

# Evaluating FIDO: Developing and pilot testing the Field Instantaneous Dog Observation tool

Amy E. Bauer, Mary Jordan, Monica Colon, Traci Shreyer, and Candace C. Croney

## Abstract

Field assessments of the health and welfare of kenneled dogs (*Canis lupus familiaris*) must be both accurate and rapid. In order to facilitate such evaluations, especially by individuals with limited training in canine behavior and welfare, a non-invasive tool was developed and pilot-tested utilizing dogs housed in commercial breeding facilities. Behavioral responses to approach were organized into three categories: red, indicating a fearful response to approach, green, indicating an affiliative or neutral response to approach, and yellow, indicating an ambivalent response to approach. The inter-rater reliability (IRR) of the tool when used by both behavioral experts and novice raters was evaluated with and without the presence of the dog's familiar caretaker. Utilizing Cohen's kappa, the experts had almost perfect levels of agreement (kappa=0.87). The novice raters had substantial levels of agreement (kappa=0.74). Overall, the dogs assessed by the novice raters had high proportions of green responses to approach and there was no statistically significant effect of caretaker presence on the proportion of green responses to approach. The assessment tool evaluated herein appears to have a high degree of IRR whether used by experts in canine behavior or by novice raters and may be a useful screening tool to determine the need for more in-depth welfare assessments.

Pet Behaviour Science | 2017, Vol.4, 1 - 15  
DOI: [10.21071/pbs.v0i4.5766](https://doi.org/10.21071/pbs.v0i4.5766)

Amy E. Bauer, Mary Jordan, Monica Colon, Traci Shreyer, and Candace C. Croney

1. Department of Comparative  
Pathobiology Purdue University  
College of Veterinary Medicine  
USA

Paper Research

2. Purdue University College of  
Veterinary Medicine  
USA

3. University of Puerto Rico at  
Mayaguez. Puerto Rico.

\* Email:  
[ccroney@purdue.edu](mailto:ccroney@purdue.edu)

United States

Keywords:  
dogs; commercial  
breeding; behavior;  
welfare; assessment

## HIGHLIGHTS

- The Field Instantaneous Dog Observation (FIDO) tool, a non-invasive, field ready tool designed to capture the immediate welfare status of dogs housed in commercial breeding facilities was developed and pilot tested for inter-rater reliability.
- Affiliative or neutral (green) responses to approach were most commonly identified by both novice and expert raters in this study.
- Both expert and novice raters achieved substantial levels of agreement when using the behavioral component of the FIDO tool.
- There was no statistically significant effect of caretaker presence on response to approach in this study.

## 1. INTRODUCTION

Accurate assessments of dog behavior and welfare help to determine both how well individual dogs cope with their environments and the adequacy of those environments. These evaluations are especially important in facilities where dogs are housed for extended periods of time such as commercial breeding kennels. Dogs in high volume breeding facilities are widely believed to live in environments lacking in enrichment and social contact with their caretakers (McMillan et al. 2011). Stress associated with these environmental inadequacies can affect physical, behavioral, and emotional health for both the breeding dogs and their puppies (McMillan et al. 2011). Achieving the goal of providing an ethical source for companion dogs (Bir et al. 2016) depends in part on a reliable welfare assessment tool that can be utilized in breeding and other kennel settings from which dogs may be acquired.

The welfare of an animal is defined as its ability to cope with its environment (Broom 1991) and is both an individual metric and one that changes over the animal's lifetime. Welfare has physical and behavioral manifestations (Broom 1991) and both aspects must be included in an assessment tool. However, because of concerns about observer subjectivity and situational variation in animal responses, behavioral metrics of welfare are often minimal or absent as components of welfare assessment tools. An example of situational variation is play behavior, which is a good indicator of positive well-being (Boissy et al. 2007; Vinke et al. 2005). However, failure to observe play during an assessment does not necessarily indicate poor welfare, particularly if a different behavior, such as eating, was being performed at the time of observation. Nevertheless, when validated and tested for reliability, many behaviors can be excellent proxy metrics of animal mental health and well-being.

