

The State of Assessment in Human-Animal Interaction Research

Kerri E. Rodriguez¹, Noémie A. Guérin¹, Robin L. Gabriels²,
James A. Serpell³, Pamela J. Schreiner⁴, & Marguerite E. O’Haire¹

¹ Center for the Human-Animal Bond, Department of Comparative Pathobiology,
College of Veterinary Medicine, Purdue University, IN, USA,

² Departments of Psychiatry and Pediatrics, School of Medicine, University of
Colorado Anschutz Medical Campus, CO, USA,

³ Center for the Interaction of Animals and Society, Department of Clinical
Sciences & Advanced Medicine, University of Pennsylvania School of
Veterinary Medicine, PA, USA, &

⁴ Center to Study Human-Animal Relationships and Environments, School of
Public Health, University of Minnesota, MN, USA

There is a growing body of research in human-animal interaction (HAI), but the field is often criticized for its lack of methodological rigor and heterogeneous outcome assessment. In addition, there is a scarcity of valid and reliable assessment tools specific to measuring constructs directly related to HAI. As the field continues to develop, there is a critical need for the consistent use of standardized, well-validated, and appropriate assessment tools. The goal of this paper is to provide an overview of the state of assessment in HAI while identifying current needs for method standardization, development, or refinement. Specifically, we review in detail the use of (1) questionnaires, (2) physiological measures, and (3) behavior assessment tools in the field of HAI while offering examples of successful application, describing methodological limitations, and providing recommendations for use in HAI research.

keywords: Human-Animal Interaction; Assessment; Animal-Assisted Intervention; Methodology; Questionnaires; Physiology; Behavior

Correspondence concerning this article should be addressed to Marguerite E. O’Haire, Center for the Human-Animal Bond, Department of Comparative Pathobiology, College of Veterinary medicine, Purdue University, 725 Harrison Street, West Lafayette, IN 47907. Email: mohaire@purdue.edu

The field of human-animal interaction (HAI) is a growing area of research aiming to quantify the bi-directional effects of the human-animal relationship on health and wellbeing. Early pioneering research in HAI established a scientific foundation to begin to examine the effects that animals may have on human lives, and as the field of HAI continues to grow it will require a coordinated effort from researchers

to strategically advance and refine the field (Fine, Tedeschi, & Elvolve, 2015; Kazdin, 2015; McCune et al., 2014). Assessment of outcomes remains one of the most critical areas in need of further refinement and standardization to develop empirical rigor in the field of HAI (Fine et al., 2015; Herzog, 2015; McCune et al., 2014; Uvnäs-Moberg, Handlin, & Petersson, 2011; Wilson & Netting, 2012).

In this paper, we refer to assessment tools as methods of gathering and interpreting data aiming to quantify the relationship between one or more constructs of interest. Examples of constructs examined in HAI can be individual attributes in humans (e.g., extraversion), attributes in animals (e.g., affiliative behaviors), or attributes of the relationship between a human and an animal (e.g., attitude towards pets). Constructs themselves are often not directly observable, but they can be quantified through the measurement of a set of behaviors or variables (Cronbach & Meehl, 1955). In turn, theories are refined by testing hypotheses about the relationship between various constructs of interest. In order to draw scientifically valid conclusions, constructs must be adequately measured by the assessment tool with regards to validity (e.g., the extent to which a tool assesses what it intends) and reliability (e.g., the ability of the tool to produce consistent results across raters and time, Cronbach & Meehl, 1955; Kimberlin & Winterstein, 2008).

The scarcity of valid and reliable assessment tools is repeatedly discussed in critical evaluations of HAI research (Esposito, McCune, Griffin, & Maholmes, 2011; Herzog, 2015; McCune et al., 2014). In addition, concerns regarding the practice of using inconsistent assessment and its implications for slowing the advancement of the field of HAI are often pointed out by systematic literature reviews on animal-assisted interventions (AAI) for various populations, including patients with intellectual disabilities (Maber-Aleksandrowicz, Avent, & Hassiotis, 2016), dementia (Bernabei et al., 2013), and autism spectrum disorder (Davis et al., 2015; O'Haire, 2016). Specifically, there is a large amount of heterogeneity in how outcomes are assessed. These variations are so large that findings from systematic reviews are often limited to speculative interpretation, with

results that are not comparable across studies. In fact, a systematic review of randomized controlled trials on the effects of animal-assisted therapy (AAT) could not perform a meta-analysis partly due to this problem (Kamioka et al., 2014). Reliable and consistent assessment is necessary to build an empirical evidence base so that studies can be compared and replicated (Asendorpf et al., 2013).

The goals of this paper are to provide an overview of the importance of assessment tools both broadly and in the context of HAI research and to provide HAI researchers with information necessary to select appropriate assessment methods for their research. Specifically, this paper provides an overview of the state of assessment in HAI research while describing (1) current common assessment techniques in HAI (2) potential measurement issues in their reliability and validity and (3) future directions towards refinement and selection of measurement in HAI. This discussion will focus on three frequently used broad categories of assessment: (I) questionnaires, (II) physiological measures, and (III) behavioral observation.

Questionnaires

Questionnaires can include self-report or report by others (e.g. caregivers, teachers, clinicians), and can be used to collect information about a wide variety of subjective experiences such as attitudes, emotions, thoughts, beliefs and perceived mental and physical health. As many psychological constructs of interest to HAI researchers are true measures of self-experience (e.g. life satisfaction, pain, or emotional affect), self-report methods in the form of both qualitative and quantitative assessment are essential to the field. With relatively low costs and resources needed to carry out successful data collection,

questionnaires are the most frequently used form of outcome assessment in HAI (O'Haire et al., 2018).

Developing and adapting questionnaires for use in HAI

A questionnaire is considered standardized when all questionnaire items are asked in the same way and order to all study participants, allowing for replication and consistency across studies. In the early years of HAI research, standardized questionnaires for constructs specific to HAI were scarce. Therefore, researchers commonly established new questionnaires or adapted previously-existing questionnaires for the needs of their individual studies. This resulted in a large pool of heterogeneous tools to study the human-animal relationship (McCune et al., 2014). A 2012 inventory of questionnaires used in HAI research identified 140 scales in total, including an array of measures of attitudes towards animals, attachment to animals, and bonding scales (Wilson & Netting, 2012). While these questionnaires may be pertinent to HAI-specific constructs, a large majority of these measures lack evaluations of construct validity, or the extent to which a questionnaire is able to accurately measure the construct that it is designed to measure (Cronbach & Meehl, 1955; May, Seivert, Cano, Casey, & Johnson, 2016).

To aid in establishing construct validity in a new tool, researchers developing or adapting questionnaires for use in HAI research should consider the intended theoretical framework in mind. For example, one of the most popular tools for assessing the human-canine relationship is the Monash Dog-Owner Relationship Scale (MDORS; Dwyer, Bennett, & Coleman, 2006). The development of the MDORS was based on Social Exchange Theory which posits that a relationship will continue to persist if the benefits of the interactions outweigh the costs

(Blau, 1964; Netting, Wilson, & New, 1987). Thus, the measure aims to quantify both the perceived costs of the relationship and the emotional closeness of the relationship to assess both the positive and negative aspects of dog ownership (Dwyer et al., 2006). In addition, it incorporates a third subscale to examine the moderating effects of dog-human interactions occurring in the relationship. Thus, the MDORS is an example of how the development of questionnaires for specific use in HAI may be guided by a theoretical basis.

When a new questionnaire is developed for use, it is important to assess the validity metrics to prevent a misalignment of questionnaires with the construct intended to be measured. There are two general types of construct validity that can be measured in order to evaluate a new assessment tool: convergent validity and discriminant validity (Campbell & Fiske, 1959). Convergent validity represents the extent to which a questionnaire correlates or agrees with other assessment tools intending to measure the same construct of interest. In HAI questionnaire assessment, tests of convergent validity are common. For example, the recently developed Pet Attachment and Life Scale (PALS; Cromer & Barlow, 2013) was tested for convergent validity with the Lexington Attachment to Pets Scale (LAPS; Johnson, Garrity, & Stallones, 1992) and the Companion Animal Bonding Scale (Poresky, Hendrix, Mosier, & Samuelson, 1987) via factor analysis and intercorrelations, verifying that PALS was measuring a similar construct with these measures. On the other hand, discriminant validity (or the extent to which a questionnaire does *not* correlate with assessment tools that do not measure the construct of interest) is not as often assessed in the field of HAI, but is additionally valuable to provide a balanced perspective of a questionnaire's construct validity (Campbell & Fiske, 1959). With the

continuous and rapid development of questionnaire measures in HAI (Wilson & Netting, 2012), it is important that researchers verify these aspects of validity (e.g. The Human-Animal Interaction Scale; Fournier, Berry, Letson, & Chanen, 2016; Animal Attitude Scale; Herzog, Grayson, & McCord, 2015).

Using existing questionnaires for HAI

For constructs related to human health and wellbeing, existing questionnaires with well-established psychometric properties, including construct validity, exist in the fields of psychology and medicine. Standardized clinical questionnaires have been used to assess the effects of AAIs on a range of populations comparing outcomes to other types of interventions or treatments in meta-analyses (e.g., Virués-Ortega, Pastor-Barriuso, Castellote, Población, & de Pedro-Cuesta, 2012). For instance, the Aberrant Behavior Checklist-Community (ABC-C; Aman, Burrow, & Wolford, 1995) has been well-researched in the field of psychopharmacology for individuals with autism spectrum disorder (ASD) and has been demonstrated to translate well to research on animal-assisted intervention in this population (e.g., Gabriels et al., 2012; Gabriels et al., 2015). In this case, the significant reductions in ABC-C irritability symptoms resulting from a 10-week randomized trial of therapeutic horseback riding (Gabriels et al., 2015) was almost half the effects of the irritability reductions in an 8-week medication trial using this same measurement tool with a same-age ASD population (R. Owen et al., 2009). This enabled inferences to be made that therapeutic horseback riding has the potential to be a complementary or alternative intervention to medications for ASD (Arnold, 2015). Therefore, the use of existing

standardized questionnaires is crucial to move the HAI field forward to understanding its efficacy in relation to other mainstream or established interventions by generating results that are comparable across research disciplines.

Learning about animals through questionnaires

One use of questionnaires specific to HAI research is to collect information about companion and service animals from humans who interact with them often, such as their handlers, trainers, or owners. These questionnaires can be used to assess an animal's health or its behavioral/temperament traits. Information about the animal is obtained indirectly through the reporting of a third party in these questionnaires, creating a risk that the collected data may be biased by the subjective perceptions of the raters. For example, an owner may have an inaccurate representation of their animal's behavior if the owner is inexperienced at recognizing certain stress or fear behaviors in their animal. A comparable situation in human research is the use of parent- or teacher-reports to assess traits in children. Research shows that parent or teacher-reports about a child are not always well-correlated with the direct report from a child (Phares, Compas, & Howell, 1989), indicating that researchers should rigorously assess the risk of bias when using data provided by a third person.

To parse out subjective bias, investigators developing questionnaires that capture animal behaviors should prioritize psychometric evaluation of their instrument. For example, the Canine Behavioral Assessment and Research Questionnaire (C-BARQ; Hsu & Serpell, 2003) is designed to collect owner-reports of dog behavior. The C-BARQ demonstrates good convergence with the clinical assessments of veterinary

behaviorists (Hsu & Serpell, 2003), predictive validity of the success of assistance dogs in training (Duffy & Serpell, 2012), and inter-rater reliability between multiple members of the same household assessing the behavior of the same dog (Jakuba et al., 2013). In this regard, the C-BARQ is a good example of a questionnaire with demonstrated validity and reliability that can be used to collect information regarding canine behavior. When psychometric properties are properly evaluated, questionnaires can be important sources of information about animal behavior and health. When evaluating constructs that may be more subjective and thus more prone to bias, such as animal welfare, we recommend that investigators endeavor to combine questionnaires with the use of direct physiological and/or behavioral assessment.

New technology in questionnaires

Recent progress in new technologies has greatly improved questionnaire assessments by increasing their ecological validity (i.e. generalizability to real world settings), shortening the amount of time necessary for research participants to complete questionnaires, and reducing data entry errors (Gibbons, 2016). One example of an emerging technology in questionnaire assessment is electronic data collection, which can help avoid skipped responses, prevent missing data or out-of-range responses, and provide standardized scoring in real-time. A 2008 meta-analysis comparing self-reported treatment efficacy outcomes from 46 clinical trials found a high correlation (0.90) between paper and computer instruments, providing evidence for their equivalency (Gwaltney, Shields, & Shiffman, 2008). HAI researchers aiming to capture self-report before, during, or after an intervention may benefit particularly from including technology-based assessment.

One development in questionnaire technology that has been advantageous in numerous fields is crowdsourcing technology such as Amazon Mechanical Turk (MTurk). MTurk provides the ability for researchers to administer a questionnaire to over 500,000 individuals across 190 countries who are paid in small monetary amounts for completion of tasks, such as surveys. This method can provide researchers with large and diverse samples (Casler, Bickel, & Hackett, 2013) as well as fast and inexpensive data collection (Buhrmester, Kwang, & Gosling, 2011). A recent 2015 analysis estimated that a typical research group can reach a subject pool of 7,300 workers per quarter year (Stewart et al., 2015). Thus, MTurk can be advantageous to HAI researchers aiming to quantify perceptions, opinions, and beliefs (e.g. Mills, Robbins, & von Keyserlingk, 2016; Rabbitt, Kazdin, & Hong, 2014) or cross-cultural comparisons (e.g. Ruby, Heine, Kamble, Cheng, & Waddar, 2013). While using online crowdsourcing can sometimes pose threats to validity of the data (e.g. respondents who provide false demographics to fit inclusion criteria, or participants who rapidly advance through surveys without full attention), several guides exist to provide researchers with best practices to aid in minimizing these issues (e.g. Sheehan, 2017).

Ecological momentary assessment (EMA) is another recently developed form of questionnaire that attempts to minimize the risk of recall-bias present in traditional questionnaires collecting retrospective data (Shiffman, Stone, & Hufford, 2008). Instead of completing assessment tools in one session and answering questions about actions or feelings over the past few weeks, research participants are prompted to respond to questions of interest during their day-to-day lives. This could take multiple forms, from a diary updated at regular time intervals, to the use of smartphone applications and reminders to respond to research questions

throughout the day. The use of EMA is particularly relevant to HAI research studies that are interested in daily life experiences, such as interactions with companion or service animals. For example, one 2017 study used EMA via a mobile phone app to assess the physical activity and emotional affect of 71 dog owners (Liao, Solomon, & Dunton, 2017). Participants were prompted randomly eight times per day for 12 days to report on their current activities (“what were you doing before the beep went off?”), their current emotional states, and if they were with their pet dog. By asking the participant to remember actions from a recent and short period, this technique reduces recall-bias and enhances the accuracy of the collected data. Liao et al.’s (2017) study, among others, demonstrate the feasibility and potential of this promising approach to HAI research.

Physiological Measures

Physiological assessment can provide an objective measure of how situations differentially affect the internal state of human and animal research participants. A range of techniques exist to help investigators explore indicators of sympathetic and hypothalamic-pituitary-adrenal (HPA) axis activity through endocrine, cardiovascular, and skin conductance measurements. In addition, the evaluation of both daytime and nighttime activity via actigraphy offers a unique objective assessment in which to better understand how animals may impact human health. In this section, we review and offer suggestions for use of common physiological measures assessed in HAI research including assessment of cardiovascular, physical, electrodermal, and endocrine activity.

Cardiovascular activity

Measures of cardiovascular health and activity have been widely used in the

field of HAI to explore a range of both theoretical and applied research questions. Examples of these measures include blood pressure, heart rate, and heart rate variability; these signs of arousal and sympathetic-to-parasympathetic nervous system balance can be measured in both humans and most animals included in HAI research (Schreiner, 2016). By quantifying these physiological metrics, researchers in HAI have begun to examine the mechanistic bio-behavioral underpinnings of human-animal interaction (e.g., Barker, Knisely, McCain, Schubert, & Pandurangi, 2010; Polheber & Matchock, 2013). For example, laboratory studies have quantified the associations of positive HAI with cardiovascular activity in humans (e.g., Polheber & Matchock, 2013), animals (e.g., McGreevy, Righetti, & Thomson, 2005), and both humans and animals (e.g., in humans stroking horses; Hama, Yogo, & Matsuyama, 1996). Objective measures such as heart rate have also been used to quantify the temporal calming aspects of AAI in populations with high arousal levels (e.g., Krause-Parello & Friedmann, 2014). In studies outside of the laboratory, measures such as systolic and diastolic blood pressure can be analyzed cross-sectionally in order to examine the impact of pet ownership on overall cardiovascular health (e.g. Parslow & Jorm, 2003).

Although valuable, these physiological assessment methods are not without their limitations. Measurement often requires complementary self-report or behavioral assessment to confidently interpret directionality, which can require a more complex study design. For example, without instantaneous self-report or behavioral data, it is unknown if a higher heart rate indicates a state of excitement (positive relation) or a state of fear (negative relation). In addition, measurement of the autonomic nervous system, especially heart rate variability, can be complicated by large

inter-individual variation and a range of external factors such as time of day, use of nicotine and caffeine, and physical activity levels as well as measurement error in the device itself that need to be adequately controlled for (Quintana & Heathers, 2014). Reviews of the literature have found that many studies examining the relationship between cardiovascular health and pet ownership lack the methodological rigor necessary to draw valid conclusions, including failure to control for confounding variables and inadequate sampling techniques (Cutt, Giles-Corti, Knuiiman, & Burke, 2007). Future HAI research incorporating assessment of cardiovascular activity will benefit from aiming to address as many of these methodological and confounding limitations as possible to generate valid and conclusive data.

Reliable measures of cardiovascular activity have become easier and less expensive in recent decades. While equipment size has previously made long-term studies of cardiovascular health costly and cumbersome, new technologies are making the recording of cardiovascular activity more portable and affordable. During the past decade, wearable wristband and chest monitors have been developed to collect moment-to-moment heart rate, heart rate variability, and blood pressure. With the latest design improvements, these wearable monitors are unobtrusive, lightweight, and can be worn as comfortably as a watch. However, researchers should be cautious of devices that have not been independently validated; while a device may be cheaper, it may also be less accurate resulting in invalid study conclusions (O'Brien et al., 2000). This is especially the case for commercially available devices that may be cost-effective and familiar to participants, but not well-validated for use in research (e.g. Fitbit HR, Apple Watch, Wang, Blackburn, Desai, & et al., 2017).

Physical activity

Physical activity has long been an important area of assessment for understanding the relationship between humans and their companion animals. Specifically, the greater physical activity associated with dog walking has been a popular research topic in recent decades (for review see Christian et al., 2013). With emerging actigraphy technology to measure activity, researchers can now objectively quantify physical activity rather than relying solely on self-report. For example, Owen et al. (2010) used wearable activity monitors to measure the physical activity (mean daily activity, mean daily steps, and activity per minute) of over 2,000 children in three different countries as part of the Child Heart and Health Study in England project. By examining daytime physical activity of children in families with and without a pet dog, researchers could determine the specific areas of activity affected by dog ownership in a large and diverse sample (C. G. Owen et al., 2010). In addition to daily activity, recent research has also used actigraphy to examine the effect of animal-assisted therapy on nighttime activity and sleep quality (Swall, Fagerberg, Ebbeskog, & Hagelin, 2014).

Activity trackers have become lighter, cheaper, and more easily used for large-scale research than ever before (Shih, Han, Poole, Rosson, & Carroll, 2015). However, as is the case with wearable technology measuring cardiovascular activity, HAI researchers using activity monitors need to be cautious of devices without established validity (Tryon, 2004). While commercially available trackers may be well-validated for step counts (Kooiman et al., 2015), they may produce unreliable data for metrics such as energy expenditure and sleep efficiency (Evenson, Goto, & Furberg, 2015). Further, care must be taken to ensure that changes in activity are directly related to

the variable of interest (e.g. dog ownership) rather than confounding factors (e.g. health-related conditions motivating activity changes). For researchers interested in sleep actigraphy, it is vital that devices be carefully chosen based on their convergent validity with laboratory-measured sleep metrics (e.g. Jean-Louis, Kripke, Cole, Assmus, & Langer, 2001). Researchers must also be aware that participants must keep a sleep log (e.g., time got into bed, time fell asleep, times woken up) to enable researchers to accurately interpret the objective actigraphy data (Littner et al., 2003).

Concomitant with improvements in technology and application of ambulatory devices, researchers in both HAI as well as in veterinary and animal behavior fields have begun to non-invasively and reliably measure actigraphy in animals. For example, several small and unobtrusive devices have been validated to reliably measure physical activity in dogs via collar and harness attachments (e.g. Westgarth & Ladha, 2017; Yam et al., 2011). Measuring animal activity along with human activity may lead to a new understanding of bidirectional health associations in HAI research. In a recent study, Patel et al. (2017) placed wearable devices on 40 adults and 40 of their pet dogs to objectively examine the impact that sharing a bedroom or a bed has on both human and dog sleep; findings indicated that canine sleep outcomes were unaffected by position, but human sleep efficiency was significantly lower when the dog was in the bed (Patel et al., 2017). Future studies may also benefit from using actigraphy to study the activity, sleep, and welfare of animals used in AAI such as service dogs (Burrows, Adams, & Millman, 2008).

Electrodermal activity

Electrodermal activity (EDA), also referred to as skin conductance or galvanic

skin response, has long been a frequently used psychophysiological measure in the field of psychology (Boucsein, 2012). While some sweat glands mostly respond to thermal stimuli (i.e., increases in temperature), the activity of eccrine sweat glands is typically associated with emotional arousal (van Dooren, de Vries, & Janssen, 2012). Measuring the electrodermal activity in areas of the body close to eccrine sweat glands is thus an effective way to measure emotional arousal in humans. In HAI research, electrodermal activity has been used in the laboratory to measure the arousal of humans when viewing videos of animals in distressing situations as an indicator of empathy (Westbury & Neumann, 2008), and to evaluate the arousal of children with autism spectrum disorder during animal-assisted intervention (O'Haire, McKenzie, Beck, & Slaughter, 2015). Similar to cardiovascular activity, technological advances have replaced cumbersome monitors with lighter, wearable devices that can measure electrodermal activity with good precision and minimal burden for the participant (e.g. Garbarino, Lai, Bender, Picard, & Tognetti, 2014). However, while these devices enable researchers to collect real-time physiological activity in a non-invasive and unobtrusive manner, most devices are still relatively expensive which may limit their widespread use.

Although electrodermal activity is a common measure of arousal in humans, it remains challenging to measure in animals and its validity varies highly with the species studied. In particular, the function and process of sweating vary widely between mammal species commonly included in HAI (Allen & Bligh, 1969). One study found that when sheep are petted by a familiar human, they show a decrease in skin humidity correlating with behavioral signs of well-being, implying that sheep in a more positive emotional state may show decreased

electrodermal activity (Reefmann, Wechsler, & Gygas, 2009). This study suggests that the use of skin humidity or conductance has potential to serve as a non-invasive measure of emotional arousal in some species. However, in other species in which electrodermal activity is not a validated measure of arousal state, other physiological assessment tools such as heart rate variability may be more appropriate for use (e.g. dogs, Katayama et al., 2016; farm animals, Von Borell et al., 2007).

Endocrine activity

Endocrine activity can complement other physiological activity measures by quantifying the levels of diverse hormones in the human body. Cortisol, an indicator of arousal and stress, and oxytocin, a hormone involved in affection and bonding, are two particular hormones popularly discussed with respect to HAI research (Beetz, Uvnäs-Moberg, Julius, & Kotrschal, 2012; Uvnäs-Moberg et al., 2011).

The hormone cortisol is produced in the adrenal cortex as part of the stress response system regulated by the hypothalamus-pituitary adrenal (HPA) axis (Stratakis & Chrousos, 1995). As a measure of both chronic and acute stress, cortisol is widely used in HAI research as a psychophysiological biomarker of human-animal interaction (Beetz et al., 2012). Cortisol has been used to measure the stress-attenuating effect of companion animals on humans (e.g., Kertes et al., 2016) and in studies of the welfare of animals included in animal-assisted intervention (e.g., Koda, Watanabe, Miyaji, Ishida, & Miyaji, 2015; A. McCullough et al., 2017). While in the past, the inclusion of cortisol in research has required invasive procedures such as blood collection, breakthroughs in salivary bioscience have made measurement non-invasive and less costly (Dreschel & Granger,

2016). Samples can even be taken in the home environment, as cortisol in saliva is stable and can be transported and mailed at room temperature (Clements & Parker, 1998). However, obtaining, assaying, and analyzing salivary cortisol data requires careful consideration of a number of factors (Levine, Zagoory-Sharon, Feldman, Lewis, & Weller, 2007), especially in studies involving children (Harmon, Hibel, Rumyantseva, & Granger, 2007). Cortisol concentrations can be significantly influenced by a large number of state and trait factors that need to be addressed for their confounding potential in statistical analyses (Hellhammer, Wüst, & Kudielka, 2009). To address this, HAI researchers are encouraged to follow the methodological standards suggested by expert guidelines (e.g. Stalder et al., 2016).

Alongside cortisol, oxytocin is growing in its inclusion in HAI research to explore the human-animal bond due to its key role in affection and bonding in mammalian species. Oxytocin has historically been measured from blood samples, but protocols have recently been developed to measure oxytocin in saliva and urine in both humans (e.g., Francis, Kirkpatrick, de Wit, & Jacob, 2016; Grewen, Davenport, & Light, 2010) and dogs (MacLean et al., 2018). While the validity of the interpretation of salivary oxytocin can be controversial (e.g., Horvat-Gordon, Granger, Schwartz, Nelson, & Kivlighan, 2005; M. E. McCullough, Churchland, & Mendez, 2013), the development of salivary sampling methods has a strong potential to reduce the stress of intravenous oxytocin sampling procedures and to positively impact animal welfare. The use of oxytocin in particular may be a good indicator of bonding between humans and dogs (Beetz et al., 2012), as the level of this hormone has been shown to increase in both owners and dogs during mutual eye contact (Nagasawa et al., 2015), and after a few

minutes of talking to and stroking a dog (Handlin et al., 2011). However, consideration needs to be given to other bonds and mood-altering lifestyle factors (e.g. pregnancy) when determining the HAI association.

Behavioral observation

Because of its focus on inter-species interactions, HAI research is inherently behavioral. Behavioral observation is uniquely complex in the field of HAI as it often involves the measurement of both human and animal behavior to quantify and describe the interactions taking place. While behavioral observation has been used less widely in HAI research compared to questionnaires (O'Haire et al., 2018), it is an important assessment that provides critical information to complement questionnaire and physiological assessments (Baumeister, Vohs, & Funder, 2007).

Using existing behavioral paradigms from other fields

Like questionnaires, several standardized behavioral paradigms already used in other fields can be adapted and reinterpreted for HAI research. The use of standardized behavioral testing may contribute to the theoretical understanding of the relationships between humans and animals. Although most standardized behavior tests must be done in a laboratory or experimental setting, this offers a controlled environment to test hypotheses and paradigms while minimizing confounding factors that may occur in the home environment. For example, in the field of psychology, the Strange Situation Procedure (Ainsworth, 1979) is a standard measure of infantile attachment; in a series of separations and reunions, behaviors of a child and its caregiver are evaluated. The Strange

Situation has similarly been used in HAI research to examine a dog's behavior when separated from, and reunited with, its owner to examine the correlates of attachment in human-canine relationships (Konok, Dóka, & Miklósi, 2011; Palmer & Custance, 2008; Rehn & Keeling, 2016; Topál, Miklósi, Csányi, & Dóka, 1998).

In addition to the field of psychology, many behavioral tasks in the fields of animal cognition and behavior may be reinterpreted to evaluate the human-animal bond. Specifically, performance in a task may inform investigators about the efficiency of the communication between an individual and the animal rather than about the cognitive capacities of the animal. For example, Hall and colleagues (2016) examined the potential relationship between a dog's point-following behavior and a child's level of attachment to their dog. Results indicated that children who reported greater feelings of attachment to their pet dogs had dogs that were more likely to successfully follow their pointing gesture to locate hidden food (Hall, Liu, Kertes, & Wynne, 2016). Utilizing these types of behavioral paradigms may help HAI researchers further understand the complexities of the human-animal bond.

Human behavior in HAI

HAI investigators may be interested in collecting behavioral data from a range of situations in which animals and humans interact. Depending on the research question, human behavioral data can be used in a variety of ways in HAI research. As an outcome measure, it may provide detailed information about the specific effects that human-animal interaction may have (e.g. how animal-assisted activities impact the social behaviors of children with autism, O'Haire, McKenzie, Beck, & Slaughter, 2013). In addition, HAI researchers may choose to quantify behavior when self-report

is not as practical for the research question of interest (e.g. the effect that a therapy dog's presence has on speed and accuracy of children's gross motor skills, Gee, Harris, & Johnson, 2007). As a mediator or moderator of outcomes, human behavior may have a significant effect on the construct being measured (e.g. a handler's affiliative and disciplinary behaviors on cortisol output of working dogs, Horváth, Dóka, & Miklósi, 2008). Finally, human behavioral data can also be used as a measure of treatment fidelity in interventions, in which audio or video recording interactions can allow researchers to verify that intervention providers are adhering to specific protocols while monitoring deviations from treatment content (Bellg et al., 2004).

Animal behavior in HAI

In HAI research, an animal's behavior can be used both to quantify the dynamics of human-animal interaction as well as acquire information about the animal itself such as monitoring animal welfare. In the context of human-animal interactions, an animal's specific directed behaviors towards a human can often provide key information necessary to interpreting results from animal-assisted interventions. For example, the display of affiliative behaviors such as physical contact-seeking and eye contact, and the absence of aggressive behaviors such as growling and teeth-baring may be of critical importance to the success of an intervention. At the same time, animal welfare may be monitored with the coding of both negative (e.g., stress), and positive (e.g., play) behaviors during an interaction (Ng, Albright, Fine, & Peralta, 2015).

Coding behavior

Similar to other fields studying behavior, the utility of behavioral data is

dependent on how it is coded (Chorney, McMurtry, Chambers, & Bakeman, 2014). Standardized coding tools attempt to systematically classify and quantify behaviors with a coding manual that is tested for reliability across multiple raters (Cone, 1999). The choice of coding tool depends on the research questions and behaviors of interest. If the behaviors of interest only include outcomes related to human behavior and ignore interactions with animals, the investigators may choose to use a standardized behavior coding tool developed in psychology or medicine (e.g., with a standardized pain behavior scale; Vagnoli et al., 2015). If behaviors of interest are animal-specific, HAI researchers will benefit by using ethograms developed in the field of animal behavior (e.g. equid ethogram, McDonnell, 2003). Because behaviors of interest in HAI research often include interactions between human and animals, existing single-species behavior coding tools are often not relevant for these HAI-specific behaviors.

As a response to the lack of standardized coding tools specific to HAI research, many studies have developed their own independent coding tools (e.g., Grandgeorge et al., 2014; Jenkins & Reed, 2013; Kotrschal & Ortbauer, 2003). One example of a coding tool specifically developed for HAI studies is the Observation of Human-Animal Interaction for Research (OHAIRE) coding system, which was developed to cover a range of human behaviors of interest in HAI research (e.g., emotional display, social communication), as well as behaviors specific to interacting with animals (O'Haire et al., 2013). The use of this coding tool in a number of different studies has demonstrated its ease of use, and its comparison with standardized questionnaires has demonstrated strong convergent validity (Guérin, 2017). In addition, the OHAIRE has established inter-rater reliability, which tests

for agreement across two or more independent raters who are blinded to the study's conditions and hypotheses (Guérin, 2017). The use of standardized behavior assessment tools with established protocols, validity, and reliability is essential to create comparable results across studies in the field of HAI.

In addition to standardized behavior coding tools, a range of technologies are available for quantifying and interpreting the behavior of both humans and animals in HAI research. Behavioral coding software such as the Noldus Observer allows researchers to synchronize and easily code video data. HAI researchers have successfully used Noldus Observer for a range of topics, including examining pet dogs and their owner's behaviors during an operational task (Kotrschal, Schöberl, Bauer, Thibeaut, & Wedl, 2009) and quantifying the effect of positive human interaction on cow behavior during milking (Ivemeyer, Knierim, & Waiblinger, 2011). New developments in Observer software have also recently allowed for the integration of physiological data with behavioral data (Observer XT; Zimmerman, Bolhuis, Willemsen, Meyer, & Noldus, 2009). Eye tracking analysis, another type of behavioral coding applicable to HAI, has similarly experienced technological advances. Borrowing from infant eye-tracking technology, HAI researchers have used increasingly reliable and non-invasive eye-tracking methods with dogs to begin to further interpret the intricacies of the dog-human relationship (e.g., Jakovcevic, Mustaca, & Bentosela, 2012; Téglás, Gergely, Kupán, Miklósi, & Topál, 2012).

Conclusion

In this paper, we have summarized the current state of assessment in human-animal interaction research in three key areas: questionnaires, physiological

measures, and behavioral observation. For each assessment area, we have identified concerns with its use, presented examples of its successful application, and provided recommendations for its use in HAI research.

Questionnaires provide invaluable sources of information about the personal experiences of research participants and the animals in their lives. When evaluating constructs that are not specific to HAI research, investigators should first turn to the appropriate standardized assessment methods developed in other research fields. HAI research will continue to benefit from adopting well-validated questionnaires for constructs related to clinical outcomes and human health and wellbeing; doing so will increase generalizability and clinical significance of outcomes. When the unique characteristics of HAI research require investigators to adapt or develop new assessment tools, they should be subject to rigorous psychometric validation, including construct validity in the forms of convergent and discriminant validity. In addition, the development of new questionnaires in the field of HAI needs to be driven by a strong theoretical background and clear research questions. Finally, HAI researchers should take advantage of new and rapidly evolving technologies in questionnaire administration to further refine and develop their assessments. While questionnaires provide critical information from the subjective point of view of research participants in HAI, the use of physiological measures and behavioral observation are important, objective complements to questionnaires.

With physiological assessment, HAI researchers have begun to investigate how human-animal interactions influence biologic responses. Cardiovascular activity, physical activity, electrodermal activity, and endocrine activity have all been measured by HAI researchers aiming to quantify the physiological effects of human-animal

interaction. However, these objective assessments are currently not as widely used in HAI as questionnaires (O'Haire et al., 2018), in part because of the additional resources and expenses necessary. As the field of HAI grows, however, so will its need to incorporate methodologically rigorous designs that combine both subjective and objective outcome measures. Through non-invasive saliva collection, HAI researchers can analyze a variety of biomarkers through emerging salivary bioscience technology. In addition, for measures with strong biological similarities between humans and their animal partners (e.g., heart rate, cortisol, actigraphy), future HAI research may benefit from using dyadic analysis to examine the potential bidirectional effects of human-animal interaction while exploring the possibility of physiological synchronicity between humans and animals who share a bond. However, if using physiological methods, researchers must carefully consider the validity, accuracy, and precision of their measurement methods as well as potential confounding variables.

Direct observation of behavior provides valuable complementary data to self-reported measures and physiological assessment. Recording how humans and animals interact with each other is important to understand the observable effects of human-animal interaction. When quantifying animal or human behavior, researchers should consider standardized coding tools readily available for the construct of interest. By utilizing existing behavior paradigms from other fields, standardized coding tools, and employing new technology for coding and behavior, HAI investigators can obtain a better understanding of HAI by capturing the unique and subtle behavioral aspects of dynamic interaction. Similar to physiological assessment, future HAI research may benefit specifically from using dyadic behavioral analysis to examine how human and animal

behaviors reflect the strength of a bond or relationship. Finally, HAI researchers should explore emerging technologies in behavior coding to obtain more precise and valuable information.

Regardless of the primary assessment method chosen, HAI researchers should aim to incorporate multiple forms of assessment in their studies to gain the most comprehensive picture of the assessed constructs. By using multiple forms of assessment in a single study, results will increase in validity by minimizing the bias of one single method and developing a more comprehensive and multimodal picture of the construct of interest. Historically, multimodal assessment has not been a prominent method in HAI research. A recent systematic literature review of animal-assisted therapy for youth found that only 24% of 45 studies used more than one method of assessment to measure outcomes. Further, the tendency to use multiple methods was not significantly related with year of publication, indicating that researchers have not adopted this practice more frequently over time (May et al., 2016). While limited funding and resources to adopt multiple assessment methods may be an issue, this practice will further aid in establishing new standards of methodological rigor in the study of human-animal interaction (Esposito et al., 2011; Herzog, 2015; McCune et al., 2014).

Whether collecting subjective or objective data, the consistent use of standardized, well-validated assessment tools is critical in furthering the credibility and strength of the field of HAI. While assessment in HAI research can also be strengthened by such strategies as using large, diverse sample sizes and decreasing sources of error, we argue that the widespread use of multiple valid, reliable, and standardized assessment tools and methods is crucial for creating a rigorous empirical evidence base for the future of the field.

References

- Ainsworth, M. S. (1979). Infant–mother attachment. *American psychologist*, *34*(10), 932.
- Allen, T., & Bligh, J. (1969). A comparative study of the temporal patterns of cutaneous water vapour loss from some domesticated mammals with epitrachial sweat glands. *Comparative biochemistry and physiology*, *31*(2), 347-363.
- Aman, M. G., Burrow, W. H., & Wolford, P. L. (1995). The Aberrant Behavior Checklist-Community: Factor validity and effect of subject variables for adults in group homes. *American Journal on Mental Retardation*.
- Arnold, L. E. (2015). The alone rangers and silver. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(7), 535-536.
- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., . . . Nosek, B. A. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*(2), 108-119.
- Barker, S. B., Knisely, J. S., McCain, N. L., Schubert, C. M., & Pandurangi, A. K. (2010). Exploratory study of stress-buffering response patterns from interaction with a therapy dog. *Anthrozoös*, *23*(1), 79-91.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396-403.
- Beetz, A., Uvnäs-Moberg, K., Julius, H., & Kotrschal, K. (2012). Psychosocial and psychophysiological effects of human-animal interactions: the possible role of oxytocin. *Frontiers in Psychology*, *3*.
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, *23*(5), 443.
- Bernabei, V., De Ronchi, D., La Ferla, T., Moretti, F., Tonelli, L., Ferrari, B., . . . Atti, A. R. (2013). Animal-assisted interventions for elderly patients affected by dementia or psychiatric disorders: A review. *Journal of Psychiatric Research*, *47*(6), 762-773.
- Boucsein, W. (2012). *Electrodermal activity*: Springer Science & Business Media.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3-5.
- Burrows, K. E., Adams, C. L., & Millman, S. T. (2008). Factors affecting behavior and welfare of service dogs for children with autism spectrum disorder. *Journal of Applied Animal Welfare Science*, *11*(1), 42-62.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, *56*(2), 81.
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156-2160.
- Chorney, J. M., McMurtry, C. M., Chambers, C. T., & Bakeman, R. (2014). Developing and modifying behavioral coding schemes in pediatric psychology: a practical guide. *Journal of pediatric psychology*, *40*(1), 154-164.
- Christian, H. E., Westgarth, C., Bauman, A., Richards, E. A., Rhodes, R. E., Evenson, K. R., . . . Thorpe Jr, R. J. (2013). Dog ownership and physical activity: a review of the evidence. *Journal of Physical Activity and Health*, *10*(5), 750-759.
- Clements, A. D., & Parker, C. R. (1998). The relationship between salivary cortisol concentrations in frozen versus mailed samples. *Psychoneuroendocrinology*, *23*(6), 613-616.
- Cone, J. D. (1999). Observational assessment: Measure development and research issues. In P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology* (2nd ed., pp. 183-223). New York: Wiley.
- Cromer, L. D., & Barlow, M. R. (2013). Factors and Convergent Validity of the Pet Attachment and Life Impact Scale (PALS). *Human-Animal Interaction Bulletin*.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, *52*(4), 281.
- Cutt, H., Giles-Corti, B., Knuiman, M., & Burke, V. (2007). Dog ownership, health and physical activity: A critical review of the literature. *Health & Place*, *13*(1), 261-272.
- Davis, T. N., Scalzo, R., Butler, E., Stauffer, M., Farah, Y. N., Perez, S., . . . Coviello, L. (2015). Animal assisted interventions for

- children with autism spectrum disorder: A systematic review. *Education and Training in Autism and Developmental Disabilities*, 50(3), 316-329.
- Dreschel, N. A., & Granger, D. A. (2016). Advancing the Social Neuroscience of Human-Animal Interaction: The Role of Salivary Bioscience. In L. S. Freund, S. McCune, L. Esposito, N. R. Gee, & P. McCardle (Eds.), *The Social Neuroscience of Human-Animal Interaction* (pp. 195-216). Washington, DC: American Psychological Association.
- Duffy, D. L., & Serpell, J. A. (2012). Predictive validity of a method for evaluating temperament in young guide and service dogs. *Applied Animal Behaviour Science*, 138(1), 99-109.
- Esposito, L., McCune, S., Griffin, J. A., & Maholmes, V. (2011). Directions in human-animal interaction research: Child development, health, and therapeutic interventions. *Child Development Perspectives*, 5(3), 205-211.
- Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*, 12(1), 159.
- Fine, A. H., Tedeschi, P., & Elvolve, E. (2015). Forward thinking: The evolving field of human-animal interactions. In A. H. Fine (Ed.), *Handbook on Animal-Assisted Therapy: Foundations and Guidelines for Animal-Assisted Interventions* (4th ed., pp. 21-36). San Diego, CA: Elsevier Inc.
- Fournier, A. K., Berry, T. D., Letson, E., & Chanen, R. (2016). The Human-Animal Interaction Scale: Development and Evaluation. *Anthrozoös*, 29(3), 455-467.
- Francis, S. M., Kirkpatrick, M. G., de Wit, H., & Jacob, S. (2016). Urinary and plasma oxytocin changes in response to MDMA or intranasal oxytocin administration. *Psychoneuroendocrinology*, 74, 92-100.
- Gabriels, R. L., Agnew, J. A., Holt, K. D., Shoffner, A., Zhaoxing, P., Ruzzano, S., . . . Mesibov, G. (2012). Pilot study measuring the effects of therapeutic horseback riding on school-age children and adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 6(2), 578-588.
- Gabriels, R. L., Zhaoxing, P., DeChant, B., Agnew, J. A., Brim, N., & Mesibov, G. (2015). Randomized controlled trial of therapeutic horseback riding in children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(7), 541-549.
- Garbarino, M., Lai, M., Bender, D., Picard, R. W., & Tognetti, S. (2014). *Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition*. Paper presented at the Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on.
- Gee, N. R., Harris, S. L., & Johnson, K. L. (2007). The role of therapy dogs in speed and accuracy to complete motor skills tasks for preschool children. *Anthrozoös*, 20(4), 375-386.
- Gibbons, C. J. G. (2016). Turning the page on pen-and-paper questionnaires: combining ecological momentary assessment and computer adaptive tests to transform psychological assessment in the 21st century. *Frontiers in Psychology*, 7, 1933.
- Grandgeorge, M., Bourreau, Y., Alavi, Z., Lemonnier, E., Tordjman, S., Deleau, M., & Hausberger, M. (2014). Interest towards human, animal and object in children with autism spectrum disorders: an ethological approach at home. *European Child & Adolescent Psychiatry*, 1-11.
- Greven, K. M., Davenport, R. E., & Light, K. C. (2010). An investigation of plasma and salivary oxytocin responses in breast-and formula-feeding mothers of infants. *Psychophysiology*, 47(4), 625-632.
- Guérin, N. A. (2017). *Psychometrics Properties and Use of a Behavioral Observation Coding Tool for Human-Animal Interaction Research*. (M.S.), Purdue University, West Lafayette, IN. (10608361)
- Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health*, 11(2), 322-333.
- Hall, N. J., Liu, J., Kertes, D., & Wynne, C. D. L. (2016). Behavioral and Self-report Measures Influencing Children's Reported Attachment to Their Dog. *Anthrozoös*, 29(1), 137-150.
- Hama, H., Yogo, M., & Matsuyama, Y. (1996). Effects of stroking horses on both humans' and horses' heart rate responses. *Japanese Psychological Research*, 38(2), 66-73.
- Handlin, L., Hydbring-Sandberg, E., Nilsson, A., Ejdebäck, M., Jansson, A., & Uvnäs-Moberg, K. (2011). Short-term interaction between dogs and their owners: effects on

- oxytocin, cortisol, insulin and heart rate—an exploratory study. *Anthrozoös*, 24(3), 301-315.
- Harmon, A. G., Hibel, L. C., Rummyantseva, O., & Granger, D. A. (2007). Measuring salivary cortisol in studies of child development: watch out—what goes in may not come out of saliva collection devices. *Developmental Psychobiology*, 49(5), 495-500.
- Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, 34(2), 163-171.
- Herzog, H. (2015). The Research Challenge: Threats to the Validity of Animal-Assisted Therapy Studies and Suggestions for Improvement. In A. H. Fine (Ed.), *Handbook on animal-assisted therapy: Foundations and guidelines for animal-assisted interventions* (Fourth ed., pp. 402-407): Academic Press.
- Herzog, H., Grayson, S., & McCord, D. (2015). Brief measures of the animal attitude scale. *Anthrozoös*, 28(1), 145-152.
- Horvat-Gordon, M., Granger, D. A., Schwartz, E. B., Nelson, V. J., & Kivlighan, K. T. (2005). Oxytocin is not a valid biomarker when measured in saliva by immunoassay. *Physiology & behavior*, 84(3), 445-448.
- Horváth, Z., Dóka, A., & Miklósi, Á. (2008). Affiliative and disciplinary behavior of human handlers during play with their dog affects cortisol concentrations in opposite directions. *Hormones and behavior*, 54(1), 107-114.
- Hsu, Y., & Serpell, J. A. (2003). Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *Journal of the American Veterinary Medical Association*, 223(9), 1293-1300.
- Ivemeyer, S., Knierim, U., & Waiblinger, S. (2011). Effect of human-animal relationship and management on udder health in Swiss dairy herds. *Journal of dairy science*, 94(12), 5890-5902.
- Jakovcevic, A., Mustaca, A., & Bentosela, M. (2012). Do more sociable dogs gaze longer to the human face than less sociable ones? *Behavioural processes*, 90(2), 217-222.
- Jakuba, T., Polcová, Z., Fedáková, D., Kottferová, J., Mareková, J., Fejsáková, M., . . . Ondrašovič, M. (2013). Differences in evaluation of a dog's temperament by individual members of the same household. *Society & Animals*, 21(6), 582-589.
- Jean-Louis, G., Kripke, D. F., Cole, R. J., Assmus, J. D., & Langer, R. D. (2001). Sleep detection with an accelerometer actigraph: comparisons with polysomnography. *Physiology & behavior*, 72(1-2), 21-28.
- Jenkins, S. R., & Reed, F. D. D. (2013). An experimental analysis of the effects of therapeutic horseback riding on the behavior of children with autism. *Research in Autism Spectrum Disorders*, 7(6), 721-740.
- Johnson, T. P., Garrity, T. F., & Stallones, L. (1992). Psychometric evaluation of the Lexington Attachment to Pets Scale (LAPS). *Anthrozoös: A Multidisciplinary Journal of The Interactions of People & Animals*, 5(3), 160-175.
- Kamioka, H., Okada, S., Tsutani, K., Park, H., Okuizumi, H., Handa, S., . . . Mutoh, Y. (2014). Effectiveness of animal-assisted therapy: A systematic review of randomized controlled trials. *Complementary Therapies in Medicine*, 22(2), 371-390.
- Katayama, M., Kubo, T., Mogi, K., Ikeda, K., Nagasawa, M., & Kikusui, T. (2016). Heart rate variability predicts the emotional state in dogs. *Behavioural processes*, 128, 108-112.
- Kazdin, A. E. (2015). Methodological Standards and Strategies for Establishing the Evidence Base of Animal-Assisted Therapies. In A. H. Fine (Ed.), *Handbook on animal-assisted therapy: Foundations and guidelines for animal-assisted interventions* (Fourth ed., pp. 377-390): Academic Press.
- Kertes, D. A., Liu, J., Hall, N. J., Hadad, N. A., Wynne, C. D. L., & Bhatt, S. S. (2016). Effect of pet dogs on children's perceived stress and cortisol stress response. *Social Development*, n/a-n/a.
- Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, 65(23), 2276-2284.
- Koda, N., Watanabe, G., Miyaji, Y., Ishida, A., & Miyaji, C. (2015). Stress levels in dogs, and its recognition by their handlers, during animal-assisted therapy in a prison. *Animal Welfare*, 24(2), 203-209.
- Konok, V., Dóka, A., & Miklósi, Á. (2011). The behavior of the domestic dog (*Canis familiaris*) during separation from and reunion with the owner: A questionnaire and an experimental study. *Applied Animal Behaviour Science*, 135(4), 300-308.
- Kooiman, T. J. M., Dontje, M. L., Sprenger, S. R., Krijnen, W. P., van der Schans, C. P., & de Groot, M. (2015). Reliability and validity of

- ten consumer activity trackers. *BMC Sports Science, Medicine and Rehabilitation*, 7(1), 24.
- Kotrschal, K., & Ortbauer, B. (2003). Behavioral effects of the presence of a dog in a classroom. *Anthrozoös*, 16(2), 147-159.
- Kotrschal, K., Schöberl, I., Bauer, B., Thibeaut, A.-M., & Wedl, M. (2009). Dyadic relationships and operational performance of male and female owners and their male dogs. *Behavioural processes*, 81(3), 383-391.
- Krause-Parello, C. A., & Friedmann, E. (2014). The Effects of an Animal-Assisted Intervention on Salivary Alpha-Amylase, Salivary Immunoglobulin A, and Heart Rate during Forensic Interviews in Child Sexual Abuse Cases. *Anthrozoös: A Multidisciplinary Journal of The Interactions of People & Animals*, 27(4), 581-590.
- Levine, A., Zagoory-Sharon, O., Feldman, R., Lewis, J. G., & Weller, A. (2007). Measuring cortisol in human psychobiological studies. *Physiology & behavior*, 90(1), 43-53.
- Liao, Y., Solomon, O., & Dunton, G. F. (2017). Does the Company of a Dog Influence Affective Response to Exercise? Using Ecological Momentary Assessment to Study Dog-Accompanied Physical Activity. *American Journal of Health Promotion*, 31(5), 388-390.
- Littner, M., Kushida, C. A., Anderson, W. M., Bailey, D., Berry, R. B., Davila, D. G., . . . Loubé, D. (2003). Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002. *Sleep*, 26(3), 337-341.
- Maber-Aleksandrowicz, S., Avent, C., & Hassiotis, A. (2016). A Systematic Review of Animal-Assisted Therapy on Psychosocial Outcomes in People with Intellectual Disability. *Research in Developmental Disabilities*, 49-50, 322-338.
- MacLean, E. L., Gesquiere, L. R., Gee, N., Levy, K., Martin, W. L., & Carter, C. S. (2018). Validation of salivary oxytocin and vasopressin as biomarkers in domestic dogs. *Journal of neuroscience methods*, 293, 67-76.
- May, D. K., Seivert, N. P., Cano, A., Casey, R. J., & Johnson, A. (2016). Animal-assisted therapy for youth: A systematic methodological critique. *Human-Animal Interaction Bulletin*, 4(1), 1-18.
- McCullough, A., Jenkins, M. A., Ruehrdanz, A., Gilmer, M. J., Olson, J., Pawar, A., . . . Pichette, D. (2017). Physiological and behavioral effects of animal-assisted interventions on therapy dogs in pediatric oncology settings. *Applied Animal Behaviour Science*.
- McCullough, M. E., Churchland, P. S., & Mendez, A. J. (2013). Problems with measuring peripheral oxytocin: can the data on oxytocin and human behavior be trusted? *Neuroscience & Biobehavioral Reviews*, 37(8), 1485-1492.
- McCune, S., Kruger, K. A., Griffin, J. A., Esposito, L., Freund, L. S., Hurley, K. J., & Bures, R. (2014). Evolution of research into the mutual benefits of human-animal interaction. *Animal Frontiers*, 4(3), 49-58.
- McDonnell, S. M. (2003). *The equid ethogram: a practical field guide to horse behavior*: Eclipse Press.
- McGreevy, P. D., Righetti, J., & Thomson, P. C. (2005). The reinforcing value of physical contact and the effect on canine heart rate of grooming in different anatomical areas. *Anthrozoös*, 18(3), 236-244.
- Mills, K. E., Robbins, J., & von Keyserlingk, M. A. (2016). Tail docking and ear cropping dogs: Public awareness and perceptions. *PloS one*, 11(6), e0158131.
- Nagasawa, M., Mitsui, S., En, S., Ohtani, N., Ohta, M., Sakuma, Y., . . . Kikusui, T. (2015). Oxytocin-gaze positive loop and the coevolution of human-dog bonds. *Science*, 348(6232), 333-336.
- Ng, Z. Y., Albright, J., Fine, A. H., & Peralta, J. (2015). Our Ethical and Moral Responsibility: Ensuring the Welfare of Therapy Animals. In A. H. Fine (Ed.), *Handbook on animal-assisted therapy: Foundations and guidelines for animal-assisted interventions* (4th ed.). San Diego, CA: Elsevier Inc.
- O'Brien, E., Coats, A., Owens, P., Petrie, J., Padfield, P. L., Littler, W. A., . . . Mee, F. (2000). Use and interpretation of ambulatory blood pressure monitoring: recommendations of the British Hypertension Society. *BMJ : British Medical Journal*, 320(7242), 1128-1134.
- O'Haire, M. E. (2016). Research on animal-assisted intervention and autism spectrum disorder, 2012-2015. *Applied Developmental Science*.
- O'Haire, M. E., Bibbo, J., Mueller, M. K., Ng, Z. Y., Buechner-Maxwell, V. A., & Hoffman, C. F. (2018). Overview of Centers and Institutes for Human-Animal Interaction in the United States. *Human-Animal*

- Interaction Bulletin*, 6(Special Edition), 3-22.
- O'Haire, M. E., McKenzie, S. J., Beck, A. M., & Slaughter, V. (2013). Social behaviors increase in children with autism in the presence of animals compared to toys. *PloS one*, 8(2), e57010.
- O'Haire, M. E., McKenzie, S. J., Beck, A. M., & Slaughter, V. (2015). Animals may act as social buffers: Skin conductance arousal in children with autism spectrum disorder in a social context. *Developmental Psychobiology*, 57(5), 584-595.
- Owen, C. G., Nightingale, C. M., Rudnicka, A. R., Ekelund, U., McMinn, A. M., van Sluijs, E. M., . . . Whincup, P. H. (2010). Family dog ownership and levels of physical activity in childhood: findings from the Child Heart and Health Study in England. *American journal of public health*, 100(9), 1669-1671.
- Owen, R., Sikich, L., Marcus, R. N., Corey-Lisle, P., Manos, G., McQuade, R. D., . . . Findling, R. L. (2009). Aripiprazole in the treatment of irritability in children and adolescents with autistic disorder. *Pediatrics*, 124(6), 1533-1540.
- Palmer, R., & Custance, D. (2008). A counterbalanced version of Ainsworth's Strange Situation Procedure reveals secure-base effects in dog-human relationships. *Applied Animal Behaviour Science*, 109(2-4), 306-319.
- Parslow, R. A., & Jorm, A. F. (2003). Pet ownership and risk factors for cardiovascular disease: another look. *The Medical Journal of Australia*, 179(9), 466-468.
- Patel, S. I., Miller, B. W., Kosiorek, H. E., Parish, J. M., Lyng, P. J., & Krahn, L. E. (2017). *The Effect of Dogs on Human Sleep in the Home Sleep Environment*. Paper presented at the Mayo Clinic Proceedings.
- Phares, V., Compas, B. E., & Howell, D. C. (1989). Perspectives on child behavior problems: Comparisons of children's self-reports with parent and teacher reports. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1(1), 68.
- Polheber, J., & Matchock, R. (2013). The presence of a dog attenuates cortisol and heart rate in the Trier Social Stress Test compared to human friends. *Journal of Behavioral Medicine*, 1-8.
- Poresky, R. H., Hendrix, C., Mosier, J. E., & Samuelson, M. L. (1987). The companion animal bonding scale: Internal reliability and construct validity. *Psychological reports*, 60(3), 743-746.
- Quintana, D. S., & Heathers, J. A. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. *Frontiers in Psychology*, 5, 805.
- Rabbitt, S. M., Kazdin, A. E., & Hong, J. E. (2014). Acceptability of animal-assisted therapy: Attitudes toward AAT, psychotherapy, and medication for the treatment of child disruptive behavioral problems. *Anthrozoös*, 27(3), 335-350.
- Reefmann, N., Wechsler, B., & Gyax, L. (2009). Behavioural and physiological assessment of positive and negative emotion in sheep. *Animal Behaviour*, 78(3), 651-659.
- Rehn, T., & Keeling, L. J. (2016). Measuring dog-owner relationships: Crossing boundaries between animal behaviour and human psychology. *Applied Animal Behaviour Science*, 183, 1-9.
- Ruby, M. B., Heine, S. J., Kamble, S., Cheng, T. K., & Waddar, M. (2013). Compassion and contamination. Cultural differences in vegetarianism. *Appetite*, 71, 340-348.
- Schreiner, P. J. (2016). Emerging Cardiovascular Risk Research: Impact of Pets on Cardiovascular Risk Prevention. *Current cardiovascular risk reports*, 10(2), 1-8.
- Sheehan, K. B. (2017). Crowdsourcing research: Data collection with Amazon's Mechanical Turk. *Communication Monographs*, 1-17.
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annu. Rev. Clin. Psychol.*, 4, 1-32.
- Shih, P. C., Han, K., Poole, E. S., Rosson, M. B., & Carroll, J. M. (2015). Use and adoption challenges of wearable activity trackers. *ICConference 2015 Proceedings*.
- Stalder, T., Kirschbaum, C., Kudielka, B. M., Adam, E. K., Pruessner, J. C., Wüst, S., . . . Hellhammer, D. H. (2016). Assessment of the cortisol awakening response: expert consensus guidelines. *Psychoneuroendocrinology*, 63, 414-432.
- Stewart, N., Ungemach, C., Harris, A. J., Bartels, D. M., Newell, B. R., Paolacci, G., & Chandler, J. (2015). The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. *Judgment and Decision making*, 10(5), 479.
- Stratakis, C. A., & Chrousos, G. P. (1995). Neuroendocrinology and pathophysiology of the stress system. *Annals of the New York Academy of Sciences*, 771(1), 1-18.

- Swall, A., Fagerberg, I., Ebbeskog, B., & Hagelin, C. L. (2014). A therapy dog's impact on daytime activity and night-time sleep for older persons with Alzheimer's disease-A case study. *Clinical Nursing Studies*, 2(4), 80.
- Téglás, E., Gergely, A., Kupán, K., Miklósi, Á., & Topál, J. (2012). Dogs' gaze following is tuned to human communicative signals. *Current Biology*, 22(3), 209-212.
- Topál, J., Miklósi, Á., Csányi, V., & Dóka, A. (1998). Attachment behavior in dogs (*Canis familiaris*): a new application of Ainsworth's (1969) Strange Situation Test. *Journal of Comparative Psychology*, 112(3), 219.
- Tryon, W. W. (2004). Issues of validity in actigraphic sleep assessment. *Sleep*, 27(1), 158-165.
- Uvnäs-Moberg, K., Handlin, L., & Petersson, M. (2011). Promises and pitfalls of hormone research in human-animal interaction.
- Vagnoli, L., Caprilli, S., Vernucci, C., Zagni, S., Mugnai, F., & Messeri, A. (2015). Can Presence of a Dog Reduce Pain and Distress in Children during Venipuncture? *Pain Management Nursing*(0).
- van Dooren, M., de Vries, J. J. G. G.-J., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & behavior*, 106(2), 298-304.
- Virués-Ortega, J., Pastor-Barriuso, R., Castellote, J. M., Población, A., & de Pedro-Cuesta, J. (2012). Effect of animal-assisted therapy on the psychological and functional status of elderly populations and patients with psychiatric disorders: a meta-analysis. *Health Psychology Review*, 6(2), 197-221.
- Von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., . . . Prunier, A. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—a review. *Physiology & behavior*, 92(3), 293-316.
- Wang, R., Blackburn, G., Desai, M., & et al. (2017). Accuracy of wrist-worn heart rate monitors. *JAMA Cardiology*, 2(1), 104-106.
- Westbury, H. R., & Neumann, D. L. (2008). Empathy-related responses to moving film stimuli depicting human and non-human animal targets in negative circumstances. *Biological psychology*, 78(1), 66-74.
- Westgarth, C. & Ladha, C. (2017). Evaluation of an open sourcemethod for calculating physical activity in dogs from harness and collar based sensors. *BMC Veterinary Research*, 13(1), 322.
- Wilson, C. C., & Netting, F. E. (2012). The status of instrument development in the human-animal interaction field. *Anthrozoös*, 25(sup1), s11-s55.
- Yam, P., Penpraze, V., Young, D., Todd, M., Cloney, A., Houston-Callaghan, K., & Reilly, J. J. (2011). Validity, practical utility and reliability of Actigraph accelerometry for the measurement of habitual physical activity in dogs. *Journal of Small Animal Practice*, 52(2), 86-91.
- Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S., & Noldus, L. P. (2009). The Observer XT: A tool for the integration and synchronization of multimodal signals. *Behavior research methods*, 41(3), 731-735.