tea steeped for five seconds with water: A half dilution (one part tea steeped for five seconds with one part boiled water) and a quarter dilution (one part tea steeped for five seconds with three parts water). Pure boiled water was also poured from the kettle into a second glass jar. This water was used to prepare the control samples for the later phases of training.

Samples were prepared for presentation by dispensing the liquids onto cotton balls inside stainless steel jars. One cotton ball was placed inside each of four stainless steel jars (5cm high, 5.5cm in diameter, and 18cm circumference). Each jar also contained a strong magnet. Using a 10mL plastic syringe, 3mL of tea was dispensed onto two of the cotton balls. In the initial tea training phase, control jars were prepared by putting one dry cotton ball inside the remaining two jars. In the later phase of tea training, a second syringe was used to dispense 3mL of boiled water onto the cotton balls in the control jars. Lids containing three small holes were placed on all of the jars. Using a permanent marker, a small check mark or "x" was drawn on one side of the jar containing the tea stimulus for identification by researchers.

2.3.1.3.2.1 Apparatus. Three of the stainless-steel jars (one containing the tea stimulus and two blank stimuli) were placed roughly 10 cm apart on a metal sheet 84cm long and 52 cm wide, with the small identification mark on the tea stimulus facing away from the dog. The metal sheet was set on top of a box so that it was raised 32cm off the ground. The magnets inside of the jars ensured that the jars did not slide when the dogs' noses contacted the jar lids. The fourth jar containing the second odour sample was placed on a table near the entrance to the room three (see Figure 2.1).

2.3.1.3.2.2 Criterion. When the dogs first began training, the three stimuli they were presented with were one jar containing cotton with tea and two jars with dry cotton. At later stages of training, an equal amount of boiled water was dispensed onto the dry cotton balls to create more valid control stimuli. The point at which water was added to the control jars was different for each dog depending on how far they had progressed in their training. For every dog, however, when water was added to the control jars, the concentration of the tea stimulus was increased to a previously more concentrated level to make the task easier. At each level of tea concentration, the dogs were required to demonstrate reliable performance before progressing to a lower concentration stimulus. Reliable performance was considered as two to three consecutive sessions of 70% correct or greater (based on chance level calculated using a binomial probability with 10 trials, 0.33 probability of success per trial, and an alpha level of 0.05). If the water in the control jars was introduced at a later stage of training, only one session with performance of 70% correct was required for concentrations they had previously been tested on, before progressing to lower concentrations. Once a dog had demonstrated

the ability to detect the lowest concentration of tea, it was then presented with tea breath samples.

2.3.1.3.3 ***Phase 2: Tea Breath.*** After training with the liquid tea samples was completed, dogs were presented with tea breath samples. A tea breath sample was prepared using a plastic cylindrical tube [poly(ethylene terephthalate), Uline.ca], measuring 20.5 cm in length and 7 cm in diameter, inside which were two cotton balls. A researcher prepared tea by placing a tea bag in a 530mL teapot and filling the teapot with boiling water to the fill line. The tea was steeped for roughly two minutes and then poured into a small glass jar. The tea was left to cool for five to seven minutes and then a researcher sipped the tea and held it in their mouth. After 30 seconds, the researcher spit out the tea and then exhaled two deep breaths through one side of the breath collection tube, then turned the tube around and exhaled a further two deep breaths through the other side of the tube. Then, using tweezers, one of the two cotton balls was removed and placed inside a stainless-steel jar with holes in the lid. This stimulus served as the cue sample (see procedure below). Two additional control tubes were prepared by inserting one cotton ball into each tube, but not exhaling any breath onto them.

2.3.1.3.3.1 Apparatus. The tea breath stimuli were presented to the dogs inside breath sample stations. Each station consisted of a wooden platform inside which a black PVC tube measuring 25cm (see Figure 2A) long stood upright. The PVC tube was longer than the breath collection tube (see Figure 2B), such that when a breath collection tube was dropped inside the PVC tube it sat 4.5cm lower than the opening of the PVC tube (see Figure 2.2). This prevented the dogs' noses from coming into direct contact

with the breath collection tubes. Three sample stations were used to present the stimuli for the study procedure (see Figure 2.3).

2.3.1.3.3.2 *Criterion*. After completing the tea stimulus training, the dogs were presented with the tea breath samples. Sessions where the dogs were presented with tea breath samples consisted of 10 trials, and the dogs completed between three and five sessions per work day. In this phase of the training program, the performance criterion was at least three consecutive sessions at 80%

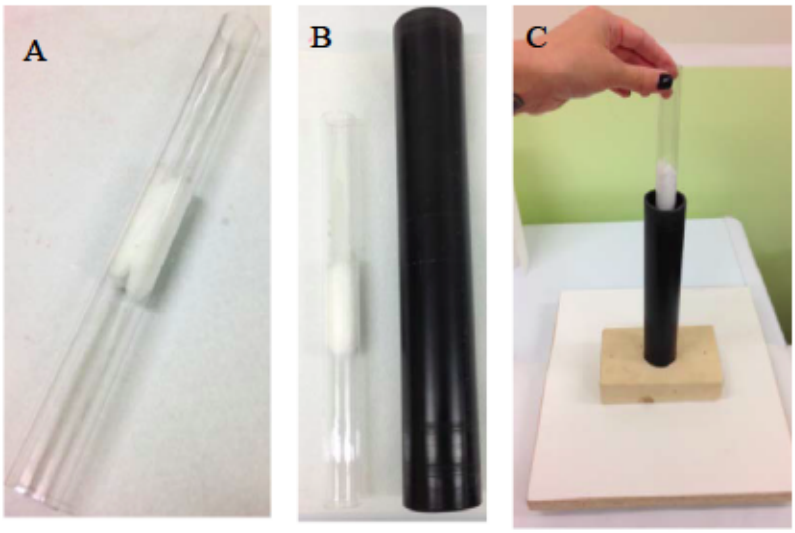

Figure 2.2. The components of a breath sample station. A. The plastic tube containing a cotton ball used for breath sample preparation. B. Illustrates that the breath sample tube is shorter than the black PVC pipe in which the breath sample tube is placed. C. The breath sample tube inside the black PVC pipe is held upright when placed inside the wooden stand. All units combined results in one breath sample station.

Figure 2.3. Three identical sample stations placed on the floor beside one another for presentation to the dogs. One sample station contains a breath sample and the other two contain blank samples. Here, Mist is identifying the station containing the breath sample using the "nose hold" behaviour.

correct or higher. However, as can be seen in Figures 5, 6, and 8, Mist, Koda and Bella completed more sessions and work days after the criterion was achieved. This was because there were often breaks longer than one week between work sessions, so to ensure that their performance was maintained over time, they completed multiple work sessions with the tea breath stimulus before progressing to the clean breath phase.

2.3.1.3.4 Phase 3: Clean Breath. Clean breath samples were prepared using the same breath collection tubes and procedure described above, except a researcher did not first hold tea in his/her mouth. After breath was exhaled through the tube, one of the cotton balls was removed and placed inside a stainless-steel jar with holes in the lid to serve as the cue (see procedure below). Two additional control breath collection tubes were prepared as well.

2.3.1.3.4.1 *Apparatus.* The same sample stations used in phase two were used here in phase three.

2.3.1.3.4.2 *Criterion.* The dogs completed two work sessions each with the clean breath stimulus. To complete the training program, they were required to demonstrate at least three consecutive sessions at 80% correct or higher.

2.3.2 Results and Discussion

Given the large degree of variability between dogs' performance across the training program, their results will be discussed individually.

2.3.2.1 Charlee. When Charlee first began the training program with the liquid tea stimuli, her performance dropped to below the criterion when a lower concentration stimulus was first introduced (see Figure 2.4). Once the stimulus reached the 45 second

steep time concentration, however, her performance was consistently high until the ¼ (five second) concentration. Here she seemed to have initial difficulties but her performance recovered after two sessions. When the water in the control stimuli was introduced, her lowest performance was observed, but it subsequently increased and remained above criterion until the last liquid tea stimulus concentration. Charlee moved away before she could begin the tea breath training.

2.3.2.2 Mist. As illustrated by Figure 2.5, Mist demonstrated the most consistent performance across all phases of the training program. With the exception of one session with the 15 minute liquid tea stimulus, her performance was maintained at above the criterion throughout the liquid tea phase. When the water was added to the control stimuli in the liquid tea phase of training, her performance dropped only slightly, but not below the criterion. When the clean breath stimuli were introduced her performance fell to below criterion, but recovered after three sessions and she reached the criterion in the subsequent three sessions.

2.3.2.3 Koda. Koda's performance was consistently above criterion for most the liquid tea training phase (see Figure 2.6). When the water was introduced to the control stimuli, we increased the concentration of the tea stimulus to seven minutes, and his performance was not affected at all. At the five second steep point, his performance fell to below criterion and remained inconsistent. His performance with the ½ (five second tea) stimulus was his lowest yet, so here we presented him with one session of a five second steep time stimulus to bring his performance back above the criterion. This strategy was successful in that as soon as we re-introduced the ½ (five second) tea stimulus, his performance remained high. During the clean breath phase of the training

program, his performance fell in one session, but in the subsequent four sessions he met the criterion and completed the training.

2.3.2.4 Nutella. Before the water was introduced to the control stimuli, Nutella's performance with the liquid tea stimuli was consistently above criterion, with the exception of three sessions below chance at the 15 minute and five minute stimuli (see Figure 2.7). When the water was introduced to the control stimuli at the ¼ (five second) stimulus, her performance dropped. After this, the stimulus concentration was increased to five seconds and then 30 seconds, at which point her performance improved but remained inconsistent until the five second stimulus. After the five second stimulus was presented, her performance on the liquid tea stimuli was maintained above the criterion. When the tea breath and clean breath stimuli were first introduced, Nutella showed initial decreases in her performance, but met criterion for both phases after eight sessions with the tea breath stimulus and five sessions of the clean breath stimulus.

2.3.2.5 Bella. As illustrated by Figure 2.8, Bella showed the most inconsistent performance throughout the training program. Interestingly, her performance did not seem to be negatively affected by the introduction of the water in the control stimuli, or by decreases in the liquid tea concentrations, as her performance did not systematically fall below criterion at these points. Bella met the criterion at each concentration of tea stimulus, including for the ¼ (five second) tea stimulus. But, given her previously inconsistent performance, we continued with additional sessions after she met the criterion for the ¼ (five second) stimulus to ensure she was prepare for the tea breath stimulus. Throughout these additional sessions, however, her performance decreased. Given the documented high motivation level of border collies, it is possible that her

inconsistent performance was a result of vigilance decrement due to task-induced stress (Helton, 2009). This is highly likely as researchers noted that if Bella were to incorrectly identify a stimulus on one trial, she often become frustrated and distressed, as evidenced by behavioural markers such as hesitation to approach the stimuli and barking. Following one incorrect trial, researchers found that she was much more likely to get subsequent trials wrong. Although Bella's performance was not at criterion at the end of the liquid tea stimulus phase, she did demonstrate the ability to detect the stimulus at criterion in earlier sessions, so the decision was made to continue to the tea breath stimulus phase. Bella required six sessions of tea breath stimuli before reaching the criterion. As with the tea stimuli, additional sessions were completed with the tea breath stimulus after Bella reached criterion to ensure she was prepared for the clean breath stimulus. Much like with the tea stimuli her performance dropped throughout these additional sessions. Given the previous pattern of performance, however, we made the decision to present her with the clean breath stimuli. Bella transferred easily to the clean breath phase and reached criterion within three sessions.

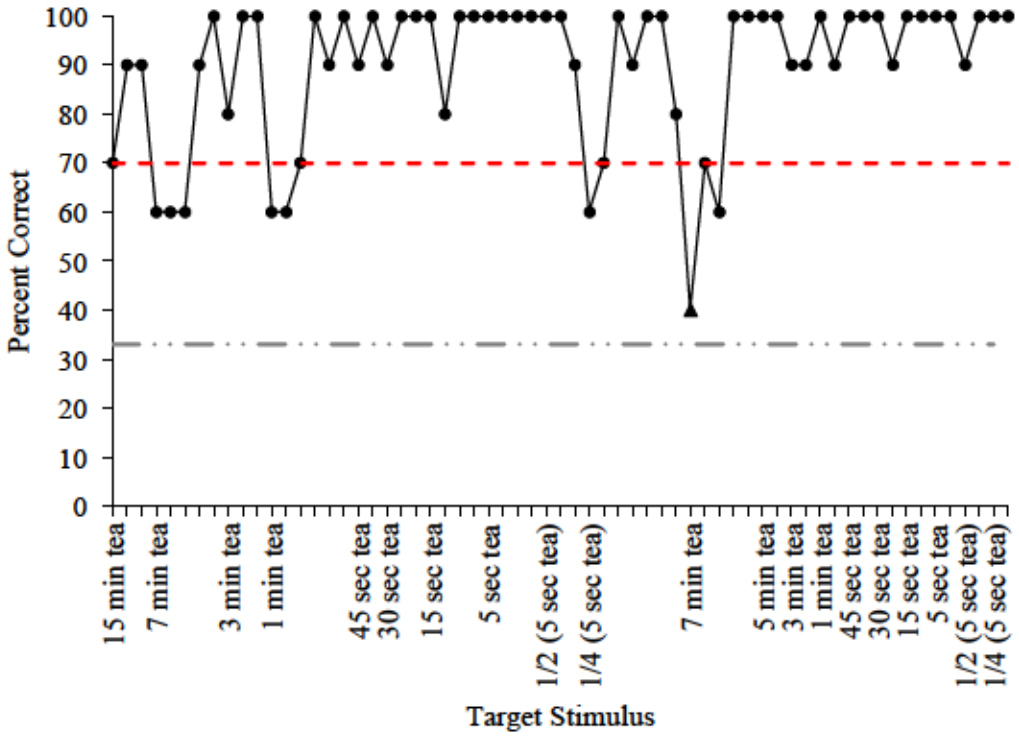

Figure 2.4. Charlee's performance in the first phase of the training program. In the first phase she was presented with liquid tea stimuli at decreasing saliency levels in a cued, 3AFC procedure. The triangle marker indicates the point at which water was added to control stimuli. The dotted line represents the criterion (70%) based on chance level calculated using a binomial probability with 10 trials, 0.33 probability of success per trial (represented by the intermittent dashed line), and an alpha level of 0.05.

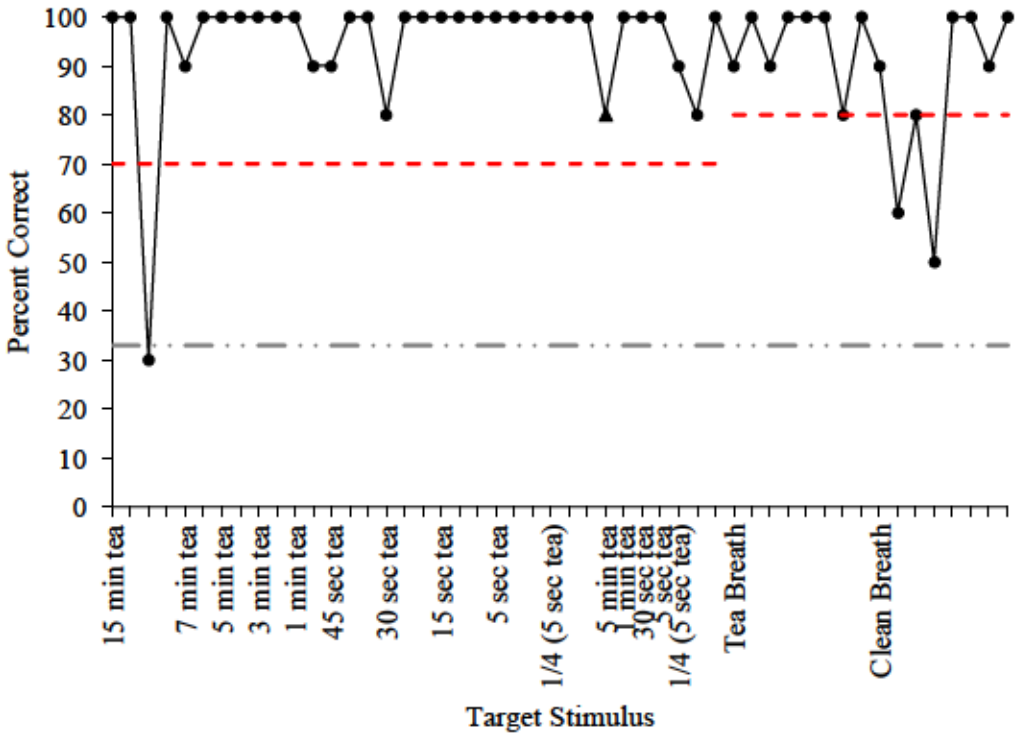

Figure 2.5. Mist's performance across the three phases of the training program. The triangle marker indicates the point at which water was introduced in the control stimuli during the liquid tea phase of the training program. The dotted line represents the criterion (70%) based on chance level calculated using a binomial probability with 10 trials, 0.33 probability of success per trial (represented by the intermittent dashed line), and an alpha level of 0.05.

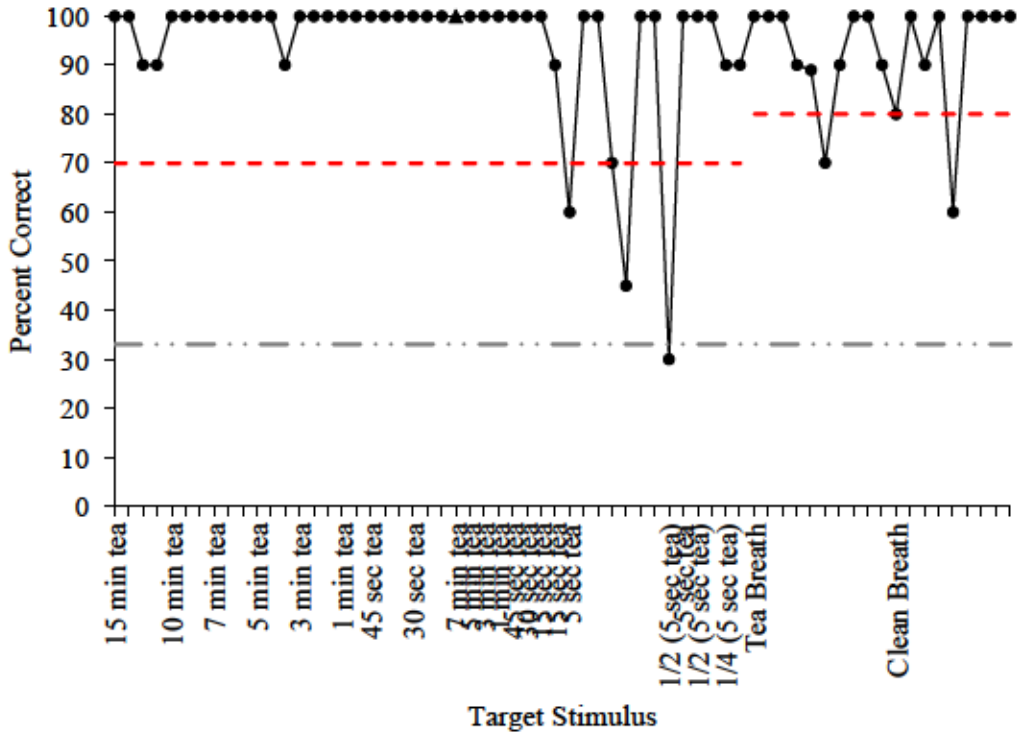

Figure 2.6. Koda's performance across the three phases of the training program. The triangle marker indicates the point at which water was introduced in the control stimuli during the liquid tea phase of the training program. The dotted line represents the criterion (70%) based on chance level calculated using a binomial probability with 10 trials, 0.33 probability of success per trial (represented by the intermittent dashed line), and an alpha level of 0.05.

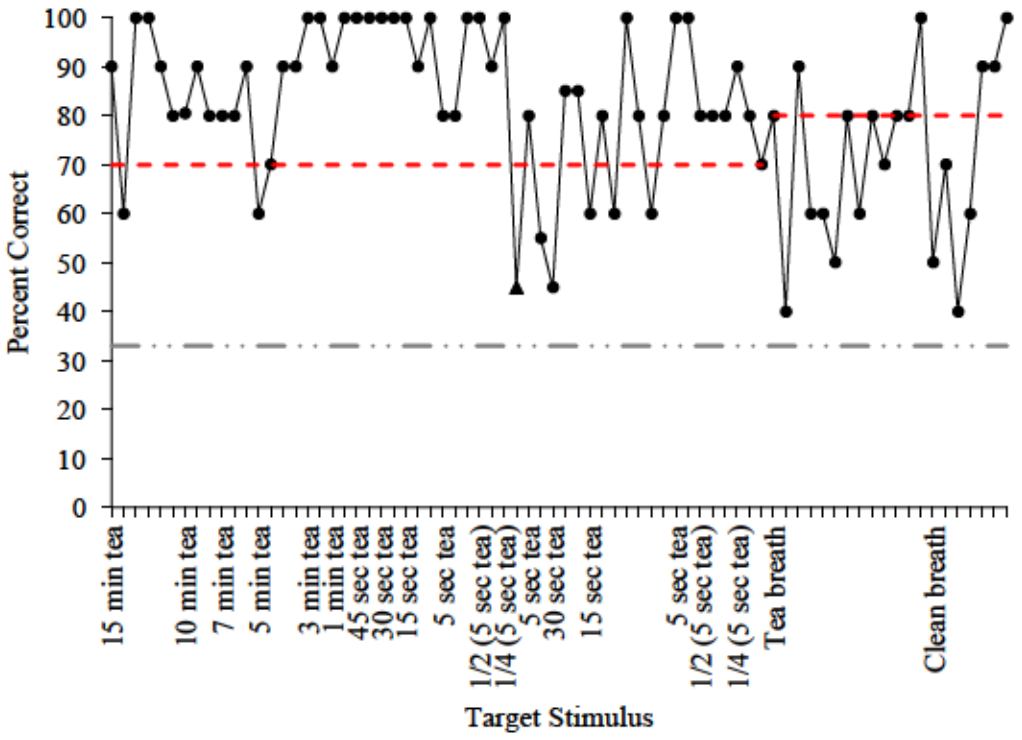

Figure 2.7. Nutella's performance across the three phases of the training program. The triangle marker indicates the point at which water was introduced in the control stimuli during the liquid tea phase of the training program. The dotted line represents the criterion (70%) based on chance level calculated using a binomial probability with 10 trials, 0.33 probability of success per trial (represented by the intermittent dashed line), and an alpha level of 0.05.

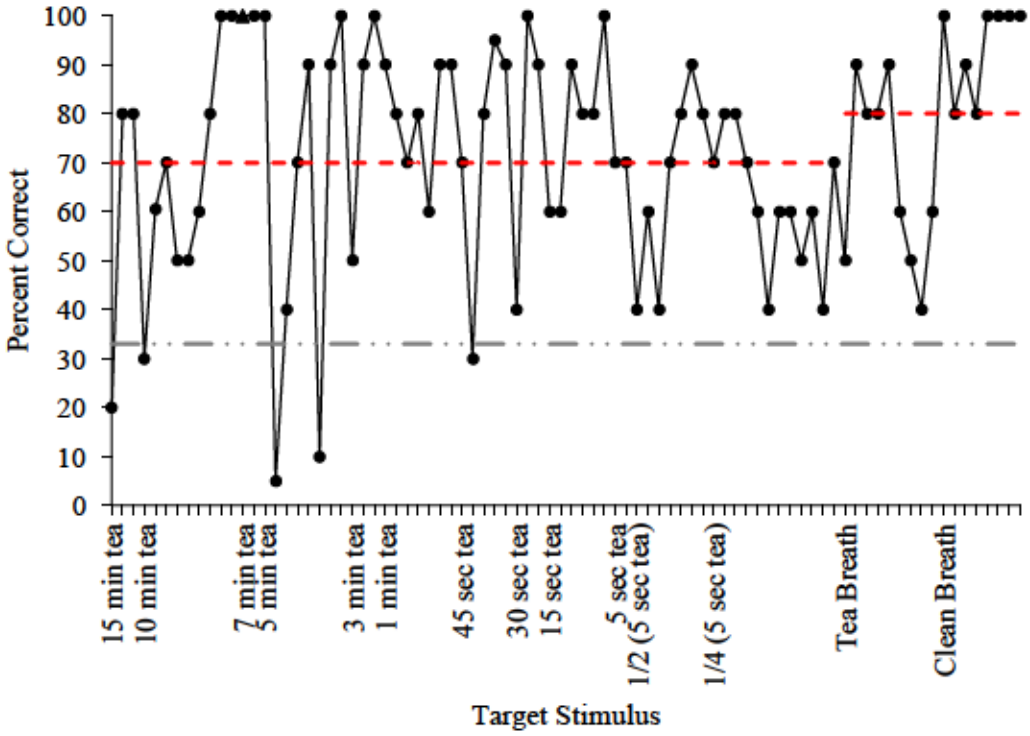

Figure 2.8. Bella's performance across the three phases of the training program. The triangle marker indicates the point at which water was introduced in the control stimuli during the liquid tea phase of the training program. The dotted line represents the criterion (70%) based on chance level calculated using a binomial probability with 10 trials, 0.33 probability of success per trial (represented by the intermittent dashed line), and an alpha level of 0.05.

2.4 Experiment 1

As reported above, four of the five dogs successfully completed the training program (the fifth dog, Charlee, moved away before she could complete the program) and could therefore reliably detect human breath samples. The goal of Experiment 1 was to test the dogs' ability to discriminate between breath samples from three different individuals using a three-stage cued, 3AFC procedure.

2.4.1 Method

2.4.1.2 Dogs. All the dogs that underwent training took part in this study, except for Charlee who moved away before the study began.

2.4.1.3 Stimuli. Breath samples were donated by three different individuals in the lab just prior to a dog arriving for its scheduled work day. Breath samples were collected with the same breath collection tubes described above except here, either end of the tubes were capped with tight fitting rubber caps. When a breath sample donor arrived at the lab, the rubber caps were removed and a breath sample was donated using the same protocol described above. Once complete, the rubber caps were fitted back on the tubes until use.

One of the three breath collection tubes contained two cotton balls (see Figure 2.9). This sample served as the target breath sample that the dog would be asked to identify. The second cotton ball allowed a cue sample to be prepared. While the breath samples were donated, two additional control samples were prepared. Control samples were breath collection tubes containing one cotton ball each, but onto which no breath was exhaled.

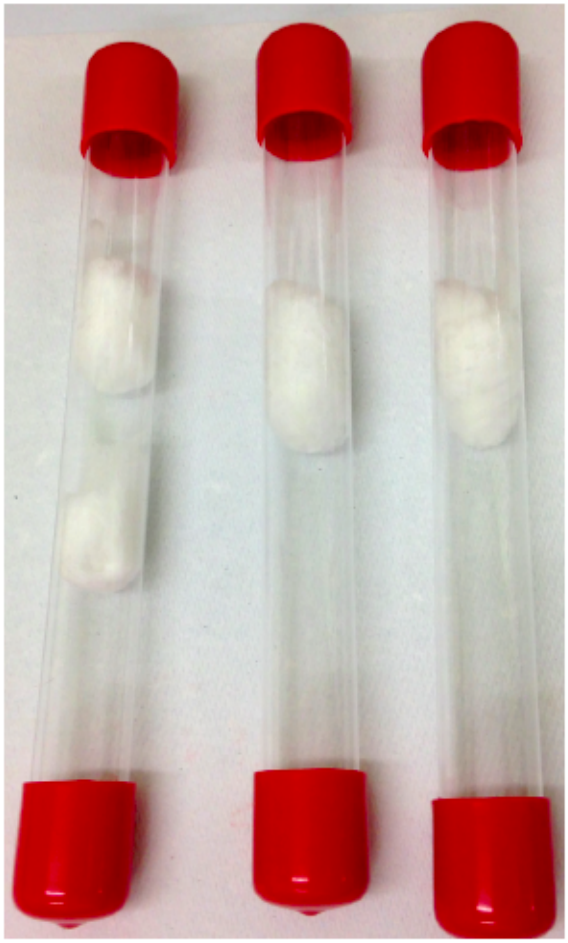

Figure 2.9. Three breath collection tubes with caps. The breath collection tube on the left, which will be used to collect the target breath sample, holds two cotton balls so that after sample donation, one can be removed and used as the sample stimulus.

2.3.1.4 Procedure. Samples were prepared for presentation by removing the tube caps and placing the breath collection tubes inside sample stations. The cue sample was prepared by using tweezers to remove one of the cotton balls from the target breath sample tube and placing it inside a stainless-steel jar with holes in the lid. Samples were presented to the dogs using the cued, 3AFC procedure outlined above, however, in this experiment the stimuli were presented in stages, where distracting stimuli were added to the presentation only if the dog met criterion for the previous stage. In the first stage, the target breath sample was presented with two control samples and the dog was required to complete one session at 80% correct before progressing to the second stage. In the second stage, one of the control samples was removed and a breath sample from a second individual was put in its place. To successfully complete stage two, the dog was required to complete three consecutive sessions at 80% correct or two sessions at 90% or 100% correct. In stage three, the last control sample was removed and the third breath sample from the third different individual was put in its place. To successfully complete stage three, the dog was required to complete three consecutive sessions at 80% correct or two consecutive sessions at 90% to 100% correct. Samples were not used for more than one work day, therefore if a dog did not complete all three stages within its scheduled work day, they had to start over at stage one with new samples the following work day. However, if a dog previously demonstrated the ability to discriminate between samples to criterion in stage two, they were only required to complete one session at 80% correct in stage two on subsequent work days before progressing to stage three. It is important to note that, because the amount of time a dog could spend at the lab was dependent upon its owner's schedule, one dog may have taken longer to complete stage three than another

simply because of time constraints, not because of its performance. Once a dog completed stage three, it was considered to have successfully discriminated between breath samples from three different individuals.

2.4.2 Results and Discussion

As illustrated by Figures 2.10 through 2.13, dogs took a varying amount of time to complete stage three of the testing. However, as discussed above, some dogs took longer to reach the criterion not because they had trouble with the task, but simply because they completed fewer work sessions per work day than another dog.

2.4.2.1 Mist. Mist required only two work days to reach the criterion for phase three, demonstrating that Mist could very easily be trained to discriminate between individuals using breath samples. However, observing her performance in Figure 2.10, it is possible that she could have met the criterion in only one work day if more sessions had been conducted on that day. Mist showed generally consistent performance throughout the training program despite some fluctuations above and below chance with the clean breath stimuli (which she recovered from quickly and then showed more typical high performance). Therefore, her performance in the training phase was predictive of her performance in Experiment 1 in that her performance was consistent and above the criterion for all stages of the test.

2.4.2.2 Koda. Koda met the criterion for phase three in only one work day, demonstrating that he could easily discriminate between breath samples from different individuals (see Figure 2.11). He showed a slight decrease in performance when the second distracting breath sample was introduced at the beginning of phase three, but his performance was maintained at above the criterion throughout all stages of testing.

Koda's performance in the training program showed some fluctuation in the lower saliency tea stimuli and in the tea breath and clean breath phases, but his performance was typically consistently above criterion with the breath stimuli. It could therefore be said that his performance in the training program was predictive of his performance in Experiment 1.

2.4.2.3 Nutella. As seen in Figure 2.12, Nutella showed an initial drop in performance when the first distracting breath sample was introduced in stage two. However, when presented with the target breath sample and one distracting sample on the second work day, her performance decreased only slightly and remained above criterion. The same pattern of performance was observed between stage one and stage two on work day three. This suggests that in the first stage of the experiment, Nutella may have been perceiving the target breath sample as an "odour" without perceiving the details of the VOC profile. Therefore, she may have initially struggled to detect differences between the different individuals' VOCs profiles because she was not detecting the distinguishing details. It appears as though after a few presentations of multiple samples, however, Nutella perceived the differences in the breath sample VOC profiles and was successfully able to discriminate between them. When the second distracting stimulus was introduced on work day three, Nutella initially showed a high degree of discriminability, as her performance was above the criterion. Her performance on the second session of stage three, however, fell to below the criterion. We struggle to explain why her performance with the same stimuli could have dropped so dramatically within one work day, but on work day three, Nutella completed stage three with no apparent difficulties. Therefore,

Nutella learned to discriminate between breath samples from different individuals with minimal difficulty.

Nutella's Performance in the training program was somewhat predictive of performance in Experiment 1. As discussed in the results of the training program, Nutella's performance decreased when different stimuli were introduced such as at the beginning of the tea breath phase and the clean breath phase. Likewise, when the new distracting breath samples were introduced here, a similar decrease in performance was observed.

2.4.2.4 Bella. As illustrated by Figure 2.13, on Bella's first work day she completed four sessions at the end of which she had met the criterion for stage two. However, observing Bella's performance in more detail, it is evident that when distracting breath samples were introduced at the beginning of stage two and stage three, she had difficulty discriminating between the stimuli, as evidenced by performance falling to below the criterion at these points. Much like Nutella, this pattern of performance suggests that, at first, Bella did not perceive the details of the VOC profiles in the breath samples. Therefore, although the introduction of new stimuli initially affected Bella's performance, she was ultimately successful in discriminating between the three breath stimuli at a high level of performance.

Bella's inconsistent performance throughout the training program may have been predictive of her performance in the current experiment, in that it was also inconsistent. However, recall that in the training program, the introduction of the water control and the point at which the tea saliencies decreased did not affect Bella's performance, whereas here in Experiment 1, decreases in Bella's performance are related

to the introduction of a distracting stimulus. Therefore, the training program did not necessarily predict a decrease in performance when the study parameters became more difficult.

2.5 Experiment 2

The results of Experiment 1 showed that all four dogs could successfully discriminate between breath samples from three different individuals at a high level of performance. The goal of Experiment 2 was to examine whether the dogs could detect more minute differences in the breath samples by testing their ability to discriminate between breath samples donated from one individual at three different times of the day.

2.5.1 Method

2.5.1.1 Stimuli. Breath samples were donated by one individual at three different times of the day. The day before the samples were to be presented to a dog, the breath sample donor came to the lab and was given a bag containing five capped breath collection tubes: three for the collection of breath samples, and two control tubes onto which no breath would be exhaled. The breath samples were collected using the same breath collection tubes and same donation procedure described above. That day that the donor individual received the tubes, he/she prepared an evening breath sample (donated just before the individual went to bed, but before brushing his/her teeth). The next day (the same day the samples would be presented to a dog), the breath sample donor collected a morning sample (donated first thing in the morning before eating or brushing one's teeth) which would serve as the target sample, therefore, this breath collection tube contained two cotton balls (one served as the cue). Finally, the individual donated one

midday sample just before he/she ate lunch. The donor then returned the tubes to the lab before the dog arrived to work that afternoon.

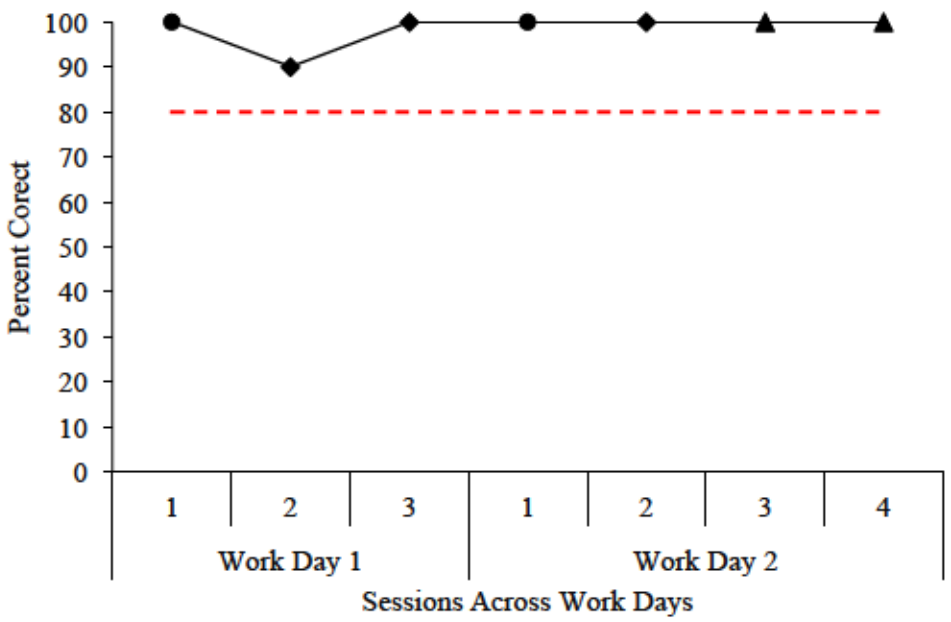

Figure 2.10. Mist's performance across stages of testing, discriminating between breath samples from three different people. The circle marker indicates stage one sessions where only the target breath sample and two blank control samples were presented. The diamond marker indicates stage two sessions where the target breath sample, one distracting breath sample from a second person, and one blank control sample was presented. The triangle marker indicates stage three sessions where the target breath sample, and two distracting breath samples from two different people were presented. The dotted line illustrates the criterion (80%).

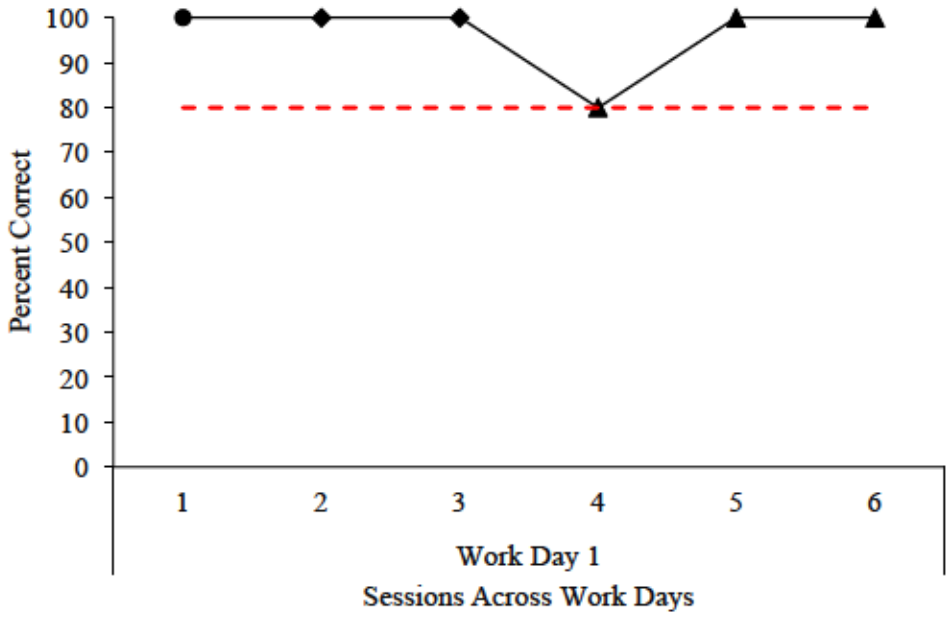

Figure 2.11. Koda's performance across stages of testing, discriminating between breath samples from three different people. The circle marker indicates stage one sessions where only the target breath sample and two blank control samples were presented. The diamond marker indicates stage two sessions where the target breath sample, one distracting breath sample from a second person, and one blank control sample was presented. The triangle marker indicates stage three sessions where the target breath sample, and two distracting breath samples from two different people were presented. The dotted line illustrates the criterion (80%).

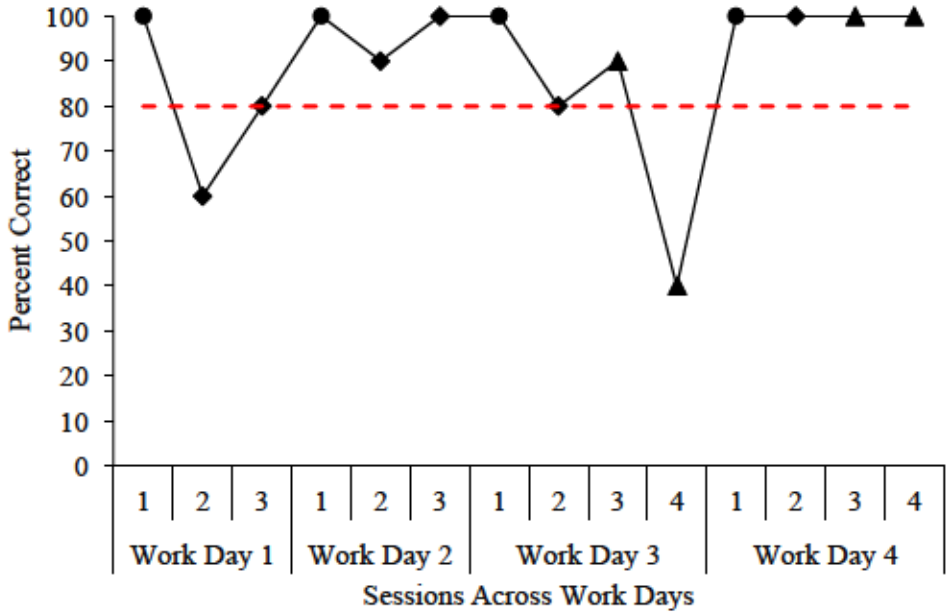

Figure 2.12. Nutella's performance across stages of testing, discriminating between breath samples from three different people. The circle marker indicates stage one sessions where only the target breath sample and two blank control samples were presented. The diamond marker indicates stage two sessions where the target breath sample, one distracting breath sample from a second person, and one blank control sample was presented. The triangle marker indicates stage three sessions where the target breath sample, and two distracting breath samples from two different people were presented. The dotted line illustrates the criterion (80%).

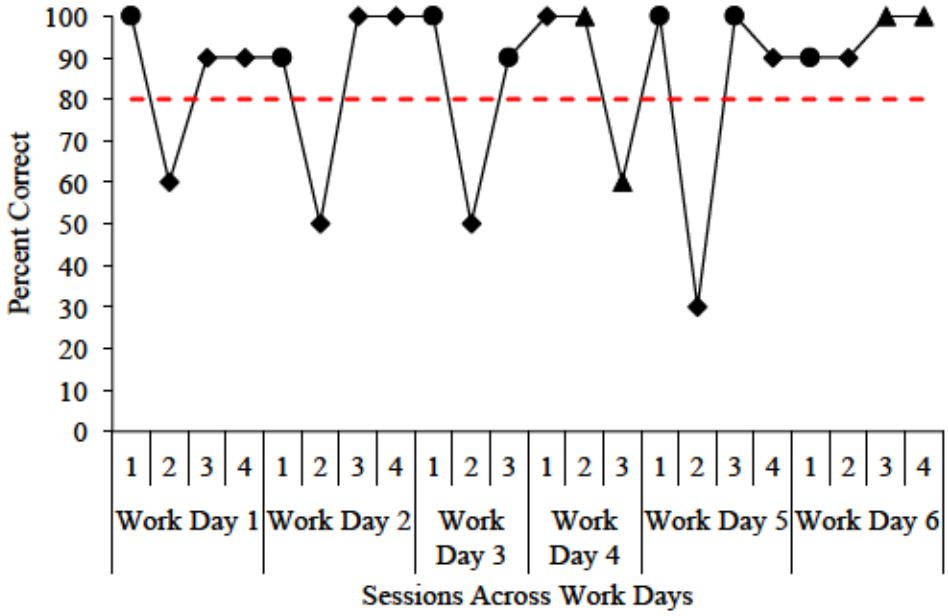

Figure 2.13. Bella's performance across stages of testing, discriminating between breath samples from three different people. The circle marker indicates stage one sessions where only the target breath sample and two blank control samples were presented. The diamond marker indicates stage two sessions where the target breath sample, one distracting breath sample from a second person, and one blank control sample was presented. The triangle marker indicates stage three sessions where the target breath sample, and two distracting breath samples from two different people were presented. The dotted line illustrates the criterion (80%).

2.5.1.2 Procedure. Samples were presented to the dogs using the cued, 3AFC procedure outlined above. Here, the samples were presented to the dogs in three stages similar to those used in the previous study, and the same criterion was required at each stage. In the first stage, the morning breath sample was presented with two control samples. In the second stage, one of the control samples was removed and the same individual's evening breath sample was put in its place. In stage three, the last control sample was removed and the midday breath sample from the same individual was put in its place. Once a dog completed stage three, it was considered to have successfully discriminated between breath samples from three different individuals. As with the previous study, no breath samples were used for more than one work day. Therefore, if a dog was unable to complete all three phases of testing within a work day, it was required to start back at stage one of testing on the next work day.

2.5.2 Results and Discussion

2.5.2.1 Mist. Mist met the criterion for stage three on the first day of testing. As seen in Figure 2.14, her performance was consistently high and never fell below 90% correct. Therefore, Mist presented no difficulties in discriminating between one individual's breath samples at three different times of the day. As discussed above, Mist demonstrated consistently high performance across all training and testing phases, therefore, her performance throughout the training program was predictive of her potential to work with human breath samples.

2.5.2.2 Koda. As illustrated by Figure 2.15, Koda's performance dropped well below the criterion at stage two when he was first required to detect the morning breath

sample and ignore the evening breath sample. These results suggest that Koda may have struggled to detect differences in the VOC content of the samples. This is to be expected, because, although an individual's breath VOC profile may fluctuate throughout the day due to metabolic changes or exogenous sources (Amann et al., 2014), the three samples likely have a large number of overlapping VOCs. Koda's performance on the second work day, however, was maintained at 100% correct through all stages of testing. But, as discussed in the method section, breath samples were not used more than once. Therefore, Koda was presented with a set of breath samples from one individual on the first work day, and a set of breath samples from a different individual on the second work day. We cannot eliminate the possibility that the individual that donated the breath samples used on the second day did not consume any food, drink, or medications throughout the day that may have made their three different breath samples particularly distinct from one another.

Koda's performance across the training program showed some inconsistent performance at lower saliency tea stimuli and with the breath stimuli and therefore could have been indicative of some difficulties with lower saliency stimuli. This could be a reason for his observed difficulty in discriminating the morning and evening breath samples in stage two on the first work day. However, as discussed above, it is possible that the breath samples used on the first work day were qualitatively different than those used on the second work day.

2.5.2.3 Nutella. All illustrated by Figure 2.16, Nutella maintained a high level of performance at throughout all stages of this experiment, suggesting that she could discriminate between the VOC profiles of the breath samples from the same individual

with ease. Nutella's performance in the training program was not indicative of her potential performance, since she demonstrated much more variability during all three training phases.

2.5.2.4 Bella. As illustrated in Figure 2.17, Bella showed no difficulty discriminating between the morning and evening breath samples in stage two, as her performance was maintained at 100%. When presented with the midday breath sample in stage three, however, her performance, although initially high, dropped to below criterion. On the second work day, however, she progressed through stages two and three maintaining performance of 100% correct throughout all sessions, indicating that she had little difficulty discriminating between the breath VOCs from one person at three different times of the day. Bella's performance throughout the training program was not predictive of her performance here since she demonstrated great variability in her performance throughout training, but showed much more reliable performance in this experiment.

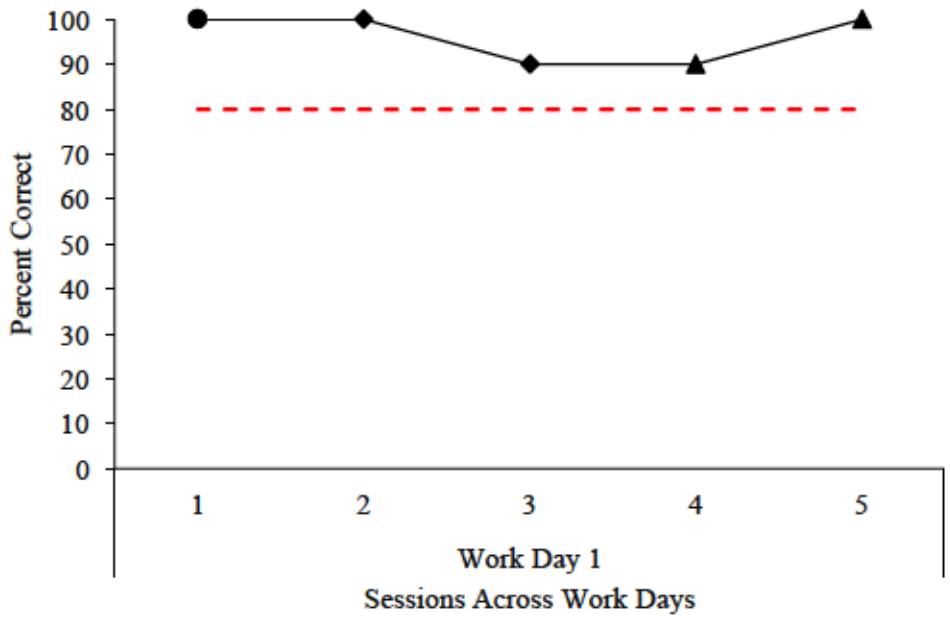

Figure 2.14. Mist's performance across stages of testing, discriminating between breath samples donated by one person at three different times of the day. The circle marker indicates stage one sessions where the target morning breath sample is presented with two blank control samples. The diamond marker indicates stage two sessions where the target morning breath sample, the distracting evening breath sample, and one blank control are presented. The triangle marker indicates stage three sessions were the target morning breath sample, and the distracting evening and midday samples are presented. The dotted line illustrates the criterion (80%).

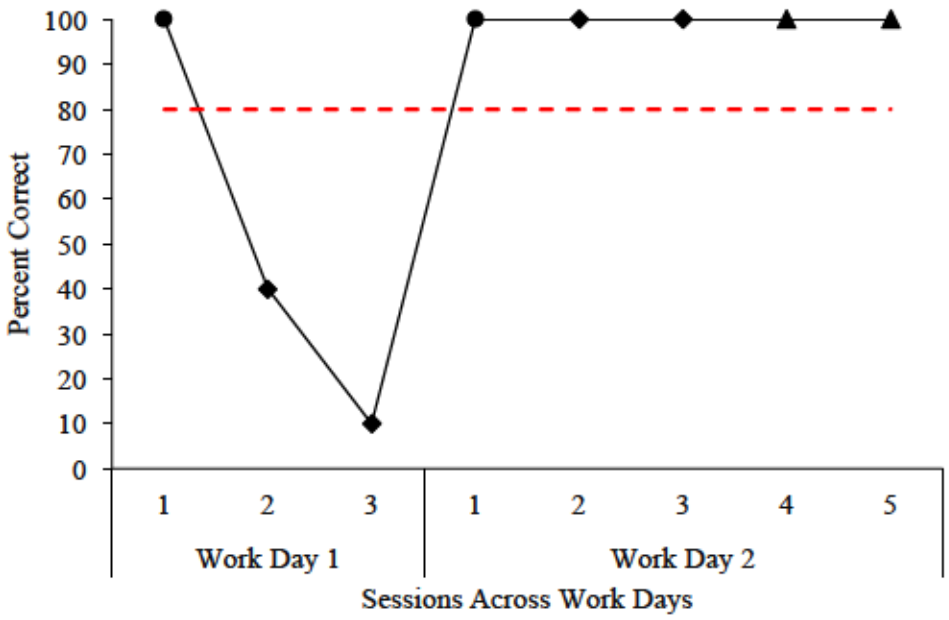

Figure 2.15. Koda's performance across stages of testing, discriminating between breath samples donated by one person at three different times of the day. The circle marker indicates stage one sessions where the target morning breath sample is presented with two blank control samples. The diamond marker indicates stage two sessions where the target morning breath sample, the distracting evening breath sample, and one blank control are presented. The triangle marker indicates stage three sessions were the target morning breath sample, and the distracting evening and midday samples are presented. The dotted line illustrates the criterion (80%).

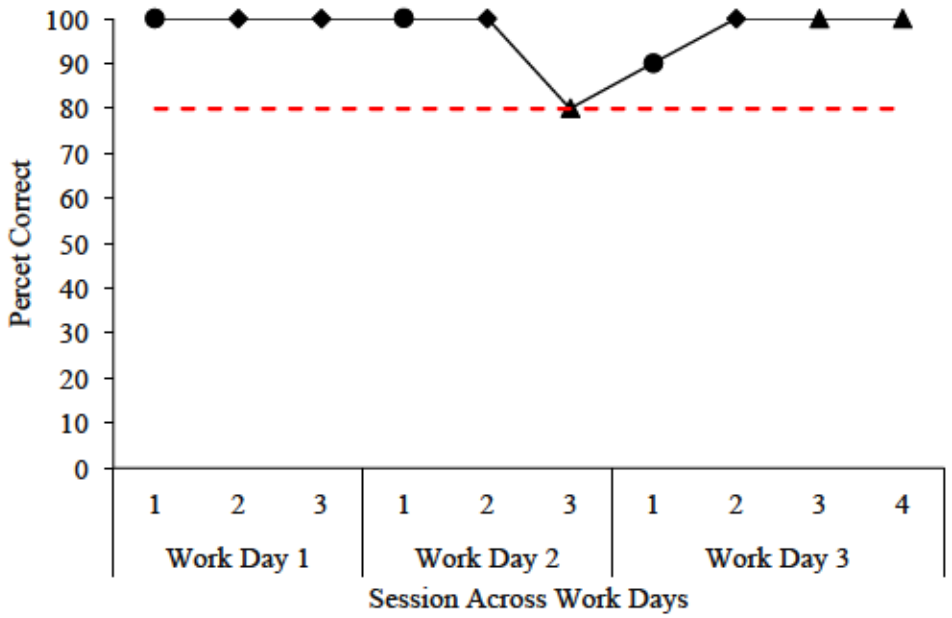

Figure 2.16. Nutella's performance across stages of testing, discriminating between breath samples donated by one person at three different times of the day. The circle marker indicates stage one sessions where the target morning breath sample is presented with two blank control samples. The diamond marker indicates stage two sessions where the target morning breath sample, the distracting evening breath sample, and one blank control are presented. The triangle marker indicates stage three sessions were the target morning breath sample, and the distracting evening and midday samples are presented. The dotted line illustrates the criterion (80%).

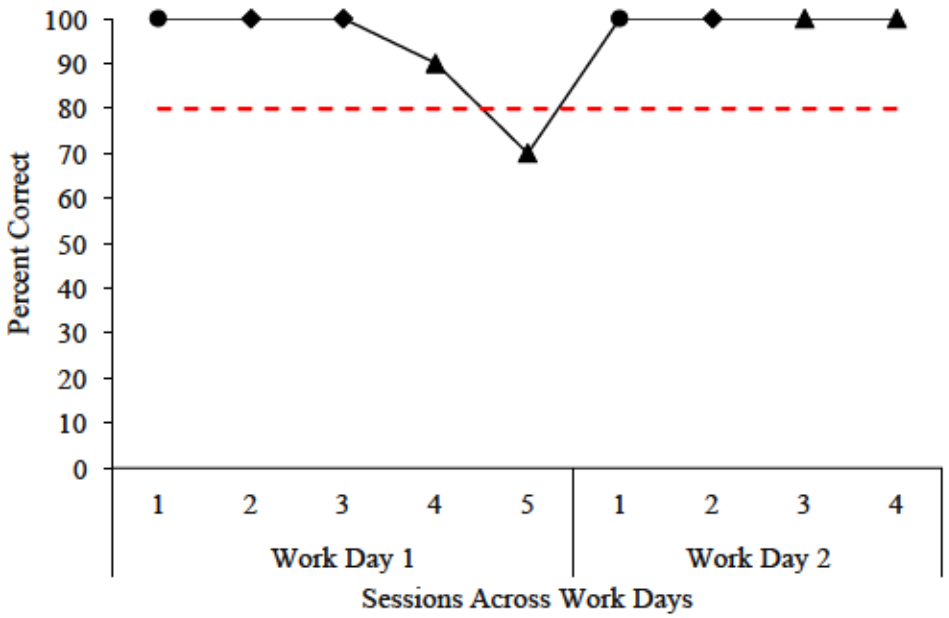

Figure 2.17. Bella's performance across stages of testing, discriminating between breath samples donated by one person at three different times of the day. The circle marker indicates stage one sessions where the target morning breath sample is presented with two blank control samples. The diamond marker indicates stage two sessions where the target morning breath sample, the distracting evening breath sample, and one blank control are presented. The triangle marker indicates stage three sessions were the target morning breath sample, and the distracting evening and midday samples are presented. The dotted line illustrates the criterion (80%).

2.6 General Discussion

The goal of the current study was to test the utility of a new training program designed to train dogs to detect and discriminate between human breath samples. The training program consisted of three major phases in which different stimuli were presented to the dogs using a cued, 3AFC procedure. In phase one, the dogs detected decreasing concentrations of orange pekoe tea on cotton against control stimuli. In phase two, the dogs detected "tea breath" stimuli, and in phase three the dogs detected clean breath. All five dogs trained successfully completed the training program. Subsequently, four dogs' ability to discriminate between breath samples was tested in two different experiments. In the first experiment, a three-stage cued, 3AFC procedure was used to test dogs' ability to discriminate between breath samples from different individuals. In the second experiment, the same procedure was used to test dogs' ability to discriminate between breath samples from one individual donated at three different times of the day. The results showed that for both experiments, all four dogs met our established criterion in less than six work days.

The dogs' performance throughout the training program appeared to predict their performance in Experiment 1 but not necessarily in Experiment 2. Results show that general trends in dogs' performance observed in the training program were maintained in Experiment 1. For example, Mist and Koda showed generally consistent performance across both, whereas Nutella was challenged when new stimuli were presented in the training phases and the testing stages. The predictive potential of Bella's performance in the training program is less clear since she showed inconsistent performance throughout the training program unrelated to decreases in stimuli saliency, but showed systematic

difficulties when distracting stimuli were introduced in stages two and three in Experiment 1. All of the dogs' performance in Experiment 2 were consistently high, with the exception of Koda, who showed poor performance on the first work day, the potential reasons for which are discussed above. Therefore, other than for Mist, whose performance was consistently high across training and testing, the dogs' performance throughout phases of the training program was not predictive of their performance in Experiment 2. As discussed above, it is assumed that the samples obtained from one person at three different times of the day would have different VOC profiles, but there would be much more overlap than the samples donated from three different people. Therefore, it is likely that the high level of performance observed in Experiment 2 was the result of practice effects, as the dogs had smelled many breath samples at that point.

Taken together, the results of the training program and Experiments 1 and 2 suggest that the training and testing procedures were successful in training the dogs to detect and discriminate between human breath samples. Furthermore, the level of performance observed in Experiments 1 and 2 was much higher than the performance of the dogs reported by Jezierski et al. (2010). The high level of accuracy demonstrated by the dogs in these studies is likely attributable to a few important factors. First, contrary to tradition, in both training and testing we did not use the five to seven stimulus line-up procedure, but rather we used what could be considered a shortened line-up of only three stimuli. While this is potentially an easier task, we argue that presenting a dog with more stimuli is unnecessary. As discussed by Gadbois and Reeve (2014), although a correct indication in a seven-stimulus line-up is less likely due to chance than a correct indication in a three-stimulus line-up, the longer line-up is more a test of a dogs' working memory

and resistance to sensory and perceptual interference than it is a test of accurate detection and discrimination. If the goal of a study is to test dogs' ability to detect and discriminate stimuli, this can be easily achieved with a smaller array of stimuli. If the ability of the dog to detect or discriminate in high distraction situations is sought, then other techniques can be applied (Gadbois & Reeve, 2014). This proposition is confirmed by the data of the current study, since the performance of our dogs is higher than what was reported by Jezierski et al. (2010).

Second, in the training phases we applied a criterion that required that the dogs' performance be above chance on multiple *consecutive* sessions (with a few exceptions) before they progressed onto subsequent training phases. If a dog met this criterion, we could be sure that they could reliably perceive the stimuli and, were motivated enough to persist with a repetitive task. In Jezierski et al. (2010), dogs progressed through training phases if their performance was above chance on 50 out of 100 trials. However, Jezierski et al. (2010) did not observe improvement in detection accuracy within a training phase. This could serve as an indication that the dog did not grasp the task fully and therefore struggled when presented with a subsequent more difficult task. This is likely, as the results showed that, overall, the dogs' performance decreased with each subsequent training phase (Jezierski et al., 2010). We agree with Schoon (1996), in suggesting that a "better than chance" criteria is not good enough if dogs are to be used in forensic and biomedical detection applications (p. 258). Therefore, we suggest that more strict criteria be adopted.

Finally, we utilized a training program that gradually decreased the saliency of the target stimuli and then gradually bridged the gap between the non-biological stimulus

(liquid tea) and the biological stimulus (tea breath and clean breath samples). In doing so, we ensured that dogs were detecting the target odour, and that they learned over time that low saliency odours, including breath that they had smelled their entire lives, were important and could signal reward.

One important limitation of the current study is that we did not include "zero trials" in which no stimulus matches the sample and the dog must therefore refrain from indicating a stimulus. However, the goal of our research was to examine basic perceptual olfactory processes in dogs and we at no point intended to have our dogs deployed in real world applications, therefore zero trials were not necessary. But, as discussed by Schoon (1996) and Jezierski et al., (2010), in real world applications of olfactory detection dogs, it is imperative that a dog be capable of signaling that there is not a match, to prevent false accusations of innocent suspects or to prevent unnecessary medical procedures. We recognize this as a limitation of the current design and encourage others to implement them in their own training programs.

2.7 Acknowledgements

This research was supported by an IWK Health Centre Category A Research Grant, and an NSERC PGS-D to Catherine Reeve.

We would like to acknowledge Dr. Elizabeth McLaughlin, Dr. Beth Cummings, Dr. Vin LoLordo, and Dr. Ray Klein for their consultation and feedback. We would further like to acknowledge the dog owners who reliably brought their dogs to our lab each week, without whom this study would not have been possible. Finally, we thank the many student volunteers who contributed their time.

2.8 References

Amann, A., de Lacy Costello, B., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., ... & Risby, T. (2014). The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *Journal of Breath Research*, *8*, 034001. doi:10.1088/1752-7155/8/3/034001

Amundsen, T., Sundstrøm, S., Buvik, T., Gederaas, O. A., & Haaverstad, R. (2014). Can dogs smell lung cancer? First study using exhaled breath and urine screening in unselected patients with suspected lung cancer. *Acta Oncologica*, *53*, 307-315. doi: 10.3109/0284186X.2013.819996

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*, e7396. doi: 10.1136/bmj.e7396

Brisbin, I. L., & Austad, S. N. (1991). Testing the individual odour theory of canine olfaction. *Animal Behaviour*, *42*, 63-69. doi: https://doi.org/10.1016/S0003-3472(05)80606-2

Buszewski, B., Rudnicka, J., Ligor, T., Walczak, M., Jezierski, T., & Amann, A. (2012). Analytical and unconventional methods of cancer detection using odor. *Trends in Analytical Chemistry*, *38*, 1-12. doi:http://dx.doi.org/10.1016/j.trac.2012.03.019

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. doi: 10.1016/j.eururo.2010.10.006

Craven, B. A., Paterson, E. G., & Settles, G. S. (2009). The fluid dynamics of canine olfaction: unique nasal airflow patterns as an explanation of macrosmia. *Journal of The Royal Society Interface*, *7*, 933-943. doi:10.1098/rsif.2009.0490

Curran, A. M., Rabin, S. I., Prada, P. A., & Furton, K. G. (2005). Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *Journal of Chemical Ecology*, *31*, 1607-1619. doi: 10.1007/s10886-005-5801-4

Dehlinger, K., Tarnowski, K., House, J. L., Los, E., Hanavan, K., Bustamante, B., ... & Ward, W. K. (2013). Can trained dogs detect a hypoglycemic scent in patients with type 1 diabetes?. *Diabetes Care*, *36*(7), e98-e99. doi: 10.2337/dc12-2342

Ehmann, R., Boedeker, E., Friedrich, U., Sagert, J., Dippon, J., Friedel, G., & Walles, T. (2012). Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *European Respiratory Journal*, *39*, 669-676. doi: 10.1183/09031936.00051711

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 urology*, *14*, 22. doi: http://www.biomedcentral.com/1471-2490/14/22

Gadbois S., Reeve C. (2014). Canine olfaction: Scent, sign, and situation. In A. Horowitz, A. (Ed.), *Domestic dog cognition and behavior (*pp 3-29). New York, NY: Springer.

Gordon, R. T., Schatz, C. B., Myers, L. J., Kosty, M., Gonczy, C., Kroener, J., ... & Arthur, N. (2008). The use of canines in the detection of human cancers. *The Journal of Alternative and Complementary Medicine*, *14*, 61-67. doi: 10.1089/acm.2006.6408

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. doi: 10.1007/s13300-015-0135-x

Helton, W. S. (2009). Canine ergonomics: Introduction to the new science of working dogs. In W. S. Helton (Ed.), *Canine ergonomics. The science of working dogs* (pp. 1–16). Boca Raton, FL: Taylor and Francis Group.

Horvath, G., Andersson, H., & Nemes, S. (2013). Cancer odor in the blood of ovarian cancer patients: a retrospective study of detection by dogs during treatment, 3 and 6 months afterward. *BMC cancer*, *13*, 396. doi: http://www.biomedcentral.com/1471-2407/13/396

Horvath, G., Andersson, H., & Paulsson, G. (2010). Characteristic odour in the blood reveals ovarian carcinoma. *BMC cancer*, *10*, 643. doi: http://www.biomedcentral.com/1471-2407/10/643

Horvath, G., Järverud, G. A. K., Järverud, S., & Horváth, I. (2008). Human ovarian carcinomas detected by specific odor. *Integrative cancer therapies*, *7*, 76-80. doi: 10.1177/1534735408319058

Jezierski, T., Gorecka-Bruzda, A., Walczak, M., Swiergiel, A. H., Chruszczewski, M. H., & Pearson, B. L. (2010). Operant conditioning of dogs (*Canis familiaris*) for identification of humans using scent lineup. *Animal Science Papers and Reports*, *28*, 81-93.

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*, 027001. doi:10.1088/1752-7155/9/2/027001

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. doi: http://dx.doi.org/10.1016/j.applanim.2017.01.001

Koivusalo, M., Vermeiren, C., Yuen, J., Reeve, C., Gadbois, S., & Katz, K. (2017). Canine scent detection as a tool to distinguish meticillin-resistant Staphylococcus aureus. *Journal of Hospital Infection*, *96*, 93-95. doi: http://dx.doi.org/10.1016/j.jhin.2017.03.005

Kusano, M., Mendez, E., & Furton, K. G. (2011). Development of headspace SPME method for analysis of volatile organic compounds present in human biological specimens. *Analytical and bioanalytical chemistry*, *400*, 1817. doi: 10.1007/s00216-011-4950-2

Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, *52*, 415-419.

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. doi: 10.1177/1534735405285096

Phillips, M., Herrera, J., Krishnan, S., Zain, M., Greenberg, J., & Cataneo, R. N. (1999). Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications*, *729*, 75-88.

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. doi:10.1016/j.applanim.2004.04.008

Rudnicka, J., Walczak, M., Kowalkowski, T., Jezierski, T., & Buszewski, B. (2014). Determination of volatile organic compounds as potential markers of lung cancer by gas chromatography–mass spectrometry versus trained dogs. *Sensors and Actuators B: Chemical*, *202*, 615-621. doi: http://dx.doi.org/10.1016/j.snb.2014.06.006

Schmidt, K., & Podmore, I. (2015). Current challenges in volatile organic compounds analysis as potential biomarkers of cancer. *Journal of Biomarkers*, *2015*, 1-16. doi: http://dx.doi.org/10.1155/2015/981458

Schoon, G. A. A. (1996). Scent identification lineups by dogs (Canis familiaris): experimental design and forensic application. *Applied Animal Behaviour Science*, *49*, 257-267.

Schoon, G. A. A. (2005). The effect of the ageing of crime scene objects on the results of scent identification line-ups using trained dogs. *Forensic Science International*, *147*, 43-47. doi:10.1016/j.forsciint.2004.04.080

Schoon, G.A. & Haak, R. (2002). *K9 suspect discrimination: Training and practicing scent identification line- ups*. Alberta, Canada: Detselig Enterprises.

Settle, R. H., Sommerville, B. A., McCormick, J., & Broom, D. M. (1994). Human scent matching using specially trained dogs. *Animal Behaviour*, *48*, 1443-1448.

Shirasu, M., & Touhara, K. (2011). The scent of disease: volatile organic compounds of the human body related to disease and disorder. *Journal of Biochemistry*, *150*, 257-266. doi: 10.1093/jb/mvr090

Sonoda, H., Kohnoe, S., Yamazato, T., Satoh, Y., Morizono, G., Shikata, K., ... & Inoue, F. (2011). Colorectal cancer screening with odour material by canine scent detection. *Gut*, *60,* 814-819. doi:10.1136/gut.2010.218305

Willis, C. M., Church, S. M., Guest, C. M., Cook, W. A., McCarthy, N., Bransbury, A. J., ... & Church, J. C. (2004). Olfactory detection of human bladder cancer by dogs: proof of principle study. *BMJ*, *329*, 712 – 714.

CHAPTER 3: ASSESSING INDIVIDUAL PERFORMANCE AND MAINTAINING BREATH SAMPLE INTEGRITY IN BIOMEDICAL DETECTION DOGS

The manuscript prepared for this study is presented below. Catherine Reeve, under the supervision of Dr. Simon Gadbois, was responsible for devising the research questions and seeking out Dr. Peter Wentzell in Dalhousie's Chemistry department to help develop a protocol to prepare the materials. With the help of Dr. Wentzell and Bjorn Wielens, Catherine developed a method for coating cottons balls in silicone oil. Catherine then designed two experiments to test whether the use of silicone-coated cottons balls improved the detectability of breath samples for sniffer dogs. With the support of her co-authors, Catherine was the lead on dog training and testing. Catherine wrote the initial draft of the manuscript, and received and incorporated feedback from her co-authors and committee members. The manuscript was submitted to a special issue of the journal Behavioural Processes on animal olfaction in May of 2017 and is in the revise and resubmit stage. The full reference for this manuscript is:

The full reference for this manuscript is:

Reeve, C., Wentzell, P., Wielens, B., Jones, C., Stehouwer, K., & Gadbois, S. (2017). *Assessing individual performance and maintaining breath sample integrity in biomedical detection dogs*. Manuscript submitted for publication to Behavioural Processes.

3.1 Abstract

In empirical tests of biomedical detection dogs, exhaled breath samples are often used because breath contains volatile organic compounds that can signal metabolic states, infection, or disease. However, in studies that present dogs with breath samples, results show a notable degree of variability in dogs' accuracy both between and within studies. Differing protocols for the collection and storage of exhaled breath samples may contribute to this observed variability. The goal of the current study was therefore to test whether there was a difference in the detectability of breath samples collected using silicone-coated versus uncoated cotton balls. This was tested in two experiments. In Experiment 1, breath samples were prepared using both silicone-coated and uncoated cotton balls, which were then left exposed to the surrounding air. Four dogs' detection of the samples was tested using a cued, three alternative forced choice (3AFC) procedure at regular intervals up to two hours after the samples were prepared. The results of Experiment 1 showed that the dogs' performance was above chance and there was no significant difference in the dogs' detection of the breath samples across conditions. In the second experiment, a series of breath samples were prepared and stored for one, two, three, and four week periods. The same four dogs' ability to detect the breath samples was tested each week using the same cued, 3AFC procedure. The results of Experiment 2 showed that when silicone-coated cotton balls were used, all four dogs could detect the breath samples at above chance levels after the samples were stored for three weeks, and three dogs could detect the samples that were stored for four weeks. When the dogs were tested on their ability to detect the breath samples prepared using uncoated cotton, two dogs' performance fell to below chance levels at one week of storage time, while the

other two dogs could detect the breath samples at above chance levels after the samples were stored for four weeks. Taken together, the results of the two experiments illustrate that silicone-coated cotton balls do not improve detectability of breath samples within two hours, but can greatly improve the detectability of breath samples stored over longer periods of time. Since the use of silicone-coated cotton balls only improved the detectability of the breath samples for two of the four dogs, these results highlight the importance of examining individual differences in dogs' performance. Furthermore, we argue that, given the inherent differences in olfactory ability across dogs, widespread use of silicone-coated cotton balls for the collection of breath samples would increase the pool of testable dogs for biomedical detection studies and would decrease the degree of variability both within and between studies.

3.2 Introduction

There is a great need for simple, non-invasive screening and diagnostic techniques to effectively prevent disease and disease complications. The effective management and treatment of disease is dependent upon early diagnosis. For example, throughout the 20th century, deaths due to most types of cancer have declined as a result of emphasis on early detection (DeSantis, Ma, Bryan, & Jemal, 2014; Siegel, DeSantis, & Jemal, 2014; Siegel, Miller, & Jemal, 2015). However, lung and pancreatic cancer do not show increased survivability because diagnoses are often made at late stages of the disease (Siegel et al., 2016). Current screening and diagnostic tools for disease often involve subjecting patients to painful and invasive procedures such as biopsies, laparoscopies, and blood tests (Amann & Smith, 2013; Burak & Liang, 1987; Jezierski, Walczak, Ligor, Rudnicka, & Buszewski, 2015; Wilson, 2015). The invasive nature of these procedures may prevent some individuals from seeking out screening tests because of fear of the procedures themselves (Burack & Liang, 1987) thereby leading to late diagnoses. Furthermore, many diagnostic tools involve the use of equipment that requires specially trained individuals to operate, analyze, and interpret the results. Therefore, these procedures are not only invasive, but can also be very expensive.

The analysis of exhaled breath has been proposed as a simple and non-invasive alternative to current diagnostic tools (Schmidt & Podmore, 2015). Breath serves as a promising channel for diagnostic purposes because cells emit compounds that, when dissolved in the blood, become volatilized and exhaled in the breath during pulmonary circulation (Amann et al., 2014) in concentrations of parts per billion (nmol/mol) and parts per trillion (pmol/mol) (Schmidt and Podmore, 2015). These volatile organic

compounds (VOCs) therefore provide a window to the metabolic processes of the body (Amann et al., 2014). Donation of a breath sample typically involves an individual exhaling into a breath collection bag, or breathing onto an absorbent material that is then contained for later analysis. This approach is less invasive than traditional diagnostic tools, permits easy, repeat donations of breath samples for most individuals, allows sampling in the hospital or at home, and the process is very inexpensive (Solga & Risby, 2013).

Currently, the most common tools for analyzing the VOC content of exhaled breath samples involve specialized extraction techniques such as solid-phase microextraction (SPME), followed by analytical techniques such as gas chromatography-mass spectrometry (GC-MS) or the use of specialized VOC sensors (Buszewski et al., 2012b; Sun, Shao, & Wang, 2016). Using GC-MS, Phillips et al. (1999) examined the VOCs in breath samples from 50 "normal" individuals. In total, Phillips et al. (1999) found over 3400 different VOCs across the participants, with individual samples averaging 204.2 VOCs. Although individuals differ greatly in the number and type of VOCs emitted in their breath, specific diseases and physiological conditions may present specific VOC profiles (Phillips et al., 1999; Wilson, 2015). The analysis of disease-specific VOCs in exhaled breath has been proposed as a method for the detection of a wide range of physiological conditions (Schmidt & Podmore, 2015), including cancers (Balseiro & Correia, 2006; Buszewski et al., 2012b; Rudnicka, Walczak, Kowalkowski, Jezierski, & Buszewski, 2014; Sun et al., 2016; Szulejko et al., 2010), lung inflammation and disease (Corradi & Mutti, 2013), liver function (Modak, 2013) and diabetic hypoglycemia (Minh, Blake, & Galassetti, 2012; Neupane et al., 2016; Smith, Španěl, Fryer, Hana, & Ferns,

2011) to name a few. Despite progress developing VOC disease profiles, current technologies are only able to detect VOCs in concentrations of parts per billion (nmol/mol) (Schmidt and Podmore, 2015), and disease conditions are often marked by numerous different VOCs or simply by *changes* in VOCs (Solga & Risby, 2013), making analysis even with the most sensitive technologies difficult (Buszewski et al., 2012b; Wilson, 2015). Moreover, these techniques can be expensive and require highly trained individuals to perform the tests and analyze the results (Ross & Esarik, 2013).

Alternatively, trained domestic dogs have been proposed as a cheaper and more accessible VOC analysis technique. A dog's nose is physiologically perfected to take in and process volatile compounds. When a dog sniffs, airborne molecules are efficiently directed into the nasal cavity where the molecules contact the olfactory epithelium (Buszewski et al., 2012b; Craven, Paterson, & Settles, 2010; Settles, Kester, & Dodson-Breibelbis, 2003). The dog's genome contains 1,094 olfactory receptor genes (Quignon et al., 2005) coding for the olfactory receptors located in the olfactory epithelium. This translates into the ability to detect some odours at 1 part per trillion (ppt, Pearsall & Verbruggen, 1982; Walker et al., 2006). Furthermore, dogs are highly trainable, and using the principles of operant conditioning, can be trained to identify specific odours (Gadbois & Reeve, 2014). Using an olfactometer, Waggoner et al. (1998) tested four dogs' ability to detect a target odour in the presence of an extraneous odours as the concentration of the extraneous odours increased. All of the dogs could successfully detect a target odour in the presence of extraneous odours, and one dog could detect the target odour even when the extraneous odour increased to a concentration 100 times stronger than the target odour. Furthermore, Walker et al. (2006) reported that two dogs

were able to detect n-amyl acetate at parts per trillion; concentrations significantly lower than those detectable by current technologies (Schmidt & Podmore, 2015). These results illustrate the incredible sensitivity of dogs' noses, and their ability to identify specific odours. It further suggests that complex mixtures of many odours, such as would be expected in a breath sample, do not necessarily impede the ability of dogs to detect specific target odours. Applied to biomedical detection, dogs' incredible olfactory abilities combined with their trainability make them promising diagnostic assistants. Empirical studies of dogs' ability to detect disease and physiological states from breath samples present promising results, but a careful examination of the literature shows inconsistencies both within and between studies.

Empirical tests of dogs' efficacy as biomedical detection tools have focused primarily on dogs' ability to detect a variety of cancers, but also on physiological states such as diabetic hypoglycemia. Dogs' performance on these tasks is typically reported using sensitivity and specificity. Sensitivity is the proportion of positives correctly identified as such, usually expressed as a percent and specificity is the proportion of negatives correctly identified as such. Here we will briefly review those studies that present dogs with breath samples specifically.

To the best of our knowledge, McCulloch et al. (2006) have conducted the only study to examine dogs' ability to detect breast cancer from breath samples. McCulloch et al. (2006) obtained breath samples from individuals with biopsy-confirmed breast cancer as well as breath samples from healthy controls, and then tested five dogs' ability to identify a cancerous sample amongst control samples. McCulloch et al. (2006) reported that the dogs could identify cancerous breath samples with 88% sensitivity and 98%

specificity. Likewise, Sonoda et al. (2011) have published the only study to examine dogs' ability to detect colorectal cancer from breath samples. Sonoda et al. (2011) presented one dog with breath samples from individuals with colorectal cancer against breath samples from healthy controls and reported the dog's detection sensitivity was 91% and specificity was 99%. Taken together, these results are impressive and encouraging with respect to the effectiveness of biomedical detection dogs.

To date, there has been a greater focus on the empirical examination of dogs' ability to detect lung cancer, and the results have been more inconsistent. A review of the current studies shows more inconsistencies in the dogs' performance than those reported for breast and colorectal cancer detection. Buszweski et al. (2012) and McCulloch et al. (2006) presented trained dogs with breath samples from individuals with lung cancer and healthy controls. Buszweski et al. (2012) reported that the dogs (the number of dogs was not reported) tested detected the cancerous samples with detection sensitivity and specificity of 82.2% and 82.4% respectively, and McCulloch et al. (2006) reported that the five dogs tested identified the lung cancer samples with 99% sensitivity and 99% specificity. Similarly, Ehmann et al. (2012) and Rudnicka et al. (2014) presented dogs with breath samples from individuals with lung cancer and control samples, but here the control samples included breath samples from healthy individuals as well as individuals with asthma (Rudnicka et al, 2014), individuals with Chronic Obstructive Pulmonary Disease (Ehmann et al., 2012; Rudnicka, 2014), or synthetic samples (Rudnicka et al., 2014). Ehmann et al. (2012) reported that the four dogs tested could indicate the lung cancer sample against the controls with detection sensitivity of 71% and specificity of 93%, while Rudnicka et al. (2014) reported that the two dogs tested had overall detection

sensitivity of 86% and specificity of 72%. Rudnicka et al. (2014) pointed out, however, that when examining each dogs' performance individually, it was apparent that one dog detected lung cancer better than the other. Finally, Amundsen, Sundstrøm, Buvik, Gederaas, and Haaverstad (2014) attempted to determine whether dogs could distinguish between malignant and benign conditions. To begin, Amundsen et al. (2014) trained dogs to identify lung cancer by presenting the dogs with cancerous tissue samples and breath samples from healthy controls; a task that the dogs completed with a high degree of sensitivity and specificity. However, when Amundsen et al. (2014) subsequently presented the dogs with a series of heterogenous breath samples that were obtained from individuals with either malignant or benign conditions (as would be expected in a real world applied setting), the dogs' were only able to detect the malignant conditions with overall sensitivity and specificity of 56% and 33% respectively. This was a more challenging task than in previous studies, and the results were much less promising. As evidenced by the findings of the studies discussed here, dogs' ability to detect lung cancer is uncertain. Although the studies discussed above were all examining dogs' ability to detect lung cancer in breath samples, considerable variability was evident in their reported results, both within and between studies.

Similarly, the only empirical study to examine dog's ability to detect diabetic hypoglycemia from breath reports differing levels of performance across dogs. Hardin, Anderson, and Cattet (2015) applied canine detection of VOCs to the detection of hypoglycemia, or low blood sugar, in individuals with Type 1 Diabetes. In this study, participants collected samples that contained both sweat and breath by rubbing gauze pads on their skin, putting the gauze pad into a bag, and then exhaling into the bag.

Samples were collected during a hypoglycemic event and when blood sugar levels were normal. Hardin et al. (2014) then tested whether six dogs could identify the hypoglycemic samples against normal glycemic samples. Of the six dogs tested, the detection sensitivity ranged from 50-87.5% and specificity ranged from 89.6-97.9%. In this case, five out of the six dogs tested could detect the hypoglycemic sample with a high degree of consistency, but one dog, although at above chance levels, had a much poorer performance than the other five dogs. Taken together, empirical studies of dogs' ability to detect disease illustrate that as the field of canine biomedical detection grows, inconsistencies in findings continue to emerge.

Inconsistencies in the field of canine biomedical detection both within and across studies are likely attributable to several factors, including differences in dog characteristics (breed, age, sex, etc.), differences in training protocols and training time, experimental setup, and sample collection and storage (reviewed by Jezierski et al, 2015; Mosher & McCulloch, 2006). Sample collection and storage is a fundamental concern because if the necessary VOCs are not adequately "captured", dogs will not be able to detect the VOCs in the samples regardless of breed, training, or experimental protocols. In the studies discussed above, the majority of breath samples were collected by having participants breathe through a sample collection tube containing polypropylene wool (Buszewski et al., 2012a; McCulloch et al., 2006) fleece, (Ehmann et al., 2012), or filters (Amundsen et al., 2014). Sonoda et al. (2011) had participants exhale into a breath collection bag and in the case of Hardin et al. (2015), participants breathed into a bag containing a gauze pad before sealing the bag. The length of time these samples were stored before being presented to dogs varied from one day (McCulloch et al., 2006) to six

months (Ehmann et al., 2012). Given that VOCs are volatile gases, the quality and longevity of breath samples must be considered carefully. Depending on the materials used to collect the breath samples, the number of VOCs contained and the amount of time before the volatiles desorb and dissipate will be variable. In an attempt to eliminate concerns about breath sample viability, some researchers have designed their breath collection materials with this in mind.

When breath is exhaled onto a given material, the containment of volatiles will be influenced by both the chemical properties (e.g., affinity to particular VOCs) and physical characteristics (e.g., surface area) of the material. If the material is a solid, such as cotton or wool, the number and types of volatile molecules that adsorb onto the surface of the solid will be limited due to the nature of the interactions between the compounds and the surface active sites, as well as by the number of sites available for bonding (Atkins & de Paula, 2006). In the studies conducted by McCulloch et al. (2006) and Ehmann et al. (2012), the authors specify that the absorbent material used to collect breath samples was polypropylene wool or fleece coated in silicone oil. When a solid material is coated in a liquid phase such as silicone oil, the capacity of the material to retain chemical compounds generally increases because it is no longer limited by the surface area, but rather the volume of the liquid material. Furthermore, mechanisms for gas/liquid phase equilibration (partitioning) are of a more general nature (dipole, induced dipole, hydrogen bonding) than those for surface binding, so a larger variety of molecules would be expected to have an affinity for silicone oil (Patel et al., 2017) than solid surfaces. These characteristics suggest that more VOCs (variety and amount) will dissolve into the silicone oil than will adsorb onto a solid material alone.

Moreover, volatile compounds dissolved into silicone oil may take longer to desorb into the surrounding air than compounds adsorb onto the surface of a solid material. After breath samples are prepared, they are typically stored within a sealed container. Once sealed, the compounds will adsorb onto the surface of the solid, or absorb into the silicone oil until they have reached equilibrium as defined by their individual partition coefficients: the ratio of the concentration of the compound on the breath collection material to its concentration in the air in the tube (Kwon, 2001). However, once the container is opened to the environment (for presentation to a sniffer dog), the volatiles on the materials will begin to desorb into the surrounding air until as the equilibrium is disrupted by diffusion of gas phase VOCs out of the tube, requiring them to be replenished by adsorbed/absorbed compounds. Given that the silicone-coated material has a higher capacity for VOCs than the material alone, once the container is opened it is likely that diffusion of the VOCs from the silicone oil into the surrounding air will persist for a longer period of time (Bir, 2000). Therefore, it is likely that breath samples collected using materials coated in silicone oil will result in a sample that, when presented to a dog, allows them to smell more volatiles for a longer period of time than samples collected with adsorbent materials alone. Considering the variability in the sample collection and storage procedures in the studies outlined above, it is possible that the variability in the performances of the dogs is a result of the varied quality of the breath samples.

The question of whether or not the use of silicone treated adsorbent materials actually increases a dog's ability to detect a breath sample over time has never been empirically examined. The goal of the current study, therefore, was to determine how the

use of silicone-coated cotton for collection of a human breath sample affected dogs' ability to detect breath samples. This question was examined across two experiments. In both experiments the dogs were presented with breath samples in two conditions: an experimental condition in which breath samples were collected using cotton coated in silicone oil and a control condition in which the breath samples were collected using uncoated cotton. In Experiment 1, we examined the detectability of breath samples after a period of two hours, and in Experiment 2 we examined the detectability of breath samples stored for up to four weeks.

3.3 Experiment 1

The goal of Experiment 1 was to determine whether the use of silicone-coated cotton balls affected dogs' ability to detect breath samples that were left exposed to the air for up to two hours. In the experimental condition dogs were presented with breath samples prepared using silicone-coated cotton balls, and in the control condition, the dogs were presented with breath samples prepared using uncoated cotton balls. In both conditions, a breath sample was prepared when a dog arrived at the lab. Once a researcher had breathed through the tube, the sample was left exposed to the surrounding air, and the dog's ability to detect the breath sample was then tested at regular intervals for a maximum of 2 hours.

3.3.1 Method

3.3.1.2 Participants. Four dogs participated in this study. Three of the dogs were purebred border collies: Nutella (4 year old female, intact), Mist (4 year old female, spayed), and Bella (3 year old female, spayed). One of the dogs, Koda, was a border collie mix (2 year old male, neutered). The dogs' owners brought the dogs to the lab once

a week for 2 to 3 hours at a time. Each dog had a designated work day that did not overlap with that of another dog. All four dogs had previously been trained to detect breath using a "Saliency Training Procedure" outlined in Gadbois and Reeve (2016).

3.3.1.3 Stimuli.

3.3.1.3.1 Silicone-coated Cotton Balls. Silicone-coated cotton balls were prepared by first placing cotton balls on a scale and recording their combined weight (ca. 29 g for 40 cotton balls). The cotton balls were then uniformly distributed in a 2.8 L glass dish (ca. 34 cm x 23 cm). Next, an equivalent weight of 100% silicone oil (Clearco Products Co., Inc) (1 g silicone oil/1g of cotton balls) was dissolved in hexane in a 14:1 (hexane:silicone oil) mass ratio (ca. 6.7% w/w solution), since we had previously determined that a 1 g cotton ball could absorb about 14 mL of liquid. The hexane was added to the silicone oil in a stepwise fashion (ca. 21.4 mL/g) from a 1000 mL beaker to ensure complete dissolution. The resulting solution was poured into the glass dish containing the cotton balls. The researcher then held each side of the dish and tilted it carefully from side to side to allow the silicone oil solution to come in contact with every cotton ball. As the cotton balls absorbed the solution they took on a slightly darker colour, so the researcher observed carefully to ensure that the solution was absorbed fully by all of the cotton balls. If necessary, individual cotton balls were turned over so that they would soak up the solution entirely.

The cotton balls soaked in silicone oil and hexane solution were then left inside a fumehood for 24 to 48 hours. This allowed the hexane to evaporate, leaving the cotton balls coated only in silicone oil. The resulting silicone-coated cotton balls were placed into a glass mason jar with the lid secured until needed.

3.3.1.3.2 Breath Samples. Breath samples were collected using PET cylindrical tubes [poly(ethylene terephthalate) ordered from Uline.ca, measuring 20.5 cm in length and 7 cm in diameter]. A single breath sample was prepared by first using metal tweezers to insert two cotton balls inside a tube (see Figure 3.1A). In the experimental condition, these cotton balls were coated in silicone oil and in the control condition, these cotton balls were not coated. One individual prepared all of the breath samples. This individual was careful not to eat any food or drink items that would result in a very strong breath odour the day of breath sample preparations. The individual providing the breath sample held the tube containing the two cotton balls in one hand, very gently placed the tube against their lips, and exhaled two deep breaths through the tube. While exhaling, the individual placed their other hand very lightly against the far end of the tube so as not to block the flow of breath, but as a barrier in case the breath pushed the cotton ball out of the tube. The individual then turned the tube around, placed the other end against their lips, gently blocked the far end of the tube, and exhaled a further two deep breaths for uniform exposure of both cotton balls. Then, using tweezers, one of the cotton balls was removed and placed inside a stainless steel jar with holes in the lid. The breath collection tube containing the other cotton ball was then placed inside a sample station: a wooden platform that held a black plastic PVC tube measuring 25 cm in length (see Figure 3.1 B and C). The PVC pipe was longer than the breath collection tube therefore the breath collection tube sat a few centimeters below the opening of the PVC pipe. This prevented the dogs from coming into contact with the breath collection tube.

Two blank tubes that were not exposed to exhaled breath were prepared using the same procedure, however only one cotton ball was inserted into each. In the experimental

condition, these "blank" cotton balls were coated in silicone oil and in the control condition, these cotton balls were uncoated. No breath was exhaled through the blank tubes. Both blank tubes were then placed inside two additional sample stations, resulting in three identical sample stations (see Figure 3.2). Researchers placed small pieces of tape with sample type indicators on the base of each sample station. The sample stations were placed on the floor so that these indicators faced away from the dog. Which sample station contained the breath sample was decided randomly (see randomization procedure below) for each work day and the PVC tubes and wooden bases were cleaned with water and ethanol between each work day.

3.3.1.4 Procedure. Testing took place inside Dalhousie's Canid Behaviour Research Lab. The lab contains three rooms: Room 1, where the dogs stayed between work sessions, Room 2, a small interior room connecting the other two rooms, and Room 3, where the testing was done (see Figure 3.3). Dogs began a work session by waiting with a handler inside Room 2 while the test stimuli were set up in Room 3 by a researcher. When a test trial began, the door to Room 3 was opened and the dog was led in.

The dogs were tested using a cued, 3 Alternative Forced Choice (3AFC) procedure. Although the dogs were only ever presented with one target sample and two blank samples, this study was part of a larger research program in which future studies would require the dogs to discriminate between multiple breath samples. Therefore, the dogs were cued at the beginning of each trial to maintain consistency. Using this procedure, the dog handler opened the door between Room 2 and Room 3 and then presented the stainless steel jar containing one cotton ball from the prepared breath sample to the dog to

sniff. The handler then directed the dog towards the three sample stations, which were presented simultaneously, lined-up beside one another (see Figure 3.2). One of the stations contained the identical breath sample and the other two stations contained blanks. The dog was required to indicate the station containing the target breath sample (that matched the sample inside the stainless steel jar) using a previously trained "nose hold" behaviour where the dog held its nose or chin against the tube for five full seconds (see Figure 3.2). If the dog chose correctly, it was rewarded with a few pieces of kibble and/or a few tosses of a ball accompanied by verbal praise. If the dog chose incorrectly, the handler uttered a gentle "nope" and the dog was lead out of the room without reward. In-between test trials, the dog was led back into Room 2 and the door to Room 3 was closed. Each work session contained 10 test trials. In between test trials, while the dog and handler were in Room 2, researchers wiped the opening of the PVC tubes with a paper towel wet with a solution of 20% ethanol. During this time, researchers also randomized the positions of the sample stations without the handler or dog watching. Within the 10 trials per session, the position of the target breath sample (left, middle, or right) relative to the blank samples was predetermined using a randomization procedure. Researchers applied specific constraints to the position, such that the breath sample could not be in one location for more than three consecutive trials, the breath sample had to be in each position at least three times within the 10 trial session, and finally, the position of the breath sample could not follow a patterned position (e.g., left, right, left right) for more than 2 iterations.

Each work day, dogs completed 4 or 5 work sessions comprised of 10 trials each. Dogs completed the first work session when the breath sample was relatively "fresh",

meaning within 5 minutes of collection. The subsequent sessions were completed within 15-20 minute intervals after the first. When a session was complete, dogs waited in Room 1 until the next work session began. During this time they had free access to water and were taken out regularly to relieve themselves if necessary.

Each dog completed two work days for each condition (for a total of 20 trials at each time interval for both conditions). First, each dog was presented with breath samples collected using silicone-coated cotton on two successive work days, and then on two subsequent work days, they were presented with samples prepared using uncoated cotton. Within each condition, the dogs' performance was averaged for each time point (e.g., a dog's number of correct trials detecting a "fresh" breath sample on uncoated cotton the first week was averaged with the same dog's number of correct trials detecting a "fresh" breath sample on uncoated cotton the second week).

All procedures were approved by the University Committee on Laboratory Animals before the study was conducted.

3.3.1.5 Analyses. Binomial tests were conducted for each dogs' performance using the number of successful trials out of their total number of trials, 0.33 probability of success per trial, and an alpha level of 0.05. Wilcoxon Signed Rank tests were completed using SPSS (version 22).

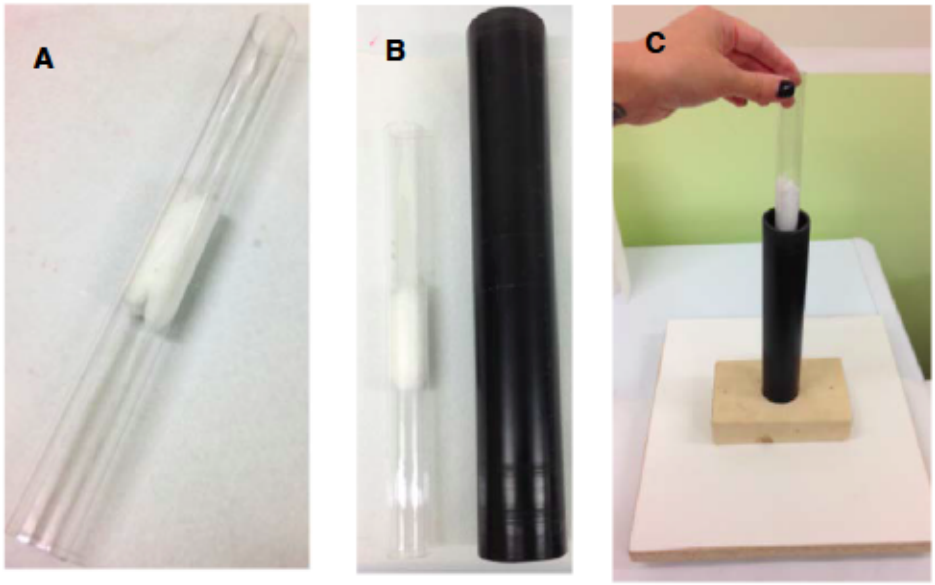

Figure 3.1. The components of a breath sample station. A. The plastic tube containing a cotton ball used for breath sample preparation. B. Illustrates that the breath sample tube is shorter than the black PVC pipe in which the breath sample tube is placed. C. The breath sample tube inside the black PVC pipe is held upright when placed inside the wooden stand. All units combined results in one breath sample station.

Figure 3.2. Three identical sample stations placed on the floor beside one another for presentation to the dogs. One sample station contains a breath sample and the other two contain blank samples. Here, Mist is identifying the station containing the breath sample using the "nose hold" behaviour.

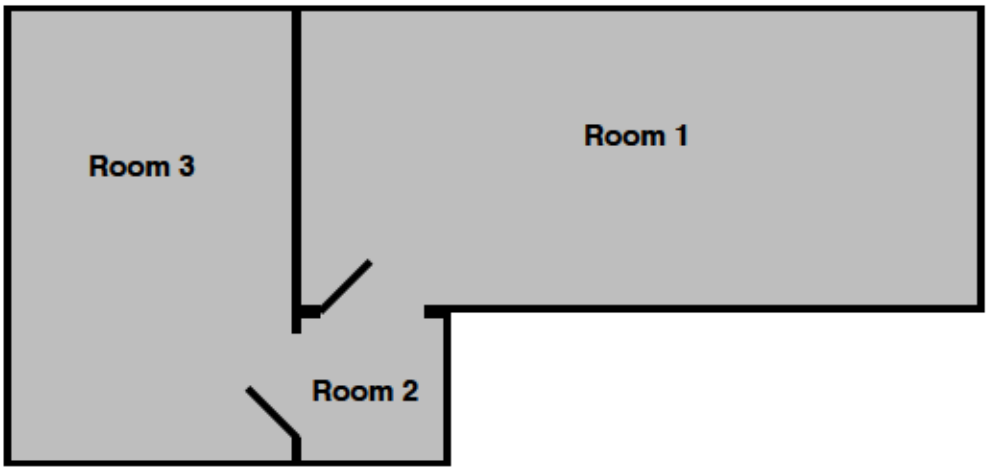

Figure 3.3. The layout of the Canid Behaviour Research Lab at Dalhousie University. Room 1 is where dogs spend time when they are not being tested. Room 2 is where dogs wait between test trials, and Room 3 is where testing takes place.

3.3.2 Results and Discussion

Wilcoxon Signed Rank tests were used to compare each individual's results across conditions and binomial tests were used to determine whether the dogs' performance was greater than what would be expected by chance. The tests revealed that for all four dogs, there was no difference between their performance detecting a breath sample prepared using silicone-coated cotton and a breath sample prepared using a uncoated cotton ball (Bella $Z(5) = -0.677$ $p = 0.498$, Nutella $Z(3) = -0.272$, $p = 0.785$, Koda $Z(3) = 0.00$ $p = 0.593$, Mist $Z(4) = -1.841$ $p = 0.066$; see Figures 3.4 A through D). Furthermore, at every interval up until two hours after the breath samples were prepared, the dogs were able to detect the breath samples at a performance level significantly better than would be expected by chance regardless of whether they were prepared using silicone-coated cotton balls or uncoated cotton balls (see Table 3.1). Therefore, it can be inferred that two hours after the breath samples were prepared, the VOCs in the breath that had adsorbed onto the surface of the cotton in the control condition did not desorb and dissipate faster than the VOCs that had dissolved into the silicone oil coating the cotton in the experimental condition. This is valuable information as it suggests that VOCs in breath samples will remain detectable throughout the duration of longer testing sessions whether or not researchers use silicone-coated materials to collect the breath samples. However, as discussed in the introduction, some researchers stored their breath samples for extended periods of time (Ehmann et al., 2012) before presenting them to dogs. Therefore, Experiment 2 sought to examine whether or not the use of silicone-coated cotton balls increased dogs' detection of breath samples with storage time of up to four weeks.

Table 3.1

Binomial tests of dogs' performance detecting breath samples prepared using cotton balls coated in silicone oil and uncoated cotton balls. Dogs were tested when the breath samples were fresh and then at regular intervals up to two hours later. All dogs completed 20 trials at each interval.

		Coated		Uncoated	
Dog	Session number	No. of successes (n=20 trials)	p - value	No. of successes (n=20 trials)	p - value
Bella	Fresh	15	= 0.00012	14	= 0.00064
	2	17	= 2.24×10^{-6}	15	= 0.00012
	3	15	= 0.00012	12	= 0.0085
	4	15	= 0.00012	14	= 0.00064
	5	15	= 0.00012	15	= 0.00012
Nutella	Fresh	14	= 0.00064	14	= 0.00064
	2	18	< 1×10^{-6}	16	= 1.93×10^{-5}
	3	15	= 0.00012	14	= 0.00064
	4	14	= 0.00064	16	= 1.93×10^{-5}
	5	16	= 1.93×10^{-5}		
Koda	Fresh	17	= 2.24×10^{-6}	18	< 1×10^{-6}
	2	19	< 1×10^{-6}	20	< 1×10^{-6}
	3	19	< 1×10^{-6}	19	< 1×10^{-6}
	4	20	< 1×10^{-6}	20	< 1×10^{-6}
	5	20	< 1×10^{-6}	12	= 0.0085
Mist	Fresh	20	< 1×10^{-6}	19	< 1×10^{-6}
	2	20	< 1×10^{-6}	16	= 1.93×10^{-5}
	3	20	< 1×10^{-6}	17	= 2.24×10^{-6}
	4	19	< 1×10^{-6}	15	= 0.00012

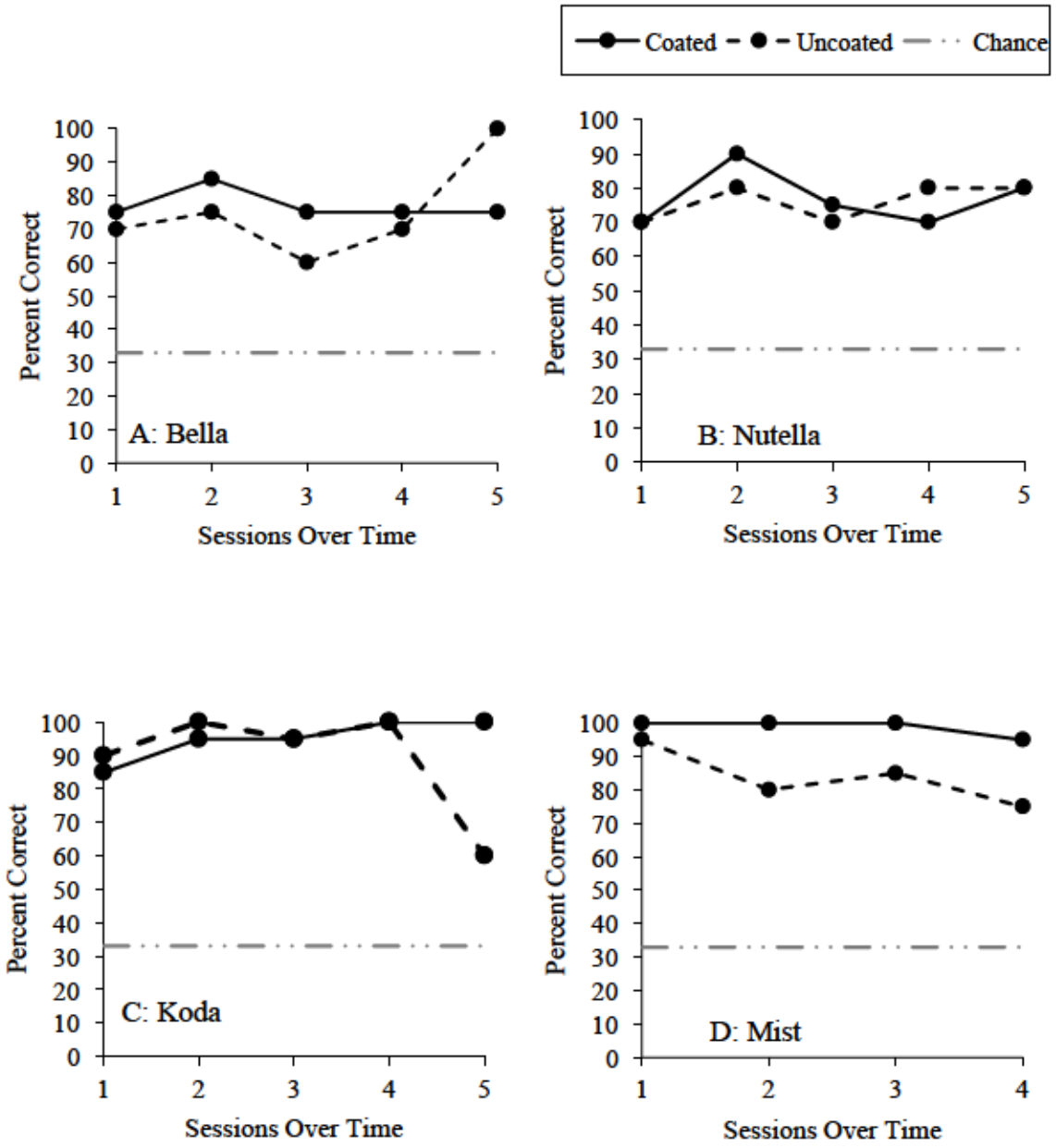

Figure 3.4. Bella (A), Nutella (B), Koda (C), and Mists's (D) performance detecting breath samples that had been prepared and left exposed for two hours. Each data point illustrates her average performance for two work days. The solid black line illustrates their performance detecting breath samples prepared using silicone-coated cotton balls and the dotted line illustrates their performance detecting breath samples prepared using uncoated cotton balls. Chance level (0.33) is illustrated with the intermittent dashed line.

3.4 Experiment 2

The results of Experiment 1 suggested that there was no difference in dogs' ability to detect a breath sample between conditions (breath samples prepared using silicone-coated cotton versus uncoated cotton), up to two hours after breath sample preparation. In Experiment 2, we were interested in whether the use of silicone-coated cotton balls affected dogs' detection of breath samples at much longer time intervals of up to four weeks. Dogs were presented with breath samples prepared using silicone-coated cotton balls (in the experimental condition) and uncoated cotton balls (in the control condition) that were fresh, or one, two, three, and four weeks old.

3.4.1 Method

3.4.1.1.Participants. The dogs' characteristics for Experiment 2 were identical to those of Experiment 1.

3.4.1.2 Stimuli. On the first day that a dog began the experimental condition of the study, a researcher prepared five sample sets using cotton balls coated in silicone oil. Each sample set contained three stimuli: one breath sample and two control samples. The breath samples were collected using the same protocol as Experiment 1, however, once a breath sample was collected, two tight fitting caps were placed on either end of the tube. Caps were placed on the control sample tubes as well. Only one individual prepared all of the breath samples. All sample sets were stored at room temperature inside cold insulated bags (President's Choice © Thermal Grocery Tote) in Dalhousie's Canid Behaviour Research Lab until use.

On the first day that a dog began the control condition of the study, a researcher prepared a further five sample sets using uncoated cotton balls. The same individual that

prepared all of the breath samples in the experimental condition also prepared all of the breath samples for the control condition (and was the same individual that prepared all the breath samples for Experiment 1).

3.4.1.3 Procedure. The dogs were tested on their ability to detect breath samples up to four weeks after samples were prepared. All dogs completed the conditions in the same order (not counterbalanced). First, all dogs were tested in the experimental condition, where they were presented with breath samples that were prepared using cotton balls that had been coated in silicone oil. Subsequently, all dogs were tested in the control condition, where they were presented with breath samples that were prepared using uncoated cotton balls. It was predicted that the breath samples prepared using uncoated cotton would be more difficult for the dogs to detect than breath samples prepared with silicone-coated cotton. Therefore, if uncoated cotton samples were presented to the dogs first, the dogs may have performed poorly and, consequently, become confused about the conditions of reward. If the dogs were then tested on their ability to detect breath samples prepared using silicone-coated cotton, poor performance could be attributed to previous confusion and therefore not be reflective of their ability to detect the breath samples.

On the first day of the study, dogs were tested on their ability to detect a fresh breath sample that was prepared that day. Each subsequent week a new sample set was used, therefore on the second week the sample set used was 1 week old. On the third week, the sample set used was two weeks old, the fourth week the sample set used was three weeks old, and finally, on the fifth week the last sample set used was four weeks old.

In the control condition, one dog, Mist, was only tested using one, two, and three-week old samples (she was not tested using a four-week old sample) due to time constraints. One dog, Bella, completed a subsequent third phase of the study in which she was presented with an additional three-week old breath sample prepared using cotton coated in silicone oil. The reason for testing Bella with a third phase will be explained below.

In preparation for testing, a single sample set was removed from storage and the caps were removed from all of the tubes. One of the cotton balls in the breath sample tube was removed and placed in a stainless steel jar with holes in the lid. Then all three tubes were placed inside sample stations (see Figures 1 and 2).

The general procedure was identical to that of Experiment 1; The dogs were tested using a cued, 3 Alternative Forced Choice procedure. Each dog completed 2-4 sessions of 10 trials each per work day and their performance was averaged across the sessions for each work day. All trials were performed double-blind.

If a dog's performance one week averaged less than 50%, and was accompanied by behavioural indications that the dog was frustrated or upset (e.g., whining, barking, refusal to approach the sample stations), testing did not continue the following week, and the condition was considered complete. This was because it was assumed that if a dog could not indicate a sample, they would be unlikely to be able to indicate an older sample the following week. It was also important to avoid the possibility that a dog became confused with the goal of the task as a result, its performance would decrease regardless of its ability to detect the breath sample.

3.4.1.4 Analyses. Binomial tests were conducted for each dog using the number of successful trials out of their total number of trials, 0.33 probability of success per trial, and an alpha level of 0.05. Wilcoxon Signed Rank tests were completed using SPSS (version 22).

3.4.2 Results and Discussion

The results of Experiment 2 showed notable individual differences in each dog's performance across conditions. Therefore, each individual dog's results will be presented and discussed independently.

In the experimental condition of Experiment 2, binomial tests revealed that Bella could detect the breath samples at a performance significantly better than expected by chance for fresh, one, two, three, and four-week old samples (see Table 3.2 and Figure 3.5A). Conversely, Bella's performance in the control condition suggested that the she could not detect the breath samples on uncoated cotton balls after one week of storage. When presented with breath samples prepared using uncoated cotton, Bella was able to detect the fresh breath sample at above chance levels. This result was not surprising, given the results of Experiment 1, which suggested that a breath sample on uncoated cotton remains detectable for at least two hours after preparation. However, when tested on her ability to detect the 1 week old sample, her performance dropped to 10%. Furthermore, as the session progressed from trial 1 to trial 10, Bella displayed behaviours that would suggest she was frustrated, such as barking, whining, and refusing to approach the sample stations. Bella's performance combined with her behaviour throughout the session indicated that the one-week old breath sample prepared on uncoated cotton was difficult, if not impossible for Bella to detect. To explore this finding further, a week after

being presented with the one-week old breath sample on uncoated cotton, Bella completed a work session where she was presented with a new three-week old breath sample prepared using silicone-coated cotton. In this session, Bella was 100% correct in identifying the breath sample against the blank samples. Therefore, Bella's results clearly suggested that, when stored for multiple weeks, the breath samples prepared using silicone-coated cotton balls were much easier to detect than the breath samples prepared using uncoated cotton balls. An examination of Nutella's performance reveals a similar pattern across conditions (see Table 3.2 and Figure 3.5B).

In the experimental condition when the breath samples were prepared using silicone-coated cotton, Nutella was able to detect fresh, one, two, and three-week old breath samples at a level significantly better than would be expected by chance (see Table 3.2). However, throughout this time her performance gradually decreased from an average of 97% with a fresh sample, to 83% at two weeks, 77% at three weeks, and then fell to 68% at four weeks. These results suggest that, despite the use of the silicone-coated cotton balls, the concentration of breath volatiles was still gradually decreasing throughout the four week period to a point that Nutella's detection of the breath was effected. Despite Nutella's gradual decrease in performance in the experimental condition, a comparison of her performance in the experimental condition to her performance in the control condition illustrates a dramatic difference. In the control condition when Nutella was presented with breath samples prepared using uncoated cotton, she was able to detect the fresh sample with an average of 97%; a level of performance to be expected considering the results of Experiment 1. But when Nutella was tested with the one-week old breath sample, her performance dropped to only 10%.

During this session, Nutella demonstrated behaviours that were not observed in the experimental condition, such as whining and barking. As with Bella, the researchers interpreted these behaviours as signs that Nutella had difficulty detecting the breath sample and was frustrated with the task. Given that every trial in this study contained a stimulus that matched the cue, Bella and Nutella were never trained to provide a "no match" behaviour whereby the dogs would indicate that none of the stimuli matched the cue stimulus. If they had been trained to provide a "no match" indication, however, it is possible that they would have exhibited less frustration behaviours when they could not perceive the breath sample. It remains clear though, that for both Bella and Nutella, the use of silicone-coated cotton balls greatly improved the detectability of the breath samples over an extended period of time. There are a few potential reasons for why, at the molecular level, this would be the case.

As discussed in the introduction, if a breath collection tube is sealed after breath sample donation, the compounds within the tube will adsorb onto the surface of the cotton, or partition into the silicone oil until each compound has reached an equilibrium between the concentration of the compounds on the cotton or in silicone oil and the air contained in the tube (defined as the partition coefficient). In Experiment 2, the breath collection tubes were sealed for up to four weeks and theoretically the concentration of compounds within the tube did not change during this time. However, the results suggest a decrease in the concentrations of gas phase VOCs over time, especially in the case of the uncoated cotton. A number of mechanisms might be responsible for this. Although unlikely, it is possible that the seal of the tubes was not sufficiently tight to prevent diffusion of VOCs out of the tube. This would affect both conditions, but the impact

would be greater for the uncoated cotton, since the capacity of the storage medium is less and would therefore be affected more quickly. A second possible explanation is irreversible adsorption of the volatiles, (either onto the sides of the tube or into the cellulose in the case of the uncoated cotton) or permeation into the plastic walls of the tube. Irreversible adsorption implies a strong chemical interaction that essentially precludes equilibration with the gas phase and, again, would be expected to have a greater effect for the uncoated cotton. A third possibility that cannot be excluded is the slow reaction of key volatile organics to form inactive (non-volatile or non-detectable) compounds. This scenario would be expected to be more likely in the case of the uncoated cotton, where the molecules are interacting directly with active sites, than in the case of the coated cotton, where the molecules are dissolved in relatively inert silicone oil. Any combination of these factors could explain the poorer detectability of the breath samples in the control condition, as demonstrated by Nutella and Bella. Taken together, these factors could explain the poorer detectability of the breath samples in the control condition, as demonstrated by Nutella and Bella.

Contrary to the results of Bella and Nutella, however, no differences were observed in either Koda's or Mist's performance across conditions; Wilcoxon Signed Rank tests revealed no significant difference between Koda ($Z(4) = -0.365$, $p = .715$) and Mist's ($Z(3) = -0.816$, $p = .414$) performance on breath samples collected using silicone-coated cotton balls and breath samples collected using uncoated cotton balls (see Figures 3.5C and 3.5D). Furthermore, the results of binomial tests revealed that their performance in both conditions, across all storage times, was significantly higher than would be expect by chance (see Table 3.2). In the experimental condition, Mist and Koda were able to

detect the breath samples prepared using silicone-coated cotton balls at above chance levels for fresh, one, two, three, and four-week old samples, consistent with the results of Bella and Nutella. Mist and Koda's performance in the control condition differed greatly from Bella and Nutella's. In the control condition of Experiment 2, Mist was able to detect the breath samples prepared using uncoated cotton balls at above chance levels when the samples were fresh, one, two, and three weeks old, and Koda was able to detect the samples at above chance levels at fresh, one, two, three, and four weeks after sample preparation. During all of the sessions, neither Mist nor Koda showed any indications that they had any difficulty detecting the samples. Given the stark difference between the performance of Koda and Mist, and Bell and Nutella, one must consider the possibility that the breath samples presented to Koda and Mist contained more easily detectable breath volatiles, compared to the samples presented to Bella and Nutella. This could occur if, on the day the breath samples were prepared for Koda and Mist, the breath donor had consumed particularly strong smelling food or drink items or taken medications that she did not consume the day she prepared breath samples for Bella and Nutella. Per the procedure, only one individual carefully prepared all breath samples to minimize this source of variability. Despite best efforts to minimize inter-sample differences, however, it cannot be ignored as a potential confound. It is also possible that the tubes used to prepare the breath samples for Koda and Mist were sealed more tightly than those used for Bella and Nutella, resulting in fewer losses of compounds from the tube. However, the same model tubes and caps were used for preparation of all of the breath samples for each dog, so it would be expected that any differences between the seal of individual tubes and caps would be minor.

Alternatively, it is possible that the differences observed between dogs is simply the result of individual differences. As will be discussed below, individual differences between dogs can greatly affect their performance on olfactory tasks (Jezierski et al., 2008). Given that individual differences are ever-present and have the potential to affect the interpretation of results, we argue that examining them in detail is an important part of studies of canine olfactory ability.

Table 3.2

Binomial tests of dogs' performance detecting breath samples prepared using cotton balls coated in silicone oil and uncoated cotton balls over storage time of up to four weeks.

		Coated			Uncoated		
Dog	Storage time (in weeks)	No. of trials	No. of successes	*p* - value	No. of trials	No. of successes	*p* - value
Bella	Fresh[1]	30	29	$< 1x10^{-6}$	30	29	$= 6.07x10^{-6}$
	One	20	15	= 0.00013	10	1	= 0.09
	Two	30	24	$< 1x10^{-6}$			
	Three	30	22	$= 6.07x10^{-6}$			
	Four	30	29	$< 1x10^{-6}$			
Nutella	Fresh	20	17	$= 2.24x10^{-6}$	30	29	$< 1x10^{-6}$
	One	20	16	$= 1.93x10^{-6}$	10	1	= 0.09
	Two	30	25	$< 1x10^{-6}$			
	Three	30	23	$= 1.04x10^{-6}$			
	Four	40	27	$= 6.59x10^{-6}$			
Koda	Fresh	30	28	$< 1x10^{-6}$	30	30	$< 1x10^{-6}$
	One	30	21	$= 3.01x10^{-5}$	30	30	$< 1x10^{-6}$
	Two	40	40	$< 1x10^{-6}$	30	30	$< 1x10^{-6}$
	Three	30	23	$= 1.04x10^{-6}$	30	27	$< 1x10^{-6}$
	Four	30	23	$= 1.04x10^{-6}$	30	29	$< 1x10^{-6}$
Mist	Fresh	30	30	$< 1x10^{-6}$	30	30	$< 1x10^{-6}$
	One	30	23	$= 1.04x10^{-6}$	30	28	$< 1x10^{-6}$
	Two	20	30	$< 1x10^{-6}$	30	30	$< 1x10^{-6}$
	Three	30	28	$< 1x10^{-6}$	30	30	$< 1x10^{-6}$

1. Dogs were tested on their ability to detect breath samples prepared the same day of testing

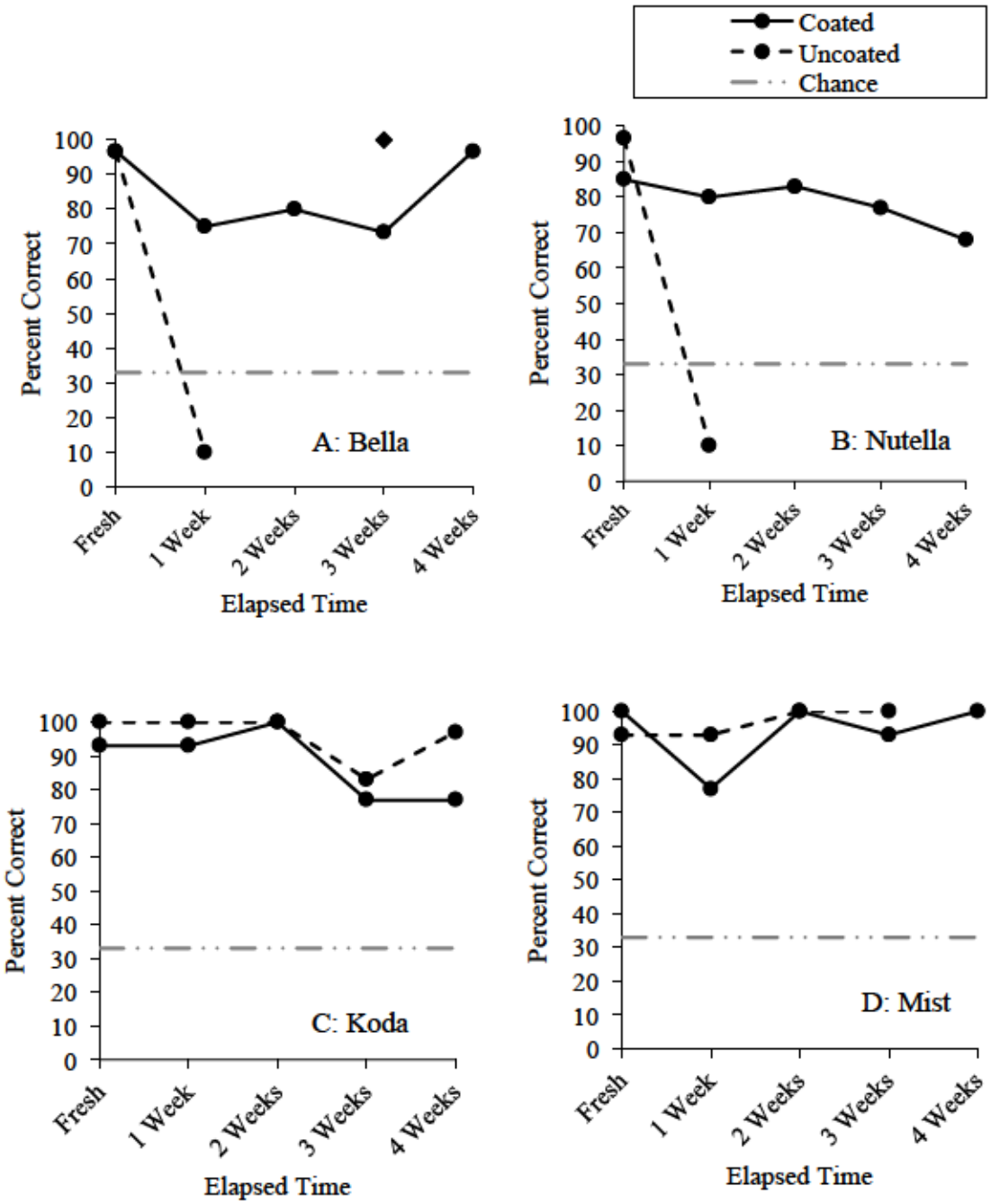

Figure 3.5. Bella (A), Nutella (B), Koda, (C), and Mist's (D) performance detecting breath samples that had been stored for up to four weeks. Each data point illustrates their average performance for one work day. The solid black line illustrates her performance detecting breath samples prepared using silicone-coated cotton balls and the dotted line illustrates their performance detecting breath samples prepared using uncoated cotton balls. Note that in Bella's figure, the diamond marker indicates her performance detecting a second three-week old breath sample. Chance level (0.33) is illustrated with the intermittent dashed line.

3.5 General Discussion and Conclusion

The goal of this study was to determine whether there was a difference in dogs' detection of breath samples in two conditions: breath samples prepared using silicone-coated cotton (experimental condition) and those prepared using uncoated cotton (control condition). This was examined with two experiments. In the first experiment, four dogs were tested for their ability to successfully detect breath samples in both conditions at regular intervals up to two hours after preparation of the breath sample. In the second experiment, the same four dogs' ability to detect breath samples stored in both conditions was tested each week, for up to four weeks after breath sample preparation. In both experiments, the dogs' ability to detect the breath samples was assessed using a cued, 3AFC task, and the dogs' performance in detecting the breath samples was compared between conditions.

The results of Experiment 1 showed that for all four dogs, there was no significant difference in the detectability of the breath samples across conditions; the dogs' performance was above chance at all time points up to two hours, regardless of whether the breath samples were prepared with silicone-coated cotton balls or uncoated cotton balls. The results of Experiment 2 showed that the use of silicone-coated cotton balls greatly improved Bella and Nutella's detection of breath samples that were stored for up to four weeks. However, this effect was not observed for Mist or Koda; Mist and Koda were able to detect both breath samples that were prepared with silicone-coated cotton and uncoated cotton at above chance levels when the samples were stored for up to four weeks. As discussed above, the differences observed between dogs' performance in Experiment 2 could be the result of differences in the breath samples themselves (more

pungent VOCs, differences in the tightness of the caps, differing dissipation of compounds). In the future, this factor could be eliminated by having a set diet on sample collection days, or by preparing all the sample sets on the same day and therefore testing the dogs on the same day each week. An examination of the literature suggests that individual dog characteristics could also be responsible for the differences in performance.

Given the performance of the four dogs across conditions in Experiment 2, individual differences between the dogs were readily apparent. This is not a unique finding; of the studies reviewed in the introduction, variability between dogs' performance was noted in all cases except for McCulloch et al. (2006). However, it is important to note that most of the dogs in these studies were either previously trained for specialized olfactory tasks (Rudnicka et al., 2014; Sonoda et al, 2011; Buszewski et al., 2012a), or were pet or rescue dogs selected for characteristics such as trainability and eagerness to sniff (Hardin et al., 2014; Ehmann et al., 2012; McCulloch et al., 2006). In the current study all four dogs were carefully selected for their trainability and motivation to work, they all had an equal amount of training within the lab, all four had demonstrated the ability to detect and discriminate between fresh breath samples with a high degree of accuracy (Reeve, Wallace, & Gadbois, unpublished data). Therefore, it is unlikely that the differences in performance observed here were due to differing levels of experience or desire to work. Rather, some studies suggest that inherent differences between dogs may influence their success in olfactory tests.

Rooney, Bradshaw, and Almey (2004) had 244 dog handlers and trainers from UK government agencies rate 30 dog behaviour characteristics for their relative

importance for a specialized search dog. Among those characteristics rated most important were: acuity of sense of smell, incentive to find an object which is out of sight, tendency to hunt by smell alone, stamina, ability to learn from being rewarded, and consistency of behaviour from day to day. Although no formal behavioural assessments were completed on the dogs in the current study, the authors would rate the four dogs high on these particular characteristics.

Another individual difference that may have affected the dogs' performance is the way in which they sniffed the samples and the stimuli. Jezierski, Walczak, and Górecka (2008) found that when dogs were tested using a matching-to-sample procedure with a five-station lineup, individual differences in dogs' "sniffing style" (willingness to sniff inside a jar), the length of time spent sniffing stimuli, and the number of stimuli sniffed influenced the number of errors the dogs made in the lineup. These variables were not assessed in the current study, but certainly could have been a factor contributing to the performance differences observed between dogs in Experiment 2 when the breath volatiles were potentially weaker and therefore more difficult to detect.

Finally, additional research suggests that, even within the same breed, differences in olfactory ability can have a genetic basis. In a study of five olfactory receptor genes, Lesniak et al. (2008) found that all five genes had allelic variations, most of which resulted in variations in the subsequent proteins. The allelic variations ultimately resulted in structural differences in olfactory receptors. Consequently, Lesniak et al. (2008) believe that some olfactory receptor genotypes may result in olfactory receptors that have a superior receptor-binding affinity for particular VOCs. Therefore, even with comparable levels of motivation and training, the olfactory task presented here

may have simply been easier for Mist and Koda because of genetic variation in the genes that code for their olfactory receptors.

An important point which is a corollary to the discussion of individual differences is that canine biomedical detection studies often employ a small sample design - researchers train and test only a few dogs at a time. As mentioned above, in most cases these dogs are not selected randomly, but are carefully selected for behavioural characteristics such as those described by Rooney et al., (2004). Therefore the dogs in biomedical detection studies are inherently not representative of the general population of dogs (Morgan & Morgan, 2009). Researchers select their dogs this way because the goal of these studies is not to illustrate that *all* dogs are capable of detecting medically relevant volatiles, but rather that some exceptional dogs can. It is extremely important in studies of canine biomedical detection that individual differences between dogs are examined carefully and that each dog's performance is reported and discussed on an individual basis. Pooling of the data would neglect these individual differences and could result in skewed results that misrepresent individual dog's capabilities (Machlis, Dodd, & Fentress, 1985). Unfortunately, this may lead readers to misinterpret the utility of dogs as biomedical detection tools.

Furthermore, the examination of individual differences can elucidate how specific study characteristics may affect the performance of some dogs (Kazdin, 2011). The experimental conditions that afford one dog a high degree of olfactory acuity may not result in a second dog demonstrating equal performance. It is reasonable to expect that breath samples collected for studies of canine biomedical detection may need to be stored for a length of time, but as the results of the current study showed, this storage

time can decrease the detectability of the samples for some dogs. The most important finding of the current study was that for some dogs, the use of silicone-coated cotton balls greatly improved the detectability of breath samples stored for extended periods. Although the use of silicone-coated cotton balls was not necessary for Mist and Koda to detect four-week old breath samples, it dramatically improved the detectability of the samples for Bella and Nutella. In the current study Bella and Nutella had had at least one full year of previous training. Therefore, the finding that silicone coated cotton increased their detectability of breath samples meant the retention of these two dogs for further studies. We extend this finding to suggest that it is very valuable for the training of biomedical detection dogs. The training time required for detection and alert dogs is extensive and, as a result, can be an expensive process. Any procedural fine-tuning that increases dog retention should be valued.

Finally, as discussed in the introduction, a larger variety of molecules would be expected to have an affinity for silicone oil (Patel et al., 2017) resulting in the silicone-coated cotton balls holding a larger number VOCs than cotton alone. Therefore, in tests of canine biomedical detection, the use of silicone-coated cotton balls for the collection of exhaled breath would result in a breath sample that is more representative of the actual exhaled breath VOC profile compared to breath samples collected using uncoated cotton.

Although beyond the scope of this paper, researchers must consider additional challenges with the use of breath samples, such as the separation of endogenous versus exogenous volatile compounds, and the possibility that only certain VOCs are contained over time (Schmidt & Podmore, 2015). Furthermore, the affinities of particular compounds to silicone oil means that the concentration of compounds contained within

the breath sample will not necessarily be identical to those expressed in the breath itself (Patel et al., 2017). Future work should utilize GC-MS to examine which particular compounds absorb into the silicone oil and if the concentration of compounds in the breath samples using a given breath collection technique change with increased time. Given the findings presented here, the authors argue that the use of silicone-coated cotton balls should be implemented in all studies of canine biomedical detection of breath samples to adequately contain breath VOCs over time, and ultimately allow researchers to study more dogs.

3.6 Acknowledgements

This research was supported by an IWK Health Centre Category A Research Grant, and an NSERC PGS-D to Catherine Reeve.

We would like to acknowledge Dr. Elizabeth McLaughlin, Dr. Beth Cummings, Dr. Vin LoLordo, and Dr. Ray Klein for their consultation and feedback. We would further like to acknowledge the dog owners who reliably brought their dogs to our lab each week, without whom this study would not have been possible. Finally, we thank the many student volunteers who contributed their time.

3.7 References

Amann, A., & Smith, D. (Eds.). (2013). *Volatile Biomarkers: Non-Invasive Diagnosis in Physiology and Medicine*. Oxford, England: Elsevier.

Amann, A., de Lacy Costello, B., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., ... & Risby, T. (2014). The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *Journal of Breath Research, 8,* 034001. doi:10.1088/1752-7155/8/3/034001

Amundsen, T., Sundstrøm, S., Buvik, T., Gederaas, O. A., & Haaverstad, R. (2014). Can dogs smell lung cancer? First study using exhaled breath and urine screening in unselected patients with suspected lung cancer. *Acta Oncologica*, *53*, 307-315. doi: 10.3109/0284186X.2013.819996

Atkins, P. & de Paula, J. (2006). *Physical Chemistry* (8th ed.). Great Britain: Oxford University Press.

Balseiro, S. C., & Correia, H. R. (2006). Is olfactory detection of human cancer by dogs based on major histocompatibility complex-dependent odour components?–A possible cure and a precocious diagnosis of cancer. *Medical Hypotheses*, *66*, 270-272. doi:10.1016/j.mehy.2005.08.027

Bir, D. (2000). Partition coefficient calculation of selected terpenes and low molecular weight solvents between tall oil fatty acid and air and polydimethyl siloxane oil and air. *Journal of the American Oil Chemists' Society*, *77*, 163-169.

Burak, R.C., & Liang, J. (1987). The early detection of cancer in the primary-care setting: factors associated with the acceptance and completion of recommended procedures. *Preventive Medicine, 16,* 739-751. doi: http://dx.doi.org/10.1016/0091-7435(87)90014-4

Buszewski, B., Ligor, T., Jezierski, T., Wenda-Piesik, A., Walczak, M., & Rudnicka, J. (2012a). Identification of volatile lung cancer markers by gas chromatography–mass spectrometry: comparison with discrimination by canines. *Analytical and Bioanalytical Chemistry*, *404*, 141-146. doi: 10.1007/s00216-012-6102-8

Buszewski, B., Rudnicka, J., Ligor, T., Walczak, M., Jezierski, T., & Amann, A. (2012b). Analytical and unconventional methods of cancer detection using odor. *Trends in Analytical Chemistry*, *38*, 1-12. doi:http://dx.doi.org/10.1016/j.trac.2012.03.019

Craven, B. A., Paterson, E. G., & Settles, G. S. (2009). The fluid dynamics of canine olfaction: unique nasal airflow patterns as an explanation of macrosmia. *Journal of The Royal Society Interface*, *7*, 933-943. doi:10.1098/rsif.2009.0490

Corradi, M. & Mutti, A. (2013). Exhaled breath analysis in occupational medicine. In A. Amann & D. Smith (Eds.), *Volatile Biomarkers. Non-invasive Diagnosis in Physiology and Medicine* (117-125). Oxford, England: Elsevier.

DeSantis, C., Ma, J., Bryan, L., & Jemal, A. (2014). Breast cancer statistics, 2013. *CA: A Cancer Journal for Clinicians, 64,* 52- 62. doi: 10.3322/caac.21203

Ehmann, R., Boedeker, E., Friedrich, U., Sagert, J., Dippon, J., Friedel, G., & Walles, T. (2012). Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *European Respiratory Journal*, *39*, 669-676. doi: 10.1183/09031936.00051711

Gadbois, S., & Reeve, C. (2014). Canine olfaction: Scent, sign, and situation. In A. Horowitz (Ed.), *Domestic Dog Cognition and Behavior* (pp. 3-29). Berlin Heidelberg: Springer.

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*(4), 509-517. doi: 10.1007/s13300-015-0135-x

Jezierski, T., Walczak, M., & Górecka, A. (2008). Information-seeking behaviour of sniffer dogs during match-to-sample training in the scent lineup. *Polish Psychological Bulletin*, *39*, 71-80. doi: 10.2478/v10059-008-0010-y

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*, 027001. doi:10.1088/1752-7155/9/2/027001

Kazdin, A. E. (2011). *Single-Case Research Designs: Methods for Clinical and Applied Settings*. New York, NY: Oxford University Press.

Kwon, Y. (2001). Partition and distribution coefficients. In *Handbook of Essential Pharmacokinetics, Pharmacodynamics, and Drug Metabolism for Industrial Scientists* (pp. 35-71). New York, NY: Kluwer Academic/Plenum Publishers.

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. doi:10.1093/jhered/esn057

Machlis, L., Dodd, P. W. D., & Fentress, J. C. (1985). The pooling fallacy: problems arising when individuals contribute more than one observation to the data set. *Ethology*, *68*, 201-214.

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. doi: 10.1177/1534735405285096

Minh, T. D. C., Blake, D. R., & Galassetti, P. R. (2012). The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes Research and Clinical Practice*, *97*, 195-205. doi:10.1016/j.diabres.2012.02.006

13

Modak, A.S. (2013). An Update on C-Breath tests: The transition to acceptability into clinical practice. In A. Amann & D. Smith (Eds.), *Volatile Biomarkers. Non-invasive Diagnosis in Physiology and Medicine* (245-262). Oxford, England: Elsevier.

Morgan, D.L. & Morgan, R.K. (2009). *Single Case Research Methods*. Thousand Oaks, CA: SAGE publications.

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. doi:10.1016/j.jveb.2010.01.002

Neupane, S., Peverall, R., Richmond, G., Blaikie, T. P., Taylor, D., Hancock, G., & Evans, M. L. (2016). Exhaled breath isoprene rises during hypoglycemia in type 1 diabetes. *Diabetes Care*, *39*, e97-e98. doi: 10.2337/dc16-0461

Patel, M.J., Popat, S.C., & Deshusses, M.A. (2017). Determination and correlation of the partition coefficients of 48 volatile organic and environmentally relevant compounds between air and silicone oil. *Chemical Engineering Journal, 310,* 72-78. doi: http://dx.doi.org/10.1016/j.cej.2016.10.086

Pearsall, M. D., & Verbruggen, H. (1982). *Scent. Training to Track, Search, and Rescue*. Loveland, CO: Alpine Publications.

Phillips, M., Herrera, J., Krishnan, S., Zain, M., Greenberg, J., & Cataneo, R. N. (1999). Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications*, *729*, 75-88.

Quignon, P., Giraud, M., Rimbault, M., Lavigne, P., Tacher, S., Morin, E., ... & Galibert, F. (2005). The dog and rat olfactory receptor repertoires. *Genome Biology*, *6*, R83. doi: 10.1186/gb-2005-6-10-r83

Rooney, N. J., Bradshaw, J. W., & Almey, H. (2004). Attributes of specialist search dogs—a questionnaire survey of UK dog handlers and trainers. *Journal of Forensic Science*, *49*, 1-7.

Ross, B.M., & Esarik, A. (2013). The analysis of oral air by selected ion flow tube mass spectrometry using indole and methylindole as examples. In A. Amann & D. Smith (Eds.), *Volatile Biomarkers. Non-invasive Diagnosis in Physiology and Medicine* (77-88). Oxford, England: Elsevier.

Rudnicka, J., Walczak, M., Kowalkowski, T., Jezierski, T., & Buszewski, B. (2014). Determination of volatile organic compounds as potential markers of lung cancer by gas chromatography–mass spectrometry versus trained dogs. *Sensors and Actuators B: Chemical*, *202*, 615-621. doi: http://dx.doi.org/10.1016/j.snb.2014.06.006

Siegel, R., DeSantis, C., & Jemal, A. (2014). Colorectal cancer statistics, 2014. *CA: A Cancer Journal for Clinicians*, *64*, 104-117. doi: 10.3322/caac.21220

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians*, *66*, 7-30. doi: 10.3322/caac.21332

Smith, D., Španěl, P., Fryer, A. A., Hanna, F., & Ferns, G. A. (2011). Can volatile compounds in exhaled breath be used to monitor control in diabetes mellitus? *Journal of Breath Research*, *5*, 022001. doi:10.1088/1752-7155/5/2/022001

Solga, S.F., and Risby, T.H. (2013). Issues and challenges in human breath research: perspectives from our experience. In A. Amann & D. Smith (Eds). *Volatile Biomarkers: Non-Invasive Diagnosis in Physiology and Medicine* (pp 19-24). Oxford, England: Elsevier.

Sonoda, H., Kohnoe, S., Yamazato, T., Satoh, Y., Morizono, G., Shikata, K., ... & Inoue, F. (2011). Colorectal cancer screening with odour material by canine scent detection. *Gut*, *60,* 814-819. doi:10.1136/gut.2010.218305

Szulejko, J. E., McCulloch, M., Jackson, J., McKee, D. L., Walker, J. C., & Solouki, T. (2010). Evidence for cancer biomarkers in exhaled breath. *IEEE Sensors Journal*, *10*, 185-210.

Sun, X., Shao, K., & Wang, T. (2016). Detection of volatile organic compounds (VOCs) from exhaled breath as noninvasive methods for cancer diagnosis. *Analytical and Bioanalytical Chemistry*, *408*, 2759-2780. doi: 10.1007/s00216-015-9200-6

Waggoner, L. P., Jones, M. H., Williams, M., Johnston, J. M., Edge, C. C., & Petrousky, J. A. (1998, December). Effects of extraneous odors on canine detection. In *Enabling Technologies for Law Enforcement and Security* (pp. 355-362). International Society for Optics and Photonics.

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. doi:10.1016/j.applanim.2005.07.009

Wilson, A. D. (2015). Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. *Metabolites*, *5*, 140-163. doi:10.3390/metabo5010140

CHAPTER 4: CAN DOGS GENERALIZE THE ODOUR OF HYPOGLYCEMIA IN BREATH SAMPLES FROM INDIVIDUALS WITH TYPE 1 DIABETES?

The manuscript prepared for this study is presented below. Catherine Reeve, under the supervision of Dr. Simon Gadbois, was responsible for devising the research question. Furthermore, under the supervision of Dr. Elizabeth Cummings, Catherine was responsible for meeting with potential sample donors with type 1 diabetes and demonstrating sample collection, obtaining consent, monitoring participants' progress, and collecting completed samples. With the help of Dr. Gadbois, Catherine designed the three experiments presented here. She completed the training and testing of the dogs largely with the help of the honours student, Sonia Smith. Catherine was responsible for data analysis and interpretation. She wrote the initial draft of the manuscript, and received and incorporated feedback from her co-authors and committee members. The manuscript will be submitted to Applied Animal Behaviour in the future. The full reference for this manuscript is as follows:

Reeve, C., Cummings, E., Smith, S., & Gadbois, S. (2017). Can dogs generalize the odour of hypoglycemia in breath samples from individual with type 1 diabetes?

4.1 Abstract

Studies examining dogs' ability to detect hypoglycemia in individuals with type 1 diabetes are few and their findings are contradictory. This study investigated whether dogs could discriminate between breath samples collected from people with type 1 diabetes during different glycemic states, and then further tested dogs' ability to generalize the odour of hypoglycemia. This was explored with three experiments. In Experiment 1, we tested four dogs' ability to discriminate between hypoglycemic, normoglycemic, and hyperglycemic breath samples from the same individual using a cued, three alternative forced choice task. In Experiment 2, two dogs were presented with three breath samples from the same individual: one hypoglycemic, one normoglycemic, and one hyperglycemic, and were trained to identify the hypoglycemic sample using a Go/No-Go procedure. Then the dogs' ability to generalize was tested by presenting them with a second sample set (that the dogs had never smelled before) from the same individual and the number of times they identified the new hypoglycemic sample was observed. In Experiment 3, we tested whether one dog could generalize the odour of hypoglycemia between two different people by further presenting additional samples from a different individual. The results of Experiment 1 showed that all four dogs discriminated between the breath samples at above chance levels. The results of Experiment 2 indicated that one dog showed evidence of being able to generalize the odour of hypoglycemia within samples from the same individual, but the second dog could not. The results of Experiment 2 showed that the dog could not generalize the odour of hypoglycemia across two different individuals. Considered together with the

current literature, the results have important implications for the training of Diabetic Alert Dogs.

4.2 Introduction

Type 1 diabetes is one of the most common chronic diseases in children (Acerini et al., 2016) and hypoglycemia, or low blood sugar, is the most common acute complication of the disease. If left untreated, severe hypoglycemic episodes can result in a loss of consciousness, a seizure, coma, or death (Frier, 2004). Given the potential severity of hypoglycemia, individuals with type 1 diabetes can develop a fear of hypoglycemia that negatively affects their management of diabetes, their psychological well-being, and their quality of life (Gonder-Frederick et al., 2011; Wild et al., 2007). Therefore, timely and accurate detection of hypoglycemia is imperative to maintain long term health in individuals with type 1 diabetes.

Individuals with type 1 diabetes may self-monitor their blood sugar levels with home blood glucose meters or they may use newer technologies such as continuous glucose monitors (CGM) and continuous subcutaneous insulin infusion (CSII) pumps with suspension of insulin infusion functions (Acerini et al., 2016; Riemsma et al., 2016). Use of these technologies can result in improved glycemic control but reports of their effect on the incidence of hypoglycemia are contradictory (for reviews see Jeitler et al., 2008; Riemsma et al., 2016). Some individuals choose not to use, or to discontinue use of these technologies because they can be uncomfortable to wear and the use of subcutaneous systems increase the likelihood of infection at insertion and tape sites (Aye et al., 2010; Wong et al., 2014). Moreover, individuals using CGMs and CSII may see these devices as a constant reminder of their disease, which may contribute to social impairments and a feeling of social stigmatization (Schabert et al., 2013). The limitations of current technologies have prompted research examining the efficacy of a potentially

more user-friendly hypoglycemia detection system: trained domestic dogs.

Reports of dogs behaving abnormally when their owners become hypoglycemic provide evidence that dogs may be able to detect hypoglycemia in their owners (Chen et al., 2008; Tauveron et al., 2006; Wells et al., 2008). Following such reports, organizations throughout North America and the U.K. have begun to train Diabetic Alert Dogs (DADs). DAD owners report a decrease in the frequency of hypoglycemic episodes, and consequently, have less fear of hypoglycemia and a higher quality of life (Gonder-Frederick et al., 2013). Despite the reported successes of DADs, empirical science elucidating how dogs detect hypoglycemia is minimal and inconclusive.

Given the incredible olfactory acuity of dogs (Craven et al., 2010; Miklosi, 2007), it is believed that dogs who detect hypoglycemia are detecting volatile organic compounds (VOCs) that signal a physiological change in their owner. When human body cells undergo changes due to metabolic fluctuations, disease, or infection, they release compounds that become dissolved in the blood and become volatized during pulmonary circulation. VOCs are then emitted in breath and sweat (Amann et al., 2014) in concentrations of parts per billion (nmol/mol) and parts per trillion (pmol/mol) (Schmidt & Podmore, 2015). Research shows that diabetic hypoglycemia may be correlated with a specific VOC profile (Minh et al., 2012; Neupane et al., 2012; Smith et al, 2012) and since dogs have been shown to detect odours at concentrations of parts per trillion (Walker et al., 2006), it is possible that hypoglycemia-detecting dogs are detecting these changes in VOCs. Working under this premise, two research teams have tested dogs' ability to detect hypoglycemia in sweat and breath.

In 2013, Dehlinger et al. tested whether three previously trained DADs could identify hypoglycemia in sweat samples from three individuals unknown to the dogs. After obtaining hypoglycemic and normoglycemic sweat samples by the same protocol used to collect the samples for the initial training of the dogs, Dehlinger et al. (2013) presented the dogs with individual sweat samples one at a time for 30-45 seconds. If the dogs detected a hypoglycemic sample, they were to ring a bell beside the sample. Dehlinger et al. (2013) found that none of the dogs could indicate the hypoglycemic samples at above chance levels.

Hardin et al. (2015) collected combined sweat and breath samples from four individuals with type 1 diabetes during hypoglycemic and normoglycemic states. After training six dogs to signal to the hypoglycemic samples, they then tested the dogs' ability to locate the hypoglycemic samples in a seven-station line-up. Hardin et al. (2015) reported that the dogs could detect the hypoglycemic samples at above chance levels, most with a high degree of sensitivity and specificity. Importantly though, Hardin et al. (2015) presented the dogs with a small number of samples that they were previously reinforced for indicating during training. Therefore, it is possible that they simply memorized the odour profile of the individual hypoglycemic samples rather than identifying a hypoglycemia-specific odour that is consistent across individuals.

Taken together, the results of Dehlinger et al. (2013), and Hardin et al. (2015) present inconclusive evidence as to whether dogs can be trained to detect hypoglycemia. Therefore, the goal of this study was to determine whether dogs could detect hypoglycemic breath samples, discriminate them from normoglycemic, and hyperglycemic breath samples, and generalize the odour of hypoglycemia. This was

tested in three experiments. In Experiment 1, four dogs were presented with breath samples from one person in each glycemic state and were tested for their ability to discriminate between the samples. In Experiment 2, two dogs were trained to identify a single hypoglycemic breath sample and then tested for their ability to generalize the odour of hypoglycemia to new samples they had never smelled before from *the same* person. In Experiment 3, one dog was again trained to identify a hypoglycemic sample and then tested for her ability to generalize the odour to new samples she had never smelled before from a *different* person.

4.3 Experiment 1

4.3.1 Method

4.3.1.1 Participant recruitment. People with type 1 diabetes were recruited through the Pediatric Diabetes Clinic at the IWK Health Centre in Halifax, NS, Canada and by word of mouth. Participants were required to have been diagnosed with type 1 diabetes at least six months prior to participation, to have had no episodes of hypoglycemia unawareness in the past six months, and in the case of children, to have sufficient parental support to complete the procedures.

4.3.1.2 Participant demographics. Twenty-three individuals provided informed consent and were given a sample collection kit. Of those 23, 11 individuals did not respond to contacts, or withdrew from the study. Three people returned some completed samples, but did not complete all. Nine people completed sample collection (Age range = 7-36, M_{age} = 17, Median = 14). Five of these also completed a second phase of sample collection. Of the nine who completed sample collection, seven were female and two were male.

Families that participated in the study received $15 for the initial visit at the IWK in which consent was obtained, and kits and instruction on sample collection were provided. A $10 gift card was issued following completion of sample collection and collection of record sheets and samples. Participants were given the opportunity to participate in a second phase of sample collection, after which they received a second gift card valued at $10 to a location of their choosing.

4.3.1.3 Breath samples. After in person training, participants were given sample collection kits that contained instructions for how to collect the samples, the supplies to collect the breath samples, sample information sheets, and three hypoglycemia symptoms questionnaires (adapted from Ross et al., 1998).

Participants were asked to collect six breaths samples in total; two breath samples during a hypoglycemic event (defined here as a blood glucose monitor test indicating blood sugar levels of less than 4 mmol/L), two breath samples during a hyperglycemic event (higher than 14 mmol/L), and two breath samples when their blood sugar was normal (between 5-10 mmol/L).

The breath collection tubes were plastic cylindrical tubes (Uline.ca), measuring 20.5 cm in length and 7 cm in diameter, inside which were silicone oil coated cotton balls (see Figure 4.1). Each breath collection tube was labeled for the glycemic states (e.g., hypoglycemic sample #1, hypoglycemic sample #2, normoglycemic sample #1, etc). Note that the breath collection tubes labelled for the collection of a hypoglycemic breath sample contained two silicone oil coated cotton balls (the reason for which will be explained below). Both ends of the tubes were capped with a tight fitting rubber cap.

Figure 4.1. Breath collection tube containing a cotton ball and with caps secured.

Participants were instructed to collect breath samples according to the home glucose meter reading after first performing their regular blood glucose check. Participants were instructed never to induce hyperglycemia or hypoglycemia, and not to collect a breath sample if they felt too unwell to do so.

To donate the samples, participants were instructed to remove the rubber caps on both ends and exhale two deep breaths through one side of the tube, and then an additional two deep breath through the opposite end of the tube. Then they were to replace the tube caps and store the collected breath sample inside the insulated bag provided. After collecting the breath sample, participants were asked to fill in the sample information sheet with the date, time, and their blood sugar level at the time they collected the sample. For hypoglycemic samples, they completed a hypoglycemic symptom questionnaire after collecting the breath sample. Completed sample sets were returned to the clinic or were picked up in person by researchers. If participants collected and returned all six samples, they were then invited to provide a further nine samples (3 hypoglycemic, 3 hyperglycemic, and 3 normoglycemic samples).

The IWK Health Centre Research Ethics Board and Dalhousie University approved all experimental protocols and procedures before the experiment began.

4.3.1.4 Dogs. Four dogs participated in Experiment 1: Nutella (4 year old female border collie, intact), Bella (3 year old female border collie), Koda (2 year old male border collie mix), and Mist (4 year old female border collie). Each dog was brought by their owner to Dalhousie's Canid Behaviour Lab once or twice a week for 2 to 3 hour work days. Only one dog worked in the lab at any given time. All four dogs had previously been trained to detect and discriminate between breath samples using the

procedure outlined in Gadbois and Reeve (2016). All procedures were approved by the University Committee on Laboratory Animals before the study was conducted (protocol number 13-004).

4.3.1.5 Sample presentation. Each participant's breath samples were organized into sample sets. Each sample set contained five breath tubes: three donated breath samples from the same individual (one hypoglycemic sample, one normoglycemic sample, and one hyperglycemic sample) and two additional blank tubes breath collection tubes in which one silicone-coated cotton ball was inserted, but onto which no breath was exhaled. The blank tubes were prepared in the lab as soon as a sample kit was received and added to the contents of the sample kit until use.

Breath samples were prepared for presentation to the dogs by first removing the caps from the breath collection tubes and placing them inside a sample station; a wooden platform holding a black PVC tube upright. The breath collection tubes were shorter in length and smaller in diameter than they PVC tubes, such that they fit inside the PVC tube and sat an inch lower than the opening of the PVC tube (see Figures 4.2A-C). This prevented dogs from coming into direct contact with the breath collection tubes. It is important to note that, before the hypoglycemic sample was placed inside a sample station, one of the cotton balls from the hypoglycemic sample was removed with tweezers and placed inside a stainless-steel jar with holes in the lid (to serve as the "cue", as described below).

4.3.1.6 Procedure. The dogs were tested at Dalhousie's Canid Behaviour Lab. The lab contains three adjoining rooms: Room 1 where the dogs stayed between work sessions, Room 2, a small interior room connecting the other two rooms, and Room 3,

where the testing was done (see Figure 4.3). Dogs began a work session by waiting with a handler inside Room 2 while the test stimuli were set up in Room 3 by a researcher. When a test trial began, the door to Room 3 was opened and the dog was led in.

The dogs were tested for their ability to discriminate between the breath samples with a cued, three alternative forced choice (3AFC) procedure. Using this procedure, the dog handler opened the door between Room 2 and Room 3 and presented the dog with the stainless-steel jar containing one cotton ball from the hypoglycemic breath sample. Once the dog sniffed the sample in the jar, the handler directed the dog towards three sample stations, which were presented simultaneously, lined up beside one another (see Figure 4.4). The breath samples inside the sample stations differed depending on the phase of testing (see Table 4.1). In the first phase, the three sample stations held one hypoglycemic breath sample and two blank tubes. The dog was required to identify the station containing the matching hypoglycemic sample. All four dogs identified their choice using a previously trained "nose hold" behaviour where they held their nose or chin against the tube for five full seconds (see Figure 4.4). If the dog correctly identified the station containing the matching hypoglycemic breath sample, they were rewarded with a few pieces of kibble and/or a few tosses of a ball accompanied by verbal praise. If the dog chose incorrectly, the handler uttered a gentle "nope" and the dog was lead out of the room without reward. In-between each test trial, the dog was lead back into Room 2 and the door to Room 3 was closed. One work session was comprised of 10 such trials. Once a work session was completed, the dogs were given a 10-15 minute break inside Room 1 where they had free access to water. During this time, the dogs were also taken outside to relieve themselves if necessary. To successfully complete phase 1, the dogs

were required to complete one full session at a rate of 80% correct or higher. If this criterion was met, the dog progressed to phase 2.

In phase two, one of the blank tubes was removed and a hyperglycemic donated by the same individual was put in its place. Now the dog was required to smell both the hypoglycemic sample and the hyperglycemic sample and continue to identify only the hypoglycemic sample. In this phase, the dog was required to identify the hypoglycemic breath sample at a rate of 80% correct for three consecutive sessions (of 10 trials each), or at a rate of 90-100% correct for two consecutive sessions (of 10 trials each). If this was completed successfully, the dog progressed to phase three. In phase three, the second blank breath sample tube was removed and was replaced with a normoglycemic breath sample donated by the same individual. Here in phase three, the dog was required to meet the same criterion as phase two to be considered to have discriminated successfully between the breath samples donated from different glycemic states.

In between test trials while the dog and handler were in Room 2, researchers wiped the opening of the PVC tubes with a paper towel wet with a solution of 20% ethanol and water. During this time, researchers also randomized the positions of the sample stations without the handler or dog watching. Within the 10 trials per session, the position of the hypoglycemic breath sample (left, middle, or right) relative to the other samples was predetermined using a pseudo randomization procedure (rolling a die). Researchers applied specific constraints to the position, such that the breath sample could not be in one location for more than three consecutive trials, the breath sample had to be in each position at least three times within the 10 trial session, and finally, the position of

the breath sample could not follow a patterned position (e.g. left, right, left right) for more than two iterations.

Most dogs completed all three phases within one work day. If, however, a dog was unable to reach criterion at any phase of testing within one work day, the sample caps were placed back on their corresponding tubes and were presented again to the dog the following work day. In total, each dog was tested with three different sample sets from three different individuals and samples were never used with more than one dog. Each dogs' performance was averaged at each phase.

4.3.2 Results and Discussion

All four dogs successfully discriminated between the samples at each phase of testing (see Figures 5A through D). Most importantly, the dogs' performance in phases two and three were at above chance levels, providing evidence that they could discriminate between the breath samples from different glycemic states. Recall, however, that a cued procedure was used here, meaning that the dogs were always shown a hypoglycemic sample "cue" before locating the matching hypoglycemic sample in the series of hypoglycemic, normoglycemic, and hyperglycemic samples. It follows, then, that the results of this experiment do not provide evidence that the dogs are able to detect a general odour of hypoglycemia, but only that they can match the odour of specific breath samples, contributing little more to the findings of Hardin et al. (2015). Therefore, to further examine whether the dogs could generalize the odour of hypoglycemia across multiple sample sets, two additional studies were conducted.

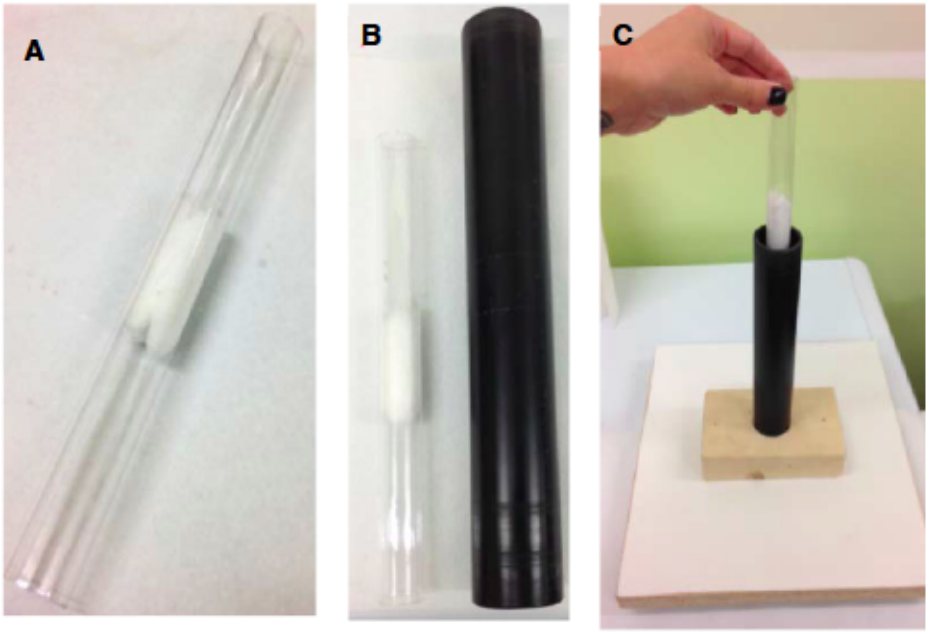

Figure 4.2. The components of a breath sample station. A. A breath collection tube with the end caps removed. B. Illustrates that the breath sample tube is shorter than the black PVC pipe in which the breath sample tube is placed. C. The breath sample tube inside the black PVC pipe is held upright when placed inside the wooden stand. All units combined results in one breath sample

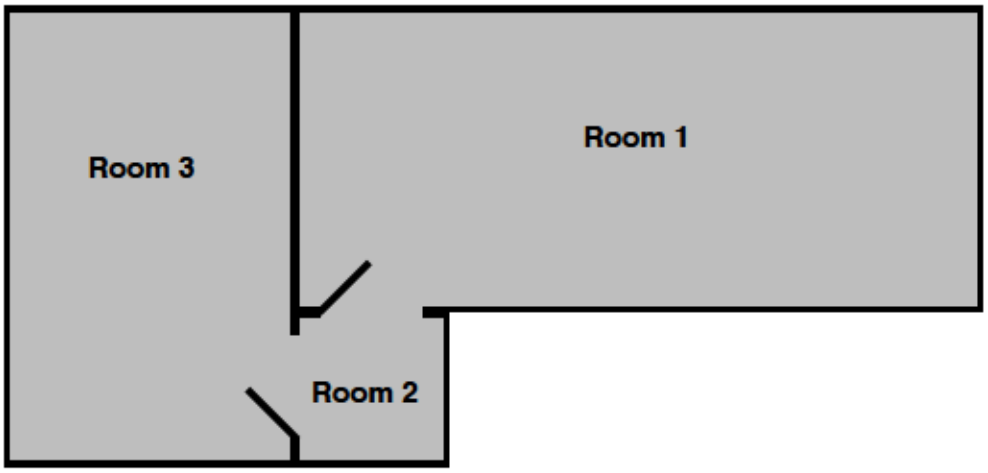

Figure 4.3. The layout of the Canid Behaviour Research Lab at Dalhousie University. Room 1 is where dogs spend time when they are not being tested. Room 2 is where dogs wait between test trials, and Room 3 is where testing takes place.

Figure 4.4. Three identical sample stations are placed on the floor beside one another for presentation to the dogs. One sample station contains a hypoglycemic sample and the other two contain the hyperglycemic and normoglycemic samples. Here, Mist is identifying the station containing the hypoglycemic breath sample using the "nose hold" behaviour.

Table 4.1.

The three testing phases used to test dogs' ability to discriminate between breath samples collected by people with type 1 diabetes during hypoglycemia, normoglycemia, and hyperglycemia.

Phase	Stimuli (hypoglycemic sample target in all sessions)	Criterion
Phase 1	Hypoglycemic sample, two blank samples	1 session ≥ 80%
Phase 2	Hypoglyecmic sample, hyperglycemic sample, one blank sample	3 sessions = 80% or 2 sessions ≥ 90
Phase 3	Hypoglycemic sample, hyperglycemic sample, normoglycemic sample	3 sessions = 80% or 2 sessions ≥ 90

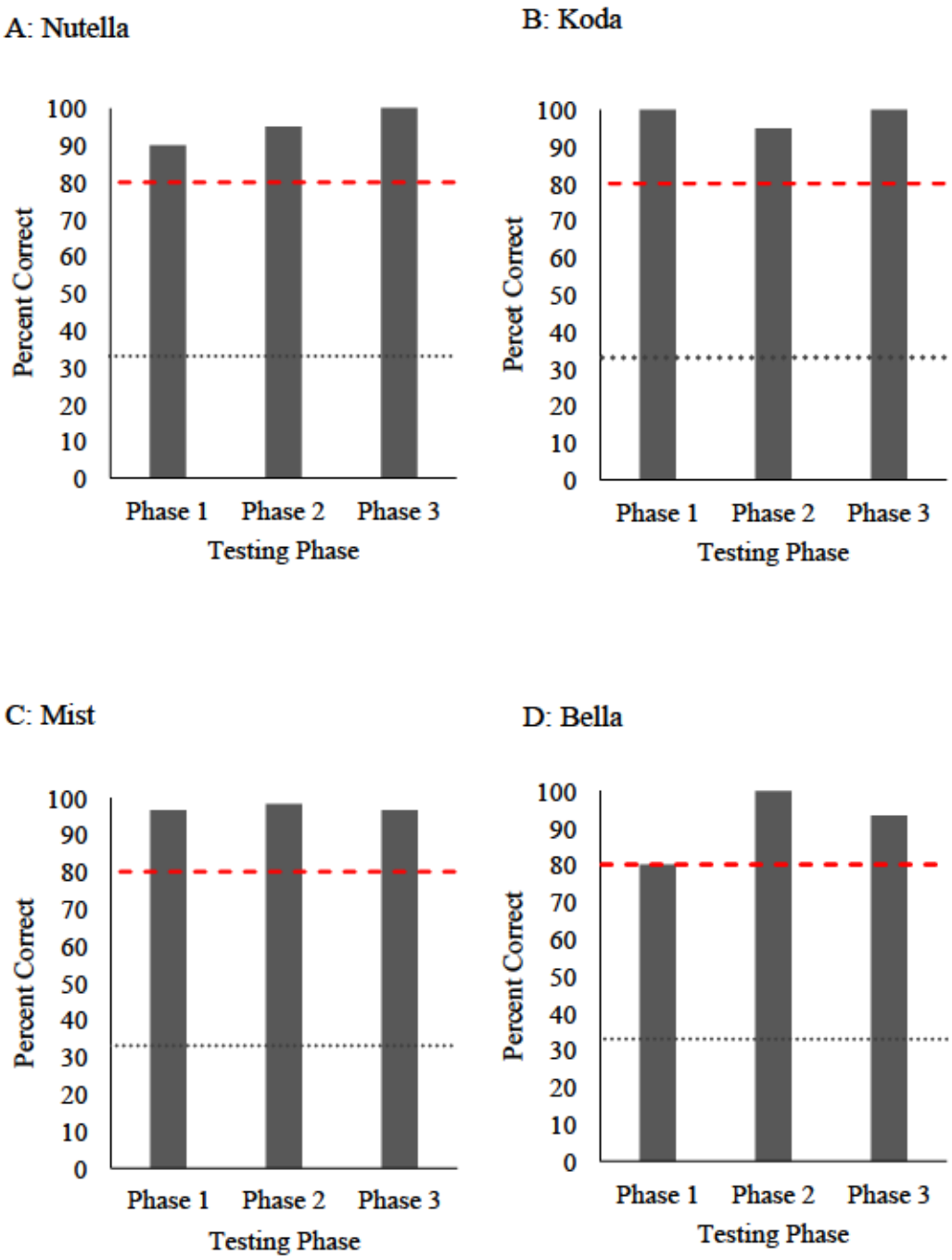

Figure 5. Nutella (A), Koda (B), Mist (C), and Bella's (D) performance identifying hypoglycemic breath samples on each phase of discrimination testing. The dashed line represents the criterion (80%). The criterion is 10% higher than chance level (70%) which was calculated using binomial probability with 10 trials, 0.33 probability of success per trial (represented by the dotted line).

4.4 Experiment 2

The results of Experiment 1 showed that the dogs could discriminate between breath samples collected at times of hypoglycemia, normoglycemia, and hyperglycemia. The goal of experiment 2 was to examine whether dogs could generalize the odour of hypoglycemia across multiple sample sets from the same individual.

4.4.1 Method

4.4.1.1 Dogs. Nutella and Koda participated in this experiment.

4.4.1.2 Samples. Breath samples were organized into sample sets, where one sample set consisted of three breath samples from the same individual: one hypoglycemic, one normoglycemic, and one hyperglycemic. Samples were prepared for presentation to the dogs by removing the tube lids and placing the tubes inside larger PVC tubes that stood upright in wooden platforms (See Figure 2). The PVC tubes were longer than the breath collection tubes (25cm in length), therefore the dogs could not directly contact the breath collection tubes. This setup resulted in three individual sample stations.

4.4.1.3 Training

4.4.1.3.1 Shaping the responses. The sample stations were presented to the dogs one at a time. When the dogs were presented with a hypoglycemic sample, they were encouraged to sniff the tube and then asked to provide a previously trained nose hold behaviour. If the dog held its nose or chin to the tube for the full five seconds, a secondary reinforcer (a clicker) was used, followed by a food reward. Conversely, when the dogs were presented with a normoglycemic or hyperglycemic breath sample, they

were encouraged to sniff the tube, were asked to sit, and then a secondary reinforcer was used (clicker), followed by a food reward.

4.4.1.3.2 Procedure. Once the behaviours were shaped, the dogs completed two work days of training within which we completed as many trials as possible given the amount of time the dog could spend at the lab. This resulted in Koda completing 80 trials and Nutella completing 120 trials. Samples were presented to the dogs in a double blind set-up. If, at the end of the trials, the dogs could provide the appropriate behavioural response corresponding with the sample type presented to them at above chance levels, the training phase was considered complete.

4.4.1.4 Testing

4.4.1.4.1 Samples. The samples presented during testing included the samples used during the training phase plus one additional sample set (one hypoglycemic, one normoglycemic, and one hyperglycemic) *from the same individual* who donated the samples used during the training phase. Therefore, the dogs could be presented with any one of six breath samples from the same individual, three of which they had previously been trained on, and three that they had never smelled before.

4.4.1.4.2 Procedure. The dogs were presented with one of the six potential samples each trial, with hypoglycemic samples presented half of the time and normoglycemic or hyperglycemic samples presented the other half of the time so that there was an equal distribution of "Go" and "No-Go" trials. Samples from the first sample set and samples from the second sample set were presented in equal proportions. All trials were double blind such that researchers who knew the identity of the sample were not visible to the dog handler or the dog. As with the training phase, the dogs

completed as many test trials as possible within two work days. This resulted in Koda completing 120 trials and Nutella completing 160 trials.

4.4.1.5. Analyses. The numbers of hits (true positives), misses (false negatives), false alarms (false positives), and correct rejections (true negatives) were documented for the training phase and for the testing phase. Using these values, each dog's sensitivity, specificity, accuracy, and precision was calculated as sensitivity = hits/(hits + misses), specificity = correct rejections/(correct rejections + false alarms), accuracy = (hits + correct rejections)/(hits + misses + correct rejections + false alarms), and precision = hit/(hits + false alarms). Signal Detection Theory was used to calculate a d' and a C (bias) value for each dog for each phase (Gadbois & Reeve, 2016).

4.4.2 Results and Discussion

4.4.2.1 Training. Nutella and Koda's distribution of hits, misses, false alarms, and correct rejections in the training phase can be seen in Table 4.2. During the initial stages of training, both Nutella and Koda obtained a performance of 100% within two 10-trial sessions, therefore demonstrating little to no acquisition phase. Following this, for both dogs, all of their training data was included in analyses. As illustrated by her results, Nutella obtained an overall accuracy level of 76%. The criterion value close to zero demonstrates that she was close to being an unbiased decision maker; she was no more likely to commit misses than she was false alarms. This is further reflected by her similar sensitivity and specificity values. These results show that Nutella could detect the hypoglycemic sample equally as often as the normoglycemic and hyperglycemic samples. Koda demonstrated an overall accuracy of 78%, but was more of a conservative decision maker than Nutella (as demonstrated by the higher C value). His conservative

decision making was reflected in the higher ratio of correct rejections to false alarms than the ratio of hits to misses, suggesting that he was confident in indicating the normoglycemic and hyperglycemic samples, but then sometimes failed to identify what *was* a hypoglycemic sample. Accordingly, his sensitivity value was lower than his specificity value.

4.4.2.2 Testing. Nutella and Koda completed 160 and 120 trials of testing, respectively. Their distribution of hits, misses, false alarms, and correct rejections can be seen in Table 4.3. Comparing the overall d' and accuracy values between Nutella and Koda reveals that Nutella was better able to generalize the odour of hypoglycemia than Koda. More specifically, examination of the dogs' responses to the new samples alone revealed that Nutella's sensitivity was much higher than Koda's. Nutella correctly identified more than half of the new hypoglycemic samples, suggesting that she detected some overlap in the VOC content between the first hypoglycemic sample and the second hypoglycemic sample from the same individual. Conversely, Koda correctly identified most of the hypoglycemic samples from the original sample set, but was unable to identify any of the new hypoglycemic samples. These results suggest that Koda likely memorized the odour of the first hypoglycemic sample and then was unable to detect any overlapping VOCs between the first hypoglycemic sample and the second hypoglycemic sample from the same person. Both Nutella and Koda, however, demonstrated a high degree of specificity, as reflected by the ratio between their correct rejection and false alarm rates. This means that both dogs were accurate in identifying normoglycemic and hyperglycemic samples; especially Koda who did not commit a single false alarm. Given

Nutella and Koda's response profiles, it follows that both dogs were conservative decision makers as indicated by their positive C (bias) values.

Table 4.2

Nutella and Koda's performance on the training phase of a Go/No-Go task detecting hypoglycemia

	Nutella	Koda
Total number of trials	120	80
Hypoglycemic samples	60	40
Normoglycemic/hyperglycemic samples	60	40
Hits	0.75 (n=45)	0.70 (n=28)
Misses	0.25 (n=15)	0.30 (n=12)
Correct Rejections	0.77 (n=46)	0.85 (n=34)
False Alarms	0.23 (n=14)	0.15 (n=6)
d'	1.41	1.56
C	0.03	0.26
Sensitivity	75%	70%
Specificity	77%	85%
Accuracy	76%	78%
Precision	76%	82%

Table 4.3

Nutella and Koda's distribution of responses with new sample set from the same individual.

	Nutella			Koda		
	Original sample set	Second sample set from same individual	Overall	Original sample set	Second sample set from same individual	Overall
Total number of samples presented	81	79	160	60	60	120
Hypoglycemic samples	38	42	80	30	30	60
Normoglycemic/ Hyperglycemic samples	43	37	80	30	30	60
Hits	0.79 (n=30)	0.62 (n=26)	0.70 (n=56)	0.87 (n=26)	0	0.43 (n=26)
Misses	0.21 (n=8)	0.38 (n=16)	0.30 (n=24)	0.13 (n=4)	1 (n=30)	0.57 (n=34)
Correct Rejections	0.86 (n=37)	0.89 (n=33)	0.88 (n=70)	0.90 (n=27)	1 (n=30)	0.95 (n=57)
False Alarms	0.14 (n=6)	0.11 (n=4)	0.12 (n=10)	0.10 (n=3)	0	0.05 (n=3)
d'	1.89	1.51	1.68	2.41	--*	1.47
C	0.14	0.47	0.31	0.08	--*	0.91
Sensitivity	79%	62%	70%	87%	0.00**	43%
Specificity	86%	89%	88%	90%	100%**	95%
Accuracy	83%	75%	79%	88%	50%	69%
Precision	83%	87%	85%	90%	--	90%

*d' and C cannot be calculated if zero hits are committed

**These are mathematically correct results, however, this result must be interpreted cautiously because Koda rejected all new samples, including the hypoglycemic sample, indicating that he could not accurately discern what was *not* a hypoglycemic sample.

4.5 Experiment 3

The results of Experiment 1 provided some evidence that Nutella could generalize the odour of hypoglycemia across multiple samples from one individual. In Experiment 2, we were interested in whether she could generalize the odour of hypoglycemia further, by presenting her with hypoglycemic samples from two different individuals.

4.5.1 Method

4.5.1.1 Dog. Nutella was the only dog that participated in this experiment.

4.5.1.2 Samples. Both sample sets from Experiment 1 plus two additional sample sets from *a different person* were added to the pool of samples. This resulted in a total of 12 samples, 6 (2 hypoglycemic, 2 normoglycemic, and 2 hyperglycemic) of which Nutella had smelled previously, and 6 (2 hypoglycemic, 2 normoglycemic, and 2 hyperglycemic) of which were completely novel to her. All samples were presented in the same sample stations described in Experiment 1.

4.5.1.3 Procedure. Nutella was presented with one of 12 potential samples at a time. The distribution of samples was such that half of the time the sample presented was a hypoglycemic sample and half of the time the sample was a normoglycemic or hyperglycemic sample. Samples from the first individual and samples from the second individual were presented in equal proportions. All trials were double blind such that researchers who knew the identity of the sample were not visible to the dog handler or the dog during testing.

4.5.2 Results and Discussion

Nutella completed a total of 80 trials in which she was presented with two sample sets from one individual that she had previously seen, and two novel sample sets from a different individual. Her distribution of responses can be seen in Table 4.4. As illustrated by the data, her overall level of accuracy was 70%. However, a closer look at the distribution of responses reveals that she demonstrated a low level of sensitivity; she only identified the new hypoglycemic samples in less than half of their presentations, as evidenced by the low ratio of hits to misses. An interesting finding was that, although Nutella had previously shown the ability to correctly identify the sample types form the first sample set in Experiment 1, she only identified less than half of these hypoglycemic samples here in Experiment 2, resulting in equal performance identifying the hypoglycemic samples she was previously trained on and the new samples she had never smelled before. The authors hypothesize that the difficulty of the task led Nutella to become confused about the conditions of the reward, or that the number of samples was overwhelming and she could not remember the odour of the original sample.

Nutella's degree of specificity was much better than her sensitivity. As shown by the high ratio of correct rejections to false alarms, Nutella was consistently able to identify the normoglycemic and hyperglycemic samples from both sample sets. Given the distribution of Nutella's responses across sample sets, however, it was difficult to assess whether she could generalize the odour of hypoglycemia from one previously known set of samples to new sample sets from a different individual. A more valid assessment would require more trials and would ideally include samples sets from additional people.

Table 4.4

Nutella's distribution of responses when presented with new sample sets from a different individual

	Original sample sets	Sample sets from different individual	Overall
Total number of samples presented	40	40	80
Hypoglycemic samples	20	20	40
Normoglycemic/ Hyperglycemic samples	20	20	40
Hits	0.45 (n=9)	0.45 (n=9)	0.45 (n=18)
Misses	0.55 (n=11)	0.55 (n=11)	0.55 (n=22)
Correct Rejections	0.95 (n=19)	0.85 (n=17)	0.90 (n=36)
False Alarms	0.05 (n=1)	0.15 (n=3)	0.20 (n=4)
d'	1.52	0.91	0.72
C	0.89	0.58	0.48
Sensitivity	45%	45%	45%
Specificity	95%	85%	90%
Accuracy	70%	65%	68%
Precision	90%	75%	82%

4.6 General Discussion

The three studies presented here examined dogs' ability to detect hypoglycemia in breath samples from people with type 1 diabetes. In Experiment 1, breath samples were collected from people with type 1 diabetes when their blood sugar was low, normal, and high and then tested four dogs' ability to discriminate between the breath samples using a cued, 3AFC procedure. In Experiment 2, we sought to determine if dogs could generalize the odour of hypoglycemia across multiple breath samples obtained from one person. To examine this, two dogs, Nutella and Koda, were presented with a set of three breath samples from the same individual: one hypoglycemic, one normoglycemic, and one hyperglycemic, and were trained to identify the hypoglycemic sample using a Go/No-Go procedure. Then, during testing, a second sample set (that the dogs had never smelled before) from the same person was presented to the dogs and their ability to identify the new hypoglycemic sample was observed. Finally, in Experiment 3 we tested whether Nutella could generalize the odour of hypoglycemia between two different people by further presenting her with additional samples from a different individual.

The results of Experiment 1 showed that all four dogs could discriminate between the breath samples from different glycemic states with a high level of accuracy. This finding is consistent with that of Hardin et al. (2015) who showed that dogs could detect a hypoglycemic breath and sweat sample in a line-up with normoglycemic samples and blank samples. But, as discussed above, the findings of Hardin et al. (2015) and the findings of Experiment 1 presented here, do not robustly illustrate dogs' ability to detect hypoglycemia. Combined with the findings of Experiments 2 and 3, however, the current

study provides some compelling evidence for dogs' detection of a general odour of hypoglycemia.

The results of Experiment 2 provided some evidence that Nutella, but not Koda, could generalize the odour of hypoglycemia across two hypoglycemic samples from the same individual. When presented with a new sample set, Nutella accurately identified more than half of the new hypoglycemic samples, and correctly identified most normoglycemic and hyperglycemic samples from both samples sets. To the best of our knowledge, no other empirical studies have explicitly tested dogs' ability to generalize the odour of hypoglycemia across multiple samples from one individual, therefore it is difficult to determine how this level of performance should be evaluated. We do interpret these findings as promising, however, since Nutella showed some degree of success with a task arguably more difficult than those of Dehlinger et al. (2013) and Hardin et al. (2015).

In Experiment 3, Nutella's distribution of responses across sample sets provided inconclusive evidence as to whether she could generalize the odour of hypoglycemia across sample sets from two different individuals. Her level of sensitivity detecting the new hypoglycemic sample was only 45%, but her sensitivity detecting the hypoglycemic sample she had previously been trained to respond to dropped to 45% as well. Therefore, her performance on each individual sample set was comparable, making interpretation of her performance difficult. Generalization of a "hypoglycemic odour" across people is likely a difficult task because, although there appears to be some VOCs that may consistently signal hypoglycemia across different individuals (Neupane et al., 2016), the enormous variability between different people's VOC profiles (Phillips et al., 1999) may

result in individual VOCs profiles emerging during hypoglycemic events. This could explain Nutella's performance here, and may have also contributed to the poor performance observed by the dogs in Dehlinger et al. (2013). Furthermore, it is worth noting that during this experiment Nutella completed 80 consecutive trials in one work day. Researchers decided to conduct the test on only one work day to prevent the breath samples from being exposed to the air over multiple work days. Although Nutella displayed motivation to continue working throughout the test work day, she also began to show behaviours suggesting that she was tiring over time such as lying down between trials. It is possible that, with further training over multiple work days, Nutella could have learned to generalize the odour of hypoglycemia across samples from different people.

Hardin et al. (2015) claim that their results provide evidence that dogs can detect hypoglycemia, but we emphasize that Hardin et al. (2015) tested the dogs' ability to detect samples they had previously been trained to detect, therefore it was not a test of generalizability. As discussed by Elliker et al. (2014) it is possible for dogs to memorize many training odours, rather than detecting a common odour across all samples. However, Hardin et al. (2015) did illustrate that the dogs could consistently detect those odours that they had previously been trained on. Therefore, the current study, combined with the findings of Dehlinger et al. (2013) and Hardin et al. (2015) contributes important findings to the literature pertaining to DADs and has notable implications. Dogs can be trained to detect VOCs from individuals with type 1 diabetes, but, as our findings illustrate, the collective VOC profile of hypoglycemia may vary across individuals. Further research is required to determine whether dogs can detect a general odour of hypoglycemia across different people. Applied to the training of DADs, the findings of

the studies presented here suggest that dogs in training may be most successful if trained to detect hypoglycemia only with their "owner-to-be".

4.7 Acknowledgements

This research was supported by an IWK Health Centre Category A Research Grant, and an NSERC PGS-D to Catherine Reeve.

We would like to acknowledge Dr. Elizabeth McLaughlin, Dr. Vin LoLordo, and Dr. Ray Klein for their consultation and feedback. We would further like to acknowledge the dog owners who reliably brought their dogs to our lab each week, without whom this study would not have been possible. Finally, we thank the many student volunteers who contributed their time.

4.8 References

Acerini, C. (2016). The rise of technology in diabetes care. Not all that is new is necessarily better. *Pediatric Diabetes*, *17,* 168-173. *doi: 10.1111/pedi.12366*

Amann, A., de Lacy Costello, B., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., ... & Risby, T. (2014). The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *Journal of Breath Research*, *8*, 1-17. doi:10.1088/1752-7155/8/3/034001

Aye, T., Block, J., & Buckingham, B. (2010). Toward closing the loop: an update on insulin pumps and continuous glucose monitoring systems. *Endocrinology and Metabolism Clinics of North America*, *39*, 609-624. doi:10.1016/j.ecl.2010.05.005.

Chen, M., Daly, M., Williams, N., Williams, S., Williams, C., & Williams, G. (2000). Non-invasive detection of hypoglycaemia using a novel, fully biocompatible and patient friendly alarm system. *Bmj*, *321*(7276), 1565-1566.

Craven, B. A., Paterson, E. G., & Settles, G. S. (2009). The fluid dynamics of canine olfaction: unique nasal airflow patterns as an explanation of macrosmia. *Journal of The Royal Society Interface*, *7*, 933-943. doi:10.1098/rsif.2009.0490

Dehlinger, K., Tarnowski, K., House, J. L., Los, E., Hanavan, K., Bustamante, B., ... & Ward, W. K. (2013). Can trained dogs detect a hypoglycemic scent in patients with type 1 diabetes?. *Diabetes Care*, *36*(7), e98-e99. doi: 10.2337/dc12-2342

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 Urology*, *14*, 22.

Frier, B. M. (2004). Morbidity of hypoglycemia in type 1 diabetes. *Diabetes Research and Clinical Practice*, *65*, S47-S52. doi:10.1016/j.diabres.2004.07.008

Gadbois, S. & Reeve, C. (2016). The semiotic canine: scent processing dogs as research assistants in biomedical and environmental research. *Dog Behaviour, 3-2016,* 26-32. doi: 10.4454/db.v2i3.43

Gonder-Frederick, L., Nyer, M., Shepard, J. A., Vajda, K., & Clarke, W. (2011). Assessing fear of hypoglycemia in children with type 1 diabetes and their parents. *Diabetes Management*, *1*(6), 627-639. doi:10.2217/DMT.11.60.

Gonder-Frederick, L., Rice, P., Warren, D., Vajda, K., & Shepard, J. (2013). Diabetic alert dogs: a preliminary survey of current users. *Diabetes Care*, *36*(4), e47-e47. doi: 10.2337/dc12-1998

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*(4), 509-517. doi: 10.1007/s13300-015-0135-x

Jeitler, K., Horvath, K., Berghold, A., Gratzer, T. W., Neeser, K., Pieber, T. R., & Siebenhofer, A. (2008). Continuous subcutaneous insulin infusion versus multiple daily insulin injections in patients with diabetes mellitus: systematic review and meta-analysis. *Diabetologia, 51,* 941-951. doi: 10.1007/s00125-008-0974-3

Miklosi, A. (2007). *Dog Behaviour, Evolution, and Cognition*. Oxford University Press: Oxford, UK.

Minh, T. D. C., Blake, D. R., & Galassetti, P. R. (2012). The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes Research and Clinical Practice*, *97*, 195-205. doi:10.1016/j.diabres.2012.02.006

Neupane, S., Peverall, R., Richmond, G., Blaikie, T. P., Taylor, D., Hancock, G., & Evans, M. L. (2016). Exhaled breath isoprene rises during hypoglycemia in type 1 diabetes. *Diabetes Care*, *39*, e97-e98. doi: 10.2337/dc16-0461

Phillips, M., Herrera, J., Krishnan, S., Zain, M., Greenberg, J., & Cataneo, R. N. (1999). Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications*, *729*, 75-88.

Riemsma, R., Ramos, I. C., Birnie, R., Büyükkaramikli, N., Armstrong, N., Ryder, S., ... & Kleijnen, J. (2016). Integrated sensor-augmented pump therapy systems [the MiniMed® Paradigm™ Veo system and the Vibe™ and G4® PLATINUM CGM (continuous glucose monitoring) system] for managing blood glucose levels in type 1 diabetes: a systematic review and economic evaluation. *Health Technology Assessment, 20*. doi 10.3310/hta20170

Schabert, J., Browne, J. L., Mosely, K., & Speight, J. (2013). Social stigma in diabetes. *The Patient-Patient-Centered Outcomes Research*, *6*(1), 1-10. doi: 10.1007/s40271-012-0001-0

Tauveron, I., Delcourt, I., Desbiez, F., Somda, F., & Thiéblot, P. (2006). Canine detection of hypoglycaemic episodes whilst driving. *Diabetic medicine*, *23*(3), 335-335.

Schmidt, K., & Podmore, I. (2015). Current challenges in volatile organic compounds analysis as potential biomarkers of cancer. *Journal of Biomarkers*, *2015*, 1-16. doi: http://dx.doi.org/10.1155/2015/981458

Smith, D., Španěl, P., Fryer, A. A., Hanna, F., & Ferns, G. A. (2011). Can volatile compounds in exhaled breath be used to monitor control in diabetes mellitus?. *Journal of breath research*, *5*(2), 022001. doi:10.1088/1752-7155/5/2/022001

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. doi:10.1016/j.applanim.2005.07.009

Wells, D. L., Lawson, S. W., & Siriwardena, A. N. (2008). Canine responses to hypoglycemia in patients with type 1 diabetes. *The Journal of Alternative and Complementary Medicine*, *14*(10), 1235-1241. doi: 10.1089/acm.2008.0288

Wild, D., von Maltzahn, R., Brohan, E., Christensen, T., Clauson, P., & Gonder-Frederick, L. (2007). A critical review of the literature on fear of hypoglycemia in diabetes: Implications for diabetes management and patient education. *Patient Education and Counseling*, *68*(1), 10-15. doi:10.1016/j.pec.2007.05.003

Wong, J. C., Foster, N. C., Maahs, D. M., Raghinaru, D., Bergenstal, R. M., Ahmann, A. J., ... & Kleis, L. (2014). Real-time continuous glucose monitoring among participants in the T1D Exchange clinic registry. *Diabetes Care*, *37*(10), 2702-2709. doi: 10.2337/dc14-0303

CHAPTER 5: DISCUSSION

The overall goal of this dissertation was to make important methodological, procedural, and empirical contributions to the field of canine biomedical detection. After a brief summary of the studies included in this dissertation and their findings, I will discuss the contributions and limitations of these studies, followed by the implications of my research findings. Finally, I will discuss avenues for future research that can further the field of canine biomedical detection.

In Chapter 2, I presented a manuscript in which the authors and I describe a novel method for training dogs to detect breath samples. Using our saliency training procedure, four dogs with no previous sniffer dog training were successfully trained to detect and discriminate between breath samples from different people and between breath samples from one person at three different times of the day. Then, with the knowledge that future studies would require that breath samples be stored over time, we developed a method to coat cotton balls in silicone oil; a procedure that, at the molecular level, increases the volume and variety of VOCs contained in a breath sample. We then tested whether the use of silicone coated cotton balls affected dogs' ability to detect breath samples over time (Chapter 3). The results showed that for some dogs, the use of silicone coated cotton balls greatly improved their detectability of breath samples after one week of storage time. Finally, we applied what we learned from the studies presented in Chapters 2 and 3 to an empirical test of dogs' detection of hypoglycemia in people with Type 1 Diabetes. In Chapter 4, I presented studies where we collected breath samples, using silicone-coated cotton balls, from people with type 1 diabetes during normoglycemia, hypoglycemia, and hyperglycemia. Then in three experiments, we tested

dogs' ability to 1) discriminate between the samples, 2) detect and generalize the odour of hypoglycemia across multiple hypoglycemic samples from one person, and 3) generalize the odour of hypoglycemia across multiple samples from different people. The results showed that the dogs could discriminate between the samples at above chance levels, and one dog showed evidence of generalizing the odour of hypoglycemia across multiple samples from one person. Further research is required to determine whether dogs can generalize the odour of hypoglycemia across different people.

5.1 Contributions and Limitations

The research presented in this dissertation has the potential for making important contributions to the field of canine biomedical detection and alert. First, in Chapter 2, the use of the novel saliency training procedure allowed us to assess the dogs' ability to detect low saliency odours, their level of motivation, and their ability to complete repetitive tasks. This level of assessment is not always completed for dogs prior to tests of biomedical detection. Furthermore, empirical assessments of training programs for biomedical detection dogs are lacking, and those that have been conducted evaluated dogs' performance on a traditional line-up procedure (Jezierski et al., 2010; Schoon, 1996). Using a less common cued, 3AFC procedure, we reported the dogs' performance as they progressed through each stage of training and then evaluated the efficacy of the training program by testing their ability to detect and discriminate human breath samples. Therefore, we contributed the first full assessment of a cued, 3AFC training program for canine biomedical detection.

Although the saliency training program and procedure was the first of its kind, a limitation of the procedure was that the use of decreasing concentrations of steeped tea as

the target stimulus was not as precise as other studies assessing dogs' detection thresholds. For example, Hall et al., (2016) assessed dogs' sensitivity to decreasing dilutions of an odour using a liquid dilution olfactometer, and Walker et al. (2006) used mass air flow controllers to decrease the concentration of an odour and measure dogs' olfactory thresholds. More precise control over the concentration of the olfactory stimulus throughout the training program would have given us a more objective measure of each dogs' olfactory sensitivity.

A second contribution of this dissertation is the procedure and tests presented in Chapter 3. Here, we presented a method for coating cotton balls in silicone oil; a procedure that allows for greater absorption of VOCs in exhaled breath samples. Previous studies of canine biomedical detection report using silicone-coated cotton (Ehmann et al., 2012; McCulloch et al., 2006), however, to the best of my knowledge, no studies describe the procedures for coating the cotton. Furthermore, no researchers have examined how the use of silicone-coated cotton effects dogs' detection of breath samples. Therefore, we are the first in the field of canine biomedical detection to present and evaluate a method for increasing breath sample integrity.

A limitation of this study is that we did not assess how the VOC profile of the breath samples changed with increased storage time. Although the results showed that for two of the four dogs the use of silicone coated cotton increased the detectability of the breath samples over time compared to breath samples prepared with uncoated cotton, we cannot be positive that the VOC profile of a freshly donated breath sample was identical to the VOC profile of a four-week old sample. This is a concern for all studies that store biological samples, and one that should be addressed in the future (as discussed below).

Finally, in Chapter 4, we present tests of dogs' detection of hypoglycemia. Included in Chapter 4 is the first study to explicitly test dogs' ability to generalize the odour of hypoglycemia across multiple hypoglycemic samples. Furthermore, Experiments 2 and 3 in Chapter 4 use a Go/No-Go procedure, which is not commonly used in tests of canine biomedical detection. I argue, however, that the Go/No-Go procedure is more ecologically valid for hypoglycemia detection since deployed DADs will never be able to compare between glycemic states. Therefore, this test is one of the first of its kind in both the test of generalization and the procedure used.

The biggest limitation of the studies presented in this chapter was that, because of human participant attrition, the dogs were only exposed to a small number of hypoglycemic samples. Although studies have shown potential identifying VOCs of hypoglycemia (Minh et al., 2012; Neupane et al., 2016; Novak et al., 2007), individual differences in people's exhaled breath VOC content (Phillips et al., 1999) could have made the recognition of identifying compounds difficult. If the dogs were shown a larger number of samples from different people, the recognition of common VOCs across all the samples might have been facilitated. Furthermore, pre-exposure to stimuli increases dogs' sensitivity for a specific odour (Hall et al., 2016; Julien, 2009). Therefore, if the dogs had smelled a larger number of samples before testing, Koda might have shown evidence of generalizing the odour of hypoglycemia, and Nutella might have obtained a higher level of performance in the generalization tests. A second limitation of this chapter is that we did not use any analytical techniques to assess the contents of the breath samples. Therefore, although Nutella showed some ability to generalize the odour of

hypoglycemia across multiple samples from the same person, we cannot know what specific VOCs she was detecting.

5.2 Implications of Findings

The results of the studies presented in this dissertation have great potential for influencing the field of canine biomedical detection. More specifically, our data support methodological and procedural changes that could result in a more accurate assessment of dogs' ability to detect and discriminate biological stimuli. Moreover, as will be discussed, our test of dogs' detection and generalization of an odour associated with hypoglycemia could influence the training of DADs for people with type 1 diabetes. What follows is a discussion of the implications of the results presented in this dissertation.

After the saliency training program to train dogs to detect and discriminate breath samples, the dogs showed consistently above chance performance on the discrimination tests. The only other empirical evaluation of a training program for canine detection of biological samples reported poorer results than the results obtained here (Jezierski et al., 2010). As discussed in Chapter 2, it is believed that the higher level of performance reported here is attributable to several factors in the training program, including the strict performance criterion applied at each phase of training, the gradual transition from the tea stimulus to the breath stimulus, and the use of the cued, 3AFC procedure as opposed to the traditional five to seven stimulus line-up. We argue that use of the traditional five to seven stimulus line-up could actually be a less valid assessment of dogs' ability to detect and discriminate odours (Gadbois & Reeve, 2014). Therefore, given the results presented in Chapter 2, we advocate a reevaluation of the common procedures used in training and testing of canine biomedical detection.

In Chapter 3, we presented a procedure for coating cotton balls in silicone oil. Furthermore, the results of the Experiments conducted in Chapter 3 demonstrate that, for some dogs, the use of silicone-coated cotton for the collection of breath samples improved the detectability of breath samples with increased storage time. These results have important implications for future studies utilizing breath samples for canine detection. Given that two dogs showed higher detectability of breath samples collected using silicone-coated cotton, our results highlighted that, even though all our dogs were carefully selected, individual differences in olfactory sensitivity were apparent between dogs. Therefore, in studies of canine biomedical detection where biological samples are stored over time, a dog's poor performance could be the result of decreased sample integrity combined with the dog's olfactory sensitivity. Therefore, it would be wise for future studies to utilize silicone coated cotton balls to collect breath samples in order to ensure that the VOCs are contained and to further obtain an accurate assessment of dogs' ability to detect the VOCs.

Finally, the results of the studies conducted in Chapter 4 suggest that dogs can generalize the odour of hypoglycemia across multiple hypoglycemic events in one person, but their ability to generalize the odour of hypoglycemia across different people is unknown. This finding has important implications for the training of DADs. Based on the results of the current empirical studies, companies that train DADs would be safest to train dogs with the person they will be living with as early as possible. If dogs are not able to generalize the odour of hypoglycemia across different people, training a DAD in a facility and then selling the dog to a person with type 1 diabetes could result in the dog missing hypoglycemic events or not being able to detect that person's hypoglycemic

events at all. Without further empirical studies, however, the extent to which dogs can generalize the odour of hypoglycemia is unknown.

5.3 Future Directions

The results of the studies presented in this dissertation generated further questions that future work should address to help advance the field of canine biomedical detection. First, as discussed in Chapter 1, a dog's internal decision making bias can influence its proportion of responses to target and distractor stimuli. Moreover, a number of factors may influence a dog's decision making bias. Future work should examine how procedural variables during training can influence or change a dog's bias in detecting physiological changes such as diabetic hypoglycemia. For example, rewarding both positive and negative responses may result in close to ideal observer, whereas emphasis on rewarding only positive indications during training could result in a more liberal decision maker. Furthermore, given that hypoglycemic events can be rare (Ly et al., 2014), future research should examine how intermittent training with high prevalence trials, as demonstrated by Wolfe et al. (2013) could affect the training and maintenance of DADs' performance.

As discussed above, an area that requires careful consideration is how procedures for the collection and storage of biological samples effects the VOC content of the sample over time. With the use of analytical techniques such as GC-MS, future work should document the VOC profile of a sample at the time of donation, and then at different time points throughout storage time. Although Buszewski et al. (2012a) found a positive correlation between the concentration of two compounds in exhaled breath from people with lung cancer and dogs' positive indication of those samples, it would be

interesting to observe whether the dog's indications were correlated with the same compounds after the breath samples were stored for a period of time. Empirical studies regarding whether VOCs change in samples over time would confirm whether the disease-specific VOCs present at the time of sample preparation are maintained over time, and in what concentrations. This would have important implications for the training and testing of biomedical detection dogs.

To that point, future research should examine whether a person's VOC profile of hypoglycemia is consistent across multiple hypoglycemic events and if the VOCs change over developmental time. As discussed by Minh et al., (2012), if a person's diabetes worsens over time, corollary pathologies could result in changes to a person's VOC profile, including the VOCs present at the time of hypoglycemia. This knowledge could influence the training of DADs and support the need for regular re-training for accurate detection.

Finally, an area of research that will need particular attention in the future is whether dogs are able to generalize the odour of hypoglycemia across different people. This work would provide much needed empirical data that could contribute to more informed training of DADs, resulting in more accurate DADs for people with type 1 diabetes.

5.4 Conclusion

The field of canine biomedical detection is a burgeoning field of research, in which studies present promising evidence that dogs can contribute to disease diagnosis. As a new field of research, however, training and testing procedures not yet optimal, and the collection and storage of biological samples need to be considered carefully. This

dissertation demonstrated that changes in training and testing procedures could more accurately assess dogs' ability to detect VOCs associated with disease, and that the integrity of exhaled breath samples can be maintained over time with the use of silicone-coated cotton balls.

Within regards to DADs, the marketing of DADs is proceeding without solid empirical studies examining the intricacies of canine detection of hypoglycemia. The studies in this dissertation demonstrated that dogs can be trained to successfully discriminate between breath samples from people with type 1 diabetes donated during hypoglycemia, normoglycemic, and hyperglycemia. Furthermore, we provided evidence that dogs can generalize the odour of hypoglycemia in breath samples obtained across multiple hypoglycemic events from one person. Although results are inconclusive as to whether dogs can generalize the odour of hypoglycemia across multiple people, further work in the area will provide findings that can contribute greatly to the successful training of DADs.

For individuals considering acquiring a DAD, surveys of DAD owners (**Gonder-Frederick et al.**, 2013) and the results of the studies conducted here in Chapter 4 provide evidence that DADs have the potential to positively impact the lives of individuals with type 1 diabetes by detecting hypoglycemic events. It is important to keep in mind, however, that further research is required to determine what exactly DADs are detecting when they signal hypoglycemic events and how the training of DADs impacts their effectiveness in a real-world setting. Regardless of the performance of DADs, prospective owners should also recognize that simply owning a dog has positive effects on physical health and mental well being (O'Haire, 2010), and are a great addition to any

home. With future studies, there is no doubt that dogs can continue to contribute to the lives of people with type 1 diabetes.

5.5 References

Buszewski, B., Ligor, T., Jezierski, T., Wenda-Piesik, A., Walczak, M., & Rudnicka, J. (2012a). Identification of volatile lung cancer markers by gas chromatography-mass spectrometry: comparison with discrimination by canines. *Analytical and Bioanalytical Chemistry, 404,* 141-146. doi: 10.1007/s00216-012-6102-8

Ehmann, R., Boedeker, E., Friedrich, U., Sagert, J., Dippon, J., Friedel, G., & Walles, T. (2012). Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *European Respiratory Journal*, *39*, 669-676. doi: 10.1183/09031936.00051711

Gonder-Frederick, L., Rice, P., Warren, D., Vajda, K., & Shepard, J. (2013). Diabetic alert dogs: a preliminary survey of current users. *Diabetes Care*, *36*(4), e47-e47. doi: 10.2337/dc12-1998

Hall, N., Smith, D.W., & Wynne, C. (2016). Effect of odorant pre-exposure on domestic dogs' sensitivity on an odorant detection task. *Applied Animal Behaviour Science, 178,* 80-87. http://dx.doi.org/10.1016/j.applanim.2016.02.003

Jezierski, T., Górecka-Bruzda, Walczak, M., Świergiel, A.H., Chruszczewski, M.H., & Pearson, B.L. (2010). Operant conditioning of dogs (*Canis familiaris*) for identification of humans using scent lineup. *Animal Science Papers and Reports, 28,* 81-93.

Julien, M. M. (2009). Early scent association for the working canine: Creating a narcotics detection canine for the average canine handler. *Journal of Veterinary Behavior: Clinical Applications and Research*, *4*, 239.

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. doi: 10.1177/1534735405285096

Minh, T. D. C., Blake, D. R., & Galassetti, P. R. (2012). The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes Research and Clinical Practice*, *97*, 195-205. doi:10.1016/j.diabres.2012.02.006

Neupane, S., Peverall, R., Richmond, G., Blaikie, T. P., Taylor, D., Hancock, G., & Evans, M. L. (2016). Exhaled breath isoprene rises during hypoglycemia in type 1 diabetes. *Diabetes Care*, *39*, e97-e98. doi: 10.2337/dc16-0461

Novak, B.J., Blake, D.R., Meinardi, S., Rowland, F.S., Pontello, A., Cooper, D.M., & Galassetti, P.R. (2007). Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *PNAS, 104,* 15613-15618. www.pnas.org cgi doi 10.1073 pnas.0706533104

O'Haire, M. (2010). Companion animals and human health: Benefits, challenges, and the road ahead. *Journal of Veterinary Behavior: Clinical Applications and Research*, *5*, 226-234. doi:10.1016/j.jveb.2010.02.002

Phillips, M., Herrera, J., Krishnan, S., Zain, M., Greenberg, J., & Cataneo, R. N. (1999). Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications*, *729*, 75-88.

Schoon, G.A.A. (1996). Scent identification lineups by dogs (*Canis familiaris*): experimental design. *Applied Animal Behaviour Science, 49,* 257-267.

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. doi:10.1016/j.applanim.2005.07.009

REFERENCES

Acerini, C. (2016). The rise of technology in diabetes care. Not all that is new is necessarily better. *Pediatric diabetes, 17,* 168-173. doi: 10.1111/pedi.12366

Amann, A., & Smith, D. (Eds.). (2013). *Volatile Biomarkers: Non-Invasive Diagnosis in Physiology and Medicine*. Oxford, England: Elsevier.

Amann, A., de Lacy Costello, B., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., ... & Risby, T. (2014). The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *Journal of Breath Research*, *8*, 034001. doi:10.1088/1752-7155/8/3/034001

Amundsen, T., Sundstrøm, S., Buvik, T., Gederaas, O. A., & Haaverstad, R. (2014). Can dogs smell lung cancer? First study using exhaled breath and urine screening in unselected patients with suspected lung cancer. *Acta Oncologica*, *53*, 307-315. doi: 10.3109/0284186X.2013.819996

Angle, C., Waggoner, L. P., Ferrando, A., Haney, P., & Passler, T. (2016). Canine detection of the Volatilome: a review of implications for pathogen and disease detection. *Frontiers in Veterinary Science*, *3,* 1-7. doi: 10.3389/fvets.2016.00047

Atkins, P. & de Paula, J. (2006). *Physical Chemistry* (8th ed.). Great Britain: Oxford University Press.

Aye, T., Block, J., & Buckingham, B. (2010). Toward closing the loop: an update on insulin pumps and continuous glucose monitoring systems. *Endocrinology and Metabolism Clinics of North America*, *39*, 609-624. doi:10.1016/j.ecl.2010.05.005

Bade-White, P. A., & Obrzut, J. E. (2009). The neurocognitive effects of type 1 diabetes mellitus in children and young adults with and without hypoglycemia. *Journal of Developmental and Physical Disabilities*, *21*, 425-440. doi: 10.1007/s10882-009-9151-y

Balseiro, S. C., & Correia, H. R. (2006). Is olfactory detection of human cancer by dogs based on major histocompatibility complex-dependent odour components?–A possible cure and a precocious diagnosis of cancer. *Medical Hypotheses*, *66*, 270-272. doi:10.1016/j.mehy.2005.08.027

Becker, D. J., & Ryan, C. M. (2000). Hypoglycemia: a complication of diabetes therapy in children. *Trends in Endocrinology & Metabolism*, *11*, 198-202.

Bir, D. (2000). Partition coefficient calculation of selected terpenes and low molecular weight solvents between tall oil fatty acid and air and polydimethyl siloxane oil and air. *Journal of the American Oil Chemists' Society*, *77*, 163-169.

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*, e7396. doi: 10.1136/bmj.e7396

Brands, A. M., Biessels, G. J., De Haan, E. H., Kappelle, L. J., & Kessels, R. P. (2005). The effects of type 1 diabetes on cognitive performance. *Diabetes Care*, *28*, 726-735.

Brisbin, I. L., & Austad, S. N. (1991). Testing the individual odour theory of canine olfaction. *Animal Behaviour*, *42*, 63-69. doi: https://doi.org/10.1016/S0003-3472(05)80606-2

Buckingham, B., Block, J., Burdick, J., Kalajian, A., Kollman, C., Choy, M., ... & Chase, P. (2005). Response to nocturnal alarms using a real-time glucose sensor. *Diabetes Technology & Therapeutics*, *7*, 440-447.

Buckingham, B., Wilson, D. M., Lecher, T., Hanas, R., Kaiserman, K., & Cameron, F. (2008). Duration of nocturnal hypoglycemia before seizures. *Diabetes Care*, *31*, 2110-2112. doi: 10.2337/dc08-0863

Burak, R.C., & Liang, J. (1987). The early detection of cancer in the primary-care setting: factors associated with the acceptance and completion of recommended procedures. *Preventive Medicine, 16,* 739-751. doi: http://dx.doi.org/10.1016/0091-7435(87)90014-4

Buszewski, B., Ligor, T., Jezierski, T., Wenda-Piesik, A., Walczak, M., & Rudnicka, J. (2012a). Identification of volatile lung cancer markers by gas chromatography–mass spectrometry: comparison with discrimination by canines. *Analytical and Bioanalytical Chemistry*, *404*, 141-146. doi: 10.1007/s00216-012-6102-8

Buszewski, B., Rudnicka, J., Ligor, T., Walczak, M., Jezierski, T., & Amann, A. (2012b). Analytical and unconventional methods of cancer detection using odor. *Trends in Analytical Chemistry*, *38*, 1-12. doi:http://dx.doi.org/10.1016/j.trac.2012.03.019

Chen, M., Daly, M., Williams, N., Williams, S., Williams, C., & Williams, G. (2000). Non-invasive detection of hypoglycaemia using a novel, fully biocompatible and patient friendly alarm system. *BMJ*, *321*, 1565-1566.

Chikazoe, J., Jimura, K., Asari, T., Yamashita, K. I., Morimoto, H., Hirose, S., ... & Konishi, S. (2009). Functional dissociation in right inferior frontal cortex during performance of go/no-go task. *Cerebral Cortex*, *19*, 146-152. doi:10.1093/cercor/bhn065

Church, J., & Williams, H. (2001). Another sniffer dog for the clinic?. *The Lancet*, *358*, 930.

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. doi: 10.1016/j.eururo.2010.10.006

Corradi, M. & Mutti, A. (2013). Exhaled breath analysis in occupational medicine. In A. Amann & D. Smith (Eds.), *Volatile Biomarkers. Non-invasive Diagnosis in Physiology and Medicine* (117-125). Oxford, England: Elsevier.

Craven, B. A., Paterson, E. G., & Settles, G. S. (2009). The fluid dynamics of canine olfaction: unique nasal airflow patterns as an explanation of macrosmia. *Journal of The Royal Society Interface*, *7,* 933-943. doi: 10.1098/rsif.2009.0490

Curran, A. M., Rabin, S. I., Prada, P. A., & Furton, K. G. (2005). Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *Journal of Chemical Ecology*, *31*, 1607-1619. doi: 10.1007/s10886-005-5801-4

Davis, S. N., Shavers, C., Mosqueda-Garcia, R., & Costa, F. (1997). Effects of differing antecedent hypoglycemia on subsequent counterregulation in normal humans. *Diabetes*, *46*, 1328-1335.

De Beaufort, C. (2006). Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. *Diabetic Medicine: a Journal of the British Diabetic Association*, *23*, 857-866. doi: 10.1111/j.1464-5491.2006.01925.x

de Lacy Costello, B., Amann, A., Al-Kateb, H., Flynn, C., Filipiak, W., Khalid, T., ... & Ratcliffe, N. M. (2014). A review of the volatiles from the healthy human body. *Journal of Breath Research*, *8*, 014001. doi:10.1088/1752-7155/8/1/014001

Dehlinger, K., Tarnowski, K., House, J. L., Los, E., Hanavan, K., Bustamante, B., ... & Ward, W. K. (2013). Can trained dogs detect a hypoglycemic scent in patients with type 1 diabetes?. *Diabetes Care*, *36*, e98-e99. doi: 10.2337/dc12-2342

DeSantis, C., Ma, J., Bryan, L., & Jemal, A. (2014). Breast cancer statistics, 2013. *CA: A Cancer Journal for Clinicians, 64,* 52- 62. doi: 10.3322/caac.21203

Ehmann, R., Boedeker, E., Friedrich, U., Sagert, J., Dippon, J., Friedel, G., & Walles, T. (2012). Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *European Respiratory Journal*, *39*, 669-676.

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 urology*, *14*, 22. doi: http://www.biomedcentral.com/1471-2490/14/22

Evans, K. K., Birdwell, R. L., & Wolfe, J. M. (2013). If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PLoS One*, *8*, e64366. doi:10.1371/journal.pone.0064366

Frier, B. M. (2004). Morbidity of hypoglycemia in type 1 diabetes. *Diabetes Research and Clinical Practice*, *65*, S47-S52. doi:10.1016/j.diabres.2004.07.008

Frier, B.M (2008). How hypoglycaemia can affect the life of a person with diabetes. *Diabetes/Metabolism Research and Reviews, 24,* 87-92. doi: 10.1002/dmrr.796

Gadbois S., & Reeve, C. (2014). Canine olfaction: Scent, sign, and situation. In: *Horowitz, A. (Ed.), Domestic dog cognition and behavior (*pp 3-29). New York, NY: Springer.

Gadbois, S., & Reeve, C. (2016). The semiotic canine: scent processing dogs as research assistants in biomedical and environmental research. *DOG BEHAVIOR*, *2*, 26-32. doi 10.4454/db.v2i3.43

Galassetti, P. R., Novak, B., Nemet, D., Rose-Gottron, C., Cooper, D. M., Meinardi, S., ... & Blake, D. R. (2005). Breath ethanol and acetone as indicators of serum glucose levels: an initial report. *Diabetes Technology & Therapeutics*, *7*, 115-123. doi: https://doi.org/10.1089/dia.2005.7.115

Gibson, E.J. (1969). *Principles of perceptual learning and development.* New York, NY: Appleton-Century-Crofts.

Gonder-Frederick, L., Nyer, M., Shepard, J. A., Vajda, K., & Clarke, W. (2011). Assessing fear of hypoglycemia in children with type 1 diabetes and their parents. *Diabetes Management*, *1*, 627-639. doi:10.2217/DMT.11.60

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. doi: 10.2337/dc12-1998

Gonder-Frederick, L., Zrebiec, J., Bauchowitz, A., Lee, J., Cox, D., Ritterband, L., ... & Clarke, W. (2008). Detection of hypoglycemia by children with type 1 diabetes 6 to 11 years of age and their parents: a field study. *Pediatrics*, *121*, e489-e495. doi: 10.1542/peds.2007-0808

Gordon, R. T., Schatz, C. B., Myers, L. J., Kosty, M., Gonczy, C., Kroener, J., ... & Arthur, N. (2008). The use of canines in the detection of human cancers. *The Journal of Alternative and Complementary Medicine*, *14*, 61-67. doi: 10.1089/acm.2006.6408

Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Gschwend, S., Ryan, C., Atchison, J., Arslanian, S., & Becker, D. (1995). Effects of acute hyperglycemia on mental efficiency and counterregulatory hormones in adolescents with insulin-dependent diabetes mellitus. *The Journal of Pediatrics*, *126*, 178-184.

Hall, N. J., Smith, D. W., & Wynne, C. D. (2013). Training domestic dogs (Canis lupus familiaris) on a novel discrete trials odor-detection task. *Learning and Motivation*, *44*, 218-228. http://dx.doi.org/10.1016/j.lmot.2013.02.004

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*(4), 509-517. doi: 10.1007/s13300-015-0135-x

Harrington, F.H., & Asa, C.S. (2010). Wolf communication. In L.D. Mech., & L. Boitani (Eds.), *Wolves: Behavior, Ecology, and Conservation* (pp. 66-103). Chicago, IL: University of Chicago Press.

Helton, W. S. (2009). Canine ergonomics: Introduction to the new science of working dogs. In W. S. Helton (Ed.), *Canine ergonomics. The science of working dogs* (pp. 1–16). Boca Raton, FL: Taylor and Francis Group.

Hershey, T., Perantie, D. C., Warren, S. L., Zimmerman, E. C., Sadler, M., & White, N. H. (2005). Frequency and timing of severe hypoglycemia affects spatial memory in children with type 1 diabetes. *Diabetes Care*, *28*, 2372-2377.

Horvath, G., Andersson, H., & Nemes, S. (2013). Cancer odor in the blood of ovarian cancer patients: a retrospective study of detection by dogs during treatment, 3 and 6 months afterward. *BMC cancer*, *13*, 396. doi: http://www.biomedcentral.com/1471-2407/13/396

Horvath, G., Andersson, H., & Paulsson, G. (2010). Characteristic odour in the blood reveals ovarian carcinoma. *BMC cancer*, *10*, 643. doi: http://www.biomedcentral.com/1471-2407/10/643

Horvath, G., Järverud, G. A. K., Järverud, S., & Horváth, I. (2008). Human ovarian carcinomas detected by specific odor. *Integrative cancer therapies*, *7*, 76-80. doi: 10.1177/1534735408319058

Howsmon, D., & Bequette, B. W. (2015). Hypo-and hyperglycemic alarms: devices and algorithms. *Journal of diabetes science and technology*, *9*, 1126-1137. doi: 10.1177/1932296815583507

Jeitler, K., Horvath, K., Berghold, A., Gratzer, T. W., Neeser, K., Pieber, T. R., & Siebenhofer, A. (2008). Continuous subcutaneous insulin infusion versus multiple daily insulin injections in patients with diabetes mellitus: systematic review and meta-analysis. *Diabetologia, 51,* 941-951. doi: 10.1007/s00125-008-0974-3

Jezierski, T., Gorecka-Bruzda, A., Walczak, M., Swiergiel, A. H., Chruszczewski, M. H., & Pearson, B. L. (2010). Operant conditioning of dogs (Canis familiaris) for identification of humans using scent lineup. *Animal Science Papers and Reports*, *1*, 81-93.

Jezierski, T., Walczak, M., & Górecka, A. (2008). Information-seeking behaviour of sniffer dogs during match-to-sample training in the scent lineup. *Polish Psychological Bulletin*, *39*, 71-80. doi: 10.2478/v10059-008-0010-y

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*, 027001. doi:10.1088/1752-7155/9/2/027001

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. doi: http://dx.doi.org/10.1016/j.applanim.2017.01.001

Kaminski, J., & Nitzschner, M. (2013). Do dogs get the point? A review of dog–human communication ability. *Learning and Motivation*, *44*, 294-302. http://dx.doi.org/10.1016/j.lmot.2013.05.001

Kazdin, A. E. (2011). *Single-Case Research Designs: Methods for Clinical and Applied Settings*. New York, NY: Oxford University Press.

Kingdom F.A.A. & Prins N. (2016). *Psychophysics. A practical introduction*. London, England: Elsevier.

Klonoff, D. C. (2005). Continuous glucose monitoring. *Diabetes care*, *28*, 1231-1239.

Koivusalo, M., Vermeiren, C., Yuen, J., Reeve, C., Gadbois, S., & Katz, K. (2017). Canine scent detection as a tool to distinguish meticillin-resistant Staphylococcus aureus. *Journal of Hospital Infection*, *96*, 93-95. doi: http://dx.doi.org/10.1016/j.jhin.2017.03.005

Kusano, M., Mendez, E., & Furton, K. G. (2011). Development of headspace SPME method for analysis of volatile organic compounds present in human biological specimens. *Analytical and bioanalytical chemistry*, *400*, 1817. doi: 10.1007/s00216-011-4950-2

Kwon, Y. (2001). Partition and distribution coefficients. In *Handbook of Essential Pharmacokinetics, Pharmacodynamics, and Drug Metabolism for Industrial Scientists* (pp. 35-71). New York, NY: Kluwer Academic/Plenum Publishers.

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. doi:10.1093/jhered/esn057

Lit, L. (2009). Evaluation learning tasks commonly applied in detection dog training. In W. S. Helton (Ed.), *Canine ergonomics. The science of working dogs* (pp. 99–114). Boca Raton, FL: Taylor and Francis Group.

Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, *52*, 415-419.

Ly, T.T., Maahs, D.M., Rewers, A., Dunger, D., Oduwole, A., & Jones, T.W. (2014). Assessment and management of hypoglycemia in children and adolescents with diabetes. *Pediatric Diabetes, 15,* 180-192.

Machlis, L., Dodd, P. W. D., & Fentress, J. C. (1985). The pooling fallacy: problems arising when individuals contribute more than one observation to the data set. *Ethology*, *68*, 201-214.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. New Jersey: Lawrence Erlbaum Associates Inc.

Marcus, D.A., & Bhowmick, A. (2013). Survey of migraine sufferers with dogs to evaluate for canine migraine-alerting behaviors. *The Journal of Alternative and Complementary Medicine, 19,* 501-508.

Matkya, K.A. (2002) Sweet dreams? – nocturnal hypoglycemia in children with type 1 diabetes. *Pediatric Diabetes, 3*, 74-81. doi: 10.1034/j.1399-5448.2002.30203.x

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. doi: 10.1177/1534735405285096

Miklosi, A. (2007). *Dog Behaviour, Evolution, and Cognition*. Oxford University Press: Oxford, UK.

Minh, T. D. C., Blake, D. R., & Galassetti, P. R. (2012). The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes Research and Clinical Practice*, *97*, 195-205. doi: 10.1016/j.diabres.2012.02.006

13

Modak, A.S. (2013). An Update on C-Breath tests: The transition to acceptability into clinical practice. In A. Amann & D. Smith (Eds.), *Volatile Biomarkers. Non-invasive Diagnosis in Physiology and Medicine* (245-262). Oxford, England: Elsevier.

Morgan, D.L. & Morgan, R.K. (2009). *Single Case Research Methods*. Thousand Oaks, CA: SAGE publications.

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. doi:10.1016/j.jveb.2010.01.002

Neupane, S., Peverall, R., Richmond, G., Blaikie, T. P., Taylor, D., Hancock, G., & Evans, M. L. (2016). Exhaled breath isoprene rises during hypoglycemia in type 1 diabetes. *Diabetes Care*, *39*, e97-e98. doi: 10.2337/dc16-0461

Novak, B. J., Blake, D. R., Meinardi, S., Rowland, F. S., Pontello, A., Cooper, D. M., & Galassetti, P. R. (2007). Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *Proceedings of the National Academy of Sciences*, *104*, 15613-15618. doi: www.pnas.org cgi doi 10.1073 pnas.0706533104

O'Connell, S. M., Cooper, M. N., Bulsara, M. K., Davis, E. A., & Jones, T. W. (2011). Reducing rates of severe hypoglycemia in a population-based cohort of children and adolescents with type 1 diabetes over the decade 2000–2009. *Diabetes Care*, *34*, 2379-2380. doi: 10.2337/dc11-0748

O'Haire, M. (2010). Companion animals and human health: Benefits, challenges, and the road ahead. *Journal of Veterinary Behavior: Clinical Applications and Research*, *5*, 226-234. doi:10.1016/j.jveb.2010.02.002

Ovalle, F., Fanelli, C. G., Paramore, D. S., Hershey, T., Craft, S., & Cryer, P. E. (1998). Brief twice-weekly episodes of hypoglycemia reduce detection of clinical hypoglycemia in type 1 diabetes mellitus. *Diabetes*, *47*, 1472-1479.

Patel, M.J., Popat, S.C., & Deshusses, M.A. (2017). Determination and correlation of the partition coefficients of 48 volatile organic and environmentally relevant compounds between air and silicone oil. *Chemical Engineering Journal, 310,* 72-78. doi: http://dx.doi.org/10.1016/j.cej.2016.10.086

Pearsall, M. D., & Verbruggen, H. (1982). *Scent. Training to Track, Search, and Rescue*. Loveland, CO: Alpine Publications.

Phillips, M. (1992). Breath tests in medicine. *Scientific American*, *267*, 74-79.

Phillips, M., Herrera, J., Krishnan, S., Zain, M., Greenberg, J., & Cataneo, R. N. (1999). Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications*, *729*, 75-88.

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. doi:10.1016/j.applanim.2004.04.008

Riemsma, R., Ramos, I. C., Birnie, R., Büyükkaramikli, N., Armstrong, N., Ryder, S., ... & Kleijnen, J. (2016). Integrated sensor-augmented pump therapy systems [the MiniMed® Paradigm™ Veo system and the Vibe™ and G4® PLATINUM CGM (continuous glucose monitoring) system] for managing blood glucose levels in type 1 diabetes: a systematic review and economic evaluation. *Health Technology Assessment, 20*, 1 - 288. doi: 10.3310/hta20170

Rooney, N. J., Bradshaw, J. W., & Almey, H. (2004). Attributes of specialist search dogs—a questionnaire survey of UK dog handlers and trainers. *Journal of Forensic Science*, *49*, 1-7.

Ross, B.M., & Esarik, A. (2013). The analysis of oral air by selected ion flow tube mass spectrometry using indole and methylindole as examples. In A. Amann & D. Smith (Eds.), *Volatile Biomarkers. Non-invasive Diagnosis in Physiology and Medicine* (77-88). Oxford, England: Elsevier.

Rudnicka, J., Walczak, M., Kowalkowski, T., Jezierski, T., & Buszewski, B. (2014). Determination of volatile organic compounds as potential markers of lung cancer by gas chromatography–mass spectrometry versus trained dogs. *Sensors and Actuators B: Chemical*, *202*, 615-621. doi: 10.1183/09031936.00051711

Schabert, J., Browne, J. L., Mosely, K., & Speight, J. (2013). Social stigma in diabetes. *The Patient-Patient-Centered Outcomes Research*, *6*, 1-10. doi: 10.1007/s40271-012-0001-0

Schmidt, K., & Podmore, I. (2015). Current challenges in volatile organic compounds analysis as potential biomarkers of cancer. *Journal of biomarkers*, *2015,* 1-16. http://dx.doi.org/10.1155/2015/981458

Schoon, G. A. A. (1996). Scent identification lineups by dogs (Canis familiaris): experimental design and forensic application. *Applied Animal Behaviour Science*, *49*, 257-267.

Schoon, G. A. A. (2005). The effect of the ageing of crime scene objects on the results of scent identification line-ups using trained dogs. *Forensic Science International*, *147*, 43-47. doi:10.1016/j.forsciint.2004.04.080

Schoon, G.A. & Haak, R. (2002). *K9 suspect discrimination: Training and practicing scent identification line- ups*. Alberta, Canada: Detselig Enterprises.

Settle, R. H., Sommerville, B. A., McCormick, J., & Broom, D. M. (1994). Human scent matching using specially trained dogs. *Animal Behaviour*, *48*, 1443-1448.

Settles, G. S., Kester, D. A., & Dodson-Dreibelbis, L. J. (2003). The external aerodynamics of canine olfaction. In *Sensors and sensing in biology and engineering* (pp. 323-335). Vienna: Springer.

Shirasu, M., & Touhara, K. (2011). The scent of disease: volatile organic compounds of the human body related to disease and disorder. *Journal of Biochemistry*, *150*, 257-266. doi: 10.1093/jb/mvr090

Shivers, J. P., Mackowiak, L., Anhalt, H., & Zisser, H. (2013). "Turn it off!": diabetes device alarm fatigue considerations for the present and the future. *Journal of Diabetes Science and Technology*, *7*, 789-794.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: a cancer journal for clinicians*, *66*, 7-30. doi: 10.3322/caac.21332

Siegel, R., DeSantis, C., & Jemal, A. (2014). Colorectal cancer statistics, 2014. *CA: A Cancer Journal for Clinicians*, *64*, 104-117. doi: 10.3322/caac.21220

Smith, D., Španěl, P., Fryer, A. A., Hanna, F., & Ferns, G. A. (2011). Can volatile compounds in exhaled breath be used to monitor control in diabetes mellitus? *Journal of Breath Research*, *5*, 022001. doi:10.1088/1752-7155/5/2/022001

Solga, S.F., and Risby, T.H. (2013). Issues and challenges in human breath research: perspectives from our experience. In A. Amann & D. Smith (Eds). *Volatile Biomarkers: Non-Invasive Diagnosis in Physiology and Medicine* (pp 19-24). Oxford, England: Elsevier.

Sonoda, H., Kohnoe, S., Yamazato, T., Satoh, Y., Morizono, G., Shikata, K., ... & Inoue, F. (2011). Colorectal cancer screening with odour material by canine scent detection. *Gut*, *60,* 814-819. doi:10.1136/gut.2010.218305

Stamey, T. A., Caldwell, M., McNEAL, J. E., Nolley, R., Hemenez, M., & Downs, J. (2004). The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? *The Journal of Urology*, *172*, 1297-1301. doi: 10.1097/01.ju.0000139993.51181.5d

Strong, V., Brown, S.W., & Walker, R. (1999). Seizure alert dogs – fact or fiction? *Seizure, 8,* 62-65.

Sun, X., Shao, K., & Wang, T. (2016). Detection of volatile organic compounds (VOCs) from exhaled breath as noninvasive methods for cancer diagnosis. *Analytical and Bioanalytical Chemistry*, *408*, 2759-2780. doi: 10.1007/s00216-015-9200-6

Szulejko, J. E., McCulloch, M., Jackson, J., McKee, D. L., Walker, J. C., & Solouki, T. (2010). Evidence for cancer biomarkers in exhaled breath. *IEEE Sensors Journal*, *10*, 185-210.

Tauveron, I., Delcourt, I., Desbiez, F., Somda, F., & Thiéblot, P. (2006). Canine detection of hypoglycaemic episodes whilst driving. *Diabetic Medicine*, *23*, 335-335.

Udell, M. A., & Wynne, C. D. (2008). A review of domestic dogs' (*Canis Familiaris*) human-like behaviors: or why behavior analysts should stop worrying and love their dogs. *Journal of the experimental analysis of behavior*, *89*, 247-261. doi: 10.1901/jeab.2008.89-247

Unger, J., & Parkin, C. (2011). Recognition, prevention, and proactive management of hypoglycemia in patients with type 1 diabetes mellitus. *Postgraduate medicine*, *123*, 71-80.

Vyplelová, P., Vokálek, V., Pinc, L., Pacáková, Z., Bartoš, L., Santariová, M., & Čapková, Z. (2014). Individual human odor fallout as detected by trained canines. *Forensic Science International*, *234*, 13-15. doi: http://dx.doi.org/10.1016/j.forsciint.2013.10.018

Waggoner, L. P., Jones, M. H., Williams, M., Johnston, J. M., Edge, C. C., & Petrousky, J. A. (1998, December). Effects of extraneous odors on canine detection. In *Enabling Technologies for Law Enforcement and Security* (pp. 355-362). International Society for Optics and Photonics. doi:10.1117/12.335008

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. doi:10.1016/j.applanim.2005.07.009

Weinzimer, S., Xing, D., Tansey, M., Fiallo-Scharer, R., Mauras, N., Wysocki, T., ... & Ruedy, K. (2009). Prolonged use of continuous glucose monitors in children with type 1 diabetes on continuous subcutaneous insulin infusion or intensive multiple-daily injection therapy. *Pediatric Diabetes*, *10*, 91-96. doi: 10.1111/j.1399-5448.2008.00476.x

Wells, D. L., Lawson, S. W., & Siriwardena, A. N. (2008). Canine responses to hypoglycemia in patients with type 1 diabetes. *The Journal of Alternative and Complementary Medicine*, *14*, 1235-1241. doi: 10.1089/acm.2008.0288

Wild, D., von Maltzahn, R., Brohan, E., Christensen, T., Clauson, P., & Gonder-Frederick, L. (2007). A critical review of the literature on fear of hypoglycemia in diabetes: Implications for diabetes management and patient education. *Patient Education and Counseling*, *68*, 10-15. doi:10.1016/j.pec.2007.05.003

Williams, H., & Pembroke, A. (1989). Sniffer dogs in the melanoma clinic? *The Lancet*, *333*, 734.

Willis, C. M., Church, S. M., Guest, C. M., Cook, W. A., McCarthy, N., Bransbury, A. J., ... & Church, J. C. (2004). Olfactory detection of human bladder cancer by dogs: proof of principle study. *BMJ*, *329*, 712 – 714.

Wilson, A. D. (2015). Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. *Metabolites*, *5*, 140-163. doi:10.3390/metabo5010140

Wolfe, J.M., Brunelli, D.N., Rubinstein, J., Horowitz, T.S. (2013). Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *Journal of Vision, 13,* 1-9. doi: 10.1167/13.3.33

Wong, J. C., Foster, N. C., Maahs, D. M., Raghinaru, D., Bergenstal, R. M., Ahmann, A. J., ... & Kleis, L. (2014). Real-time continuous glucose monitoring among participants in the T1D Exchange clinic registry. *Diabetes Care*, *37*(10), 2702-2709. doi: 10.2337/dc14-0303

Zijlstra, E., Heise, T., Nosek, L., Heinemann, L., & Heckermann, S. (2013). Continuous glucose monitoring: quality of hypoglycaemia detection. *Diabetes, Obesity and Metabolism*, *15*, 130-135.