Along these lines, assessments of fear have been included as key behavioral indices in evaluations of welfare in production species (Muri et al. 2013; Powell et al. 2016). Likewise, fear in dogs has been used as an indicator of their early rearing experiences that has implications for their overall quality of life. For example, dogs with little human contact during the

socialization window have been reported to have a lasting, generalized fear of people, resulting in related deficits that also significantly impact their puppies (Serpell and Jagoe 1995; Lindsay 2000). Incorporation of behavior into welfare assessments is therefore essential to fully understand how well animals are coping in a given environment.

Animal-based assessments often focus on ethograms, lists of specifically defined behaviors associated with both negative and positive emotional states (Kiddie and Collins 2014; Protopopova et al. 2014). Challenges with ethogram-type assessments include observer drift (interpretation of behavior changing over time with increased familiarity) (Kiddie and Collins 2014; Protopopova 2016), fixation on specific behaviors that results in improper interpretation of the animal's overall state (Vas et al. 2008), and limited time for assessment of each specific behavior in the field setting. However, these challenges can be addressed via validation of the behaviors of interest as well as robust observer training. In addition, a benefit of ethogram-type assessments is their detail, which allows the prevalence of and the ability of raters to identify each behavior to be determined. Behaviors that are rare or difficult to interpret consistently can be identified and either removed from the assessment or focused on in training.

A rapid, non-invasive, field ready welfare assessment tool that is reliable is greatly needed for commercial breeding facilities as it could be utilized by researchers, facility inspectors, and staff interested in promoting high welfare standards for dogs. Among the concerns reported about dogs in such kennels is heightened fearfulness when confronted with unfamiliar people that has been noted in dogs originating from commercial breeding facilities (McMillan et al. 2011; McMillan 2017). However, to our knowledge, no studies have been conducted in the commercial breeding environments in which the dogs are kept. A tool that permits objective evaluations of dogs in these types of kennels would facilitate improved understanding of the quality of their early rearing environments relative to the development of behaviors observed later.

The first objective of this study, therefore, was to develop a rapid, non-invasive, field ready assessment tool useable by individuals without extensive training in canine behavior. The second objective was to pilot test the field readiness of the newly created Field Instantaneous Dog Observation (FIDO) tool, and determine inter-rater reliability (IRR). It was important to do this with experts as well as with raters with limited training in canine behavior (novice raters) to ensure broad usefulness of the tool. While dogs kept in commercial breeding facilities have not been well studied, the responses of dogs to unfamiliar people have been found to differ based on the presence of a familiar person (Pullen et al. 2012; Kis et al. 2014). Thus, a final objective was to evaluate the dogs' responses to approach both with and without their caretakers present.

## 2. METHODS

### Facilities

All observations were conducted in United States Department of Agriculture (USDA) licensed commercial breeding facilities. Kennel size varied by facility to meet or exceed the USDA requirements for dogs (USDA, 2017). All dogs were provided unrestricted access to an outdoor run during the day.

### Statement of approval

The procedures described in this manuscript were reviewed and approved by the Purdue University Institutional Animal Care and Use Committee.

### Initial tool development

The assessment tool developed in this study includes measurements of physical health and behavioral well-being. The physical health variables are described in Table 1 (Hurt, 2016). The behavioral component of the tool utilizes response to approach by a stranger as a surrogate measure of socialization toward people. An ethogram consisting of 43 different variables, including the dog's position in the kennel when initially approached by a stranger (Table 2) was developed for test use based on established indicators of affiliative, maintenance and defensive behaviors in dogs (Beerda

et al. 1997; Lindsay 2001; Boissy et al. 2007; Sontag and Overall 2014).

**Table 1. Measures of physical health and cleanliness**

Measure	Scale
Body Condition Score (BCS)	1=Emaciated 2=Thin 3=Moderate 4=Stout 5=Obese
Body Cleanliness (Percent of the body covered in debris)	0=0% 1=1-25% 2=26-50% 3=51-75% 4= >76%
Nasal Discharge	Present or absent
Ocular discharge	Present or absent
Tear staining	Present or absent
Sneezing	Present or absent
Coughing	Present or absent
Missing fur or poor coat	Present or absent
Wounds, sores or lesions	Present or absent
Lameness	Present or absent

Ninety seven dogs maintained at five commercial breeding kennels that were licensed by the USDA in Indiana, USA were scored on the physical and behavioral metrics compiled. During this initial scoring, two female observers approached the front of the dog's home kennel. The testers stood side by side 0.61m away from the kennel door, and turned their bodies so that their sides faced the door and dog to avoid direct eye contact that could have been aversive or threatening to the dog. The experimenter closest to the door bent at the waist and extended a hand toward the dog. Using the ethogram, the dog's immediate response to approach was recorded. The dog's physical metrics were then recorded. Following this scoring, the ethogram was revised to facilitate ease and accuracy of data collection obtained by live-observations in the field. Categories were adjusted such that manifestations of the freeze, flight, and fight responses (Lindsay 2001), were captured in a single category, Fearful. Behaviors that were never or only rarely observed during pilot testing were removed and behaviors that were observed and classified as "other" were added (Table 3).

<b>Category</b>	
Behavior	Description of behavior
<b>Friendly/affiliative</b>	
Approach	The dog moves toward the observer
Play bow	The dog lowers the front half of its body while keeping the hindquarters elevated to indicate playfulness
Solicit attention	The dog attempts to gain the observer's attention (e.g. scrambling at the front of the pen)
Other	
<b>Maintenance</b>	
Eat	The dog is consuming food
Rest	The dog is laying down in a relaxed position
Groom	The dog is cleaning itself
Other	
<b>Fidget/high arousal</b>	
Displacement behaviors	The dog is demonstrating a normal behavior that is not appropriate to its current situation
Increased RR or harsh panting	The dog is breathing at a more rapid rate than expected
Dilate pupils	The dog's pupils are wider than would be expected given ambient light
Excessive vocalizations	The dog vocalizes more than would normally be expected for the situation
Other	
<b>Fearful/avoidant</b>	
Move to back of enclosure/ outside	The dog moves to the back of or leaves the kennel when approached
Climbing the walls	The dog is attempting to scale the walls of its enclosure
Attempt to hide	The dog moves to be unseen by the observer
Attempt to escape	The dog is attempting to get away from the observer
Ambivalent (approach & avoid)	The dog alternately moves toward and away from the observer
Displacement behaviors	The dog is demonstrating a normal behavior that is out of context and reflects conflict
Trembling/ shaking	The dog's body is shaking
Crouched, slinking body posture	The dog's body is held low and close to the ground
Other	
<b>Fearful/shutdown</b>	
Freeze	The dog holds its body completely still in the presence of the observer
Catatonic	The dog is completely unresponsive in the presence of the observer
Trembling/ shaking	The dog's body is shaking
Other	
<b>Fear aggression - offensive</b>	
Lunge	The dog moves toward the observer quickly and aggressively
Vertical lip retraction	The lips are pulled up to expose the front and canine teeth
Tail-up	The tail is held tense, above hip level or vertical
Front of enclosure	The dog is already at or moves to the front of the kennel when approached
Forward body posture	The dog's weight and gaze are directed toward the observer
Hard body	The dog's entire body is tense and rigid
Other	
<b>Fear aggression- defensive</b>	
Horizontal lip retraction	The lips are pulled back to expose the molars
Withdrawal	The dog moves away from the observer
Ambivalent (approach & avoid)	The dog alternately moves toward and away from the observer
Backward body posture	The dog's weight is directed away from the observer
Tail down	The tail is held tense, below hip level
Other	
<b>Stereotypies</b>	
Circling	The dog repeatedly moves in a circular pattern in the same direction
Pacing	The dog walks back and forth in the kennel in a repeated pattern
Wall bouncing	The dog jumps and pushes off of the walls of the kennel repeatedly
Other	

**Table 2. Behavioral metrics used in assessment development**

Category	
Behavior	Description of behavior
Maintenance	
Eat	Dog is consuming food
Rest	Dog is laying down in a relaxed position
Drink	Dog is consuming water
Affiliative	
Approach	Dog moves toward the observer
Neutral wag	The tail is held in a relaxed position at hip level and wagged
Solicit	The dog attempts to gain the observer's attention (e.g. scrambling at the front of the pen)
Play Bow	The dog lowers the front half of its body while keeping the hindquarters elevated to indicate playfulness
Lean near	The dog leans on the door nearest to the observer
Fearful	
Non-responsive	The dog sits frozen and does not respond to approach
Tail tuck	The tail is tucked under the belly
Low wag	The tail is held tense, below hip level, and wagged
High wag	The tail is held tense, above hip level or vertical, and wagged
Tremble	The dog's body is shaking
Crouch	The dog's body is held low and close to the ground
Hide	The dog moves to be unseen by the observer
Growl	The dog growls in response to approach
Bark	The dog barks in response to approach
Hard body	The dog's entire body is tense and rigid
Hackles raised	The fur on the back of the dogs neck is erect (piloerection)
Lip lift	The lip is retracted either horizontally or vertically, exposing teeth
Lunge	The dog moves toward the observer quickly and aggressively
Repetitive	
Circling	The dog repeatedly moves in a circular pattern in the same direction
Pacing	The dog walks back and forth in the kennel in a repeated pattern
Wall bouncing	The dog jumps and pushes off of the walls of the kennel repeatedly
Continuous vocalization	The dog vocalizes without stopping
Calming behaviors	
Paw lift	One paw is lifted from the ground
Avert gaze	The dog looks away from the observer
Yawn	The dog yawns
Lip lick	The dog licks its lips
Shake off	The dog shakes its body after interaction with the observer
Positional front	
Start in front of kennel	The dog is already at the front of the kennel when approached
Move toward front of kennel	The dog moves to the front of the kennel when approached
Positional back	
Start in back of kennel	The dog is at the back of the kennel when approached
Move toward back of kennel	The dog moves to the back of the kennel when approached
Move from front to back repeatedly	The dog moves from the front of the kennel to the back of the kennel and then returns to the front of the kennel repeatedly (or from the back to the front)

Other	
Frantic	The dog is highly excited
Play with waterer	The dog manipulates the waterer when approached
Play with feeder	The dog manipulates the feeder when approached

**Table 3: Categories of behavior and the behaviors included in each category pilot tested by expert raters**

## Study one: Expert pilot testing and refinement

### Methods

Two female experts in canine behavior and welfare science (one with a PhD in Applied Animal Ethology and the other an applied animal behaviorist with an MA in animal behavior) used the revised ethogram-style form of the assessment to evaluate the response to approach of 38 dogs from two USDA licensed commercial breeding facilities located in Missouri, USA (Facilities 1 and 2, Table 4).

	Facilities					Total
	1	2	3	4	5	
Evaluator type (E Expert; N Novice)	E	E	N	N	N	
Number observed both with and without caretaker	N/A <sup>a</sup>	N/A	18	16	17	51
Number observed without caretaker only	N/A	N/A	0	1	1	2
Number observed with caretaker only	N/A	N/A	1	2	2	5
Facility total	20	18	19	19	20	
Total by evaluator type						
• Expert						38
• Novice						58
Total Overall						96

a. N/A - not applicable

**Table 4: Number of dogs observed at each facility. In all cases, the facility type was 'Breeder'**

The approach test was performed in the manner described above. The raters indicated which behaviors in Table 3 were performed in response to approach. These behavioral responses were then condensed into the behavioral categories. Finally, the results were categorized into one of three rating codes: red, yellow, or green (RYG). Responses classified as red included fearful behaviors (both offensive and defensive



aggression indicators and shut down responses, such as freezing) and stereotypic behaviors. Responses classified as green included affiliative behaviors such as approaching or soliciting attention from the rater, or neutral responses such as continued engagement in maintenance behaviors (e.g., eating, resting) despite the presence of the rater. Dogs categorized as displaying a yellow response did not clearly respond to approach in a manner that could be classified as either green or red.

IRR, the ability of different raters utilizing the tool to reach the same conclusions (Meagher, 2009) was evaluated. Percent agreement was calculated for the physical metrics, individual behaviors, categories of behavior, and responses to approach classified as red, yellow or green based upon the recorded behaviors and notes made by the raters. Percent agreement does not account for agreement due to chance, but can be an important baseline index of reliability (Burn and Weir 2011). In order to account for the presence of chance agreement, Cohen's kappa (Landis and Koch 1977) was used as a measure of IRR for each of the documented behaviors and the behavioral categories and was calculated using SPSS Statistics for Windows, Version 22.0 (IBM Corp. Released 2013. Armonk, NY: IBM Corp.). For the physical metrics recorded on an ordinal scale [Body Condition Score (BCS) and Body Cleanliness] a weighted kappa statistic was calculated in SAS, Version 9.4 (SAS Institute Inc., Cary, NC) using proc freq with 0.0001 added to 0 values to facilitate the analysis. A weighted kappa statistic was also calculated for the RYG portion of the tool. Kappa statistics <0.00 indicate poor agreement, 0.00-0.20 indicate slight agreement, 0.21-0.40 indicate fair agreement, 0.41-0.60 indicate moderate agreement, 0.61-0.80 indicate substantial agreement, and 0.81-1.00 indicate almost perfect agreement (Landis and Koch 1977).

Cohen's kappa is influenced by homogeneity in responses and as the prevalence of a given response approaches either 0 or 100%, interpretation of the kappa estimate becomes more challenging. To aid in interpretation of the Cohen's kappa values, prevalence indices (PIs) were calculated for the physical and behavioral components of the FIDO tool as a measure of homogeneity of response using the formula:

$$PI = |g-o|/N$$

where g indicates the number of agreed upon identifications of the behavior, o indicates the number of agreed upon instances where the behavior was not identified and n is the total number of observations (Burn and Weir 2011). The closer the PI is to 0, the more balanced the sample is in regard to the distribution of responses and the less influence prevalence has on the kappa estimate. The closer the PI is to 1, the more likely it is that sample homogeneity influenced the kappa estimate (Burn and Weir 2011).

The scoring tool was further refined (Table 5) so that only the behaviors regularly observed during the initial scoring sessions were included in the RYG categories on which novices were later trained. The dog's position in the pen at the time of approach was removed at this stage of refinement since that information could generally be determined by the response to approach. Testers were able to note the occurrence of any additional behaviors observed.

<b>Category</b>	
Behavior	Description of behavior
<b>Red</b>	
Fight	The dog demonstrates offensively aggressive behavior (e.g. lunging at the observer)
Flight	The dog moves to the back of or leaves the kennel when approached
Hard and forward body language	The dog's entire body is tense and rigid and oriented toward the observer
Frozen or catatonic	The dog is completely still or unresponsive in the presence of the observer
Stereotypic behaviors	The dog performs a pattern of behavior repeatedly
<b>Yellow</b>	
Ambivalent body language	The dog displays behavioral cues that may be interpreted in contradictory ways (e.g. growling or barking while tail wagging)
Ambivalent approach	The dog alternately moves toward and away from the observer
<b>Green</b>	
Friendly approach	Dog moves toward the observer
Solicits attention	The dog attempts to gain the observer's attention (e.g. scrambling at the front of the pen)
Neutral	The dog is engaged in another activity (e.g. eating, resting, or play) and is undisturbed by the presence of the observer
<b>Frantic/Overstimulated</b>	
The dog is highly excited	

**Table 5: Behaviors included in the RYG evaluation**

## Results

There was almost perfect agreement between the expert evaluators on the physical health metrics (Table 6).

Measurement	% Agreement	Kappa (95% CI <sup>a</sup> )	PI <sup>b</sup>
BCS	100	1.00	0.95 c
Body Cleanliness	94.74	0.48 (0.12, 1.00)	0.89 d
Nasal Discharge	100	*	1.0
Ocular discharge	100	1.00	0.95
Tear staining	97.37	0.89 (0.69, 1.00)	0.71
Sneezing	100	*	1.0
Coughing	100	*	1.0
Missing fur or poor coat	100	*	1.0
Wounds, sores or lesions	100	*	1.0
Lameness	100	*	1.0

**Table 6: Percent agreement, kappa estimate of inter-rater reliability, and prevalence indices for the physical health attributed measured by expert raters. \* indicates that agreement entirely or almost entirely in 1 category precluded calculation of a kappa statistic. a. CI - confidence interval b. PI - prevalence index / c. Dogs were only rated as 3 or 4. d Dogs were only rated as 0 or 1**

Percent agreement was 99.67% and kappa was 0.92 (95% CI: 0.77, 1.00). Although BCS and Body Cleanliness were evaluated on an ordinal scale, only two levels of each variable were observed in this sample, allowing the use of Cohen’s kappa and calculation of a PI for these variables.

When the 1444 pairs of individual behaviors recorded in the initial ethogram-style assessment were evaluated, the overall percent agreement was 96.95% (N=1400), kappa was 0.86 (95% CI: 0.82, 0.90), and the PI was 0.77. Table 7 provides the percent agreement, kappa, and PI for each of the behaviors in the ethogram-style assessment. A kappa value was unable to be calculated for 42.10% (16/38) of the individual behaviors because the raters either did not observe the behavior or it was very rarely identified. For an additional 23.68% of the behaviors (9/38), the raters were in perfect agreement.

When the data were collapsed into behavioral categories, 304 pairs of evaluations were available for analysis. Table 8 provides the percent agreement, kappa, and PI for each of the categories in this

assessment. The overall percent agreement was 94.41% (N=287), kappa was 0.88 (95% CI: 0.82, 0.93), and the PI was 0.26. A kappa statistic was able to be calculated for all of the categories. The raters were in perfect agreement in two categories, Maintenance and Repetitive behaviors. The categories Calm and Fearful had low PIs. Thus, the substantial agreement indicated by their kappa values (0.74 and 0.73 respectively) could be considered reliable.

When the data was further collapsed into RYG categories, an overall rating was made for each dog, giving a total of 38 pairs of evaluations. A green response to approach was identified by both raters for the majority of dogs (Table 9). The overall percent agreement was 92.10% (N=35) with a kappa value of 0.87 (95% CI: 0.72, 1.000), indicating almost perfect agreement between the raters (Landis & Koch, 1977). A PI was not calculated because the weighted kappa was used to take into account all 3 outcome categories.

Category	% Agreement	Kappa (95% CI <sup>a</sup> )	PI <sup>b</sup>
<b>Behavior</b>			
<b>Maintenance</b>			
Eat	100	1.00	0.89
Rest	97.37	*	0.97
Drink	100	1.00	0.89
<b>Affiliative</b>			
Approach	86.84	0.54(0.18-0.89)	0.66
Neutral wag	84.21	0.66(0.43-0.90)	0.26
Solicit	86.84	0.74(0.53-0.94)	0.13
Play Bow	100	1.00	0.68
Lean near	100	*	1
<b>Fearful</b>			
Non-responsive	100	*	1
Tail tuck	100	1.00	0.95
Low wag	97.37	0.65(0.03-1.00)	0.92
High wag	100	1.00	0.95
Tremble	100	1.00	0.95
Crouch	100	*	1
Hide	100	*	1
Growl	94.74	0.48(-0.12-1.00)	0.89
Bark	89.47	0.77(0.56-0.98)	0.32
Hard body	97.37	0.84(0.54-1.00)	0.81
Hackles raised	100	*	1
Lip lift	100	*	1
Lunge	100	*	1

Category	% Agreement	Kappa (95% CI <sup>a</sup> )	PI <sup>b</sup>
<b>Repetitive</b>			
Circling	100	1.00	0.84
Pacing	100	*	1
Wall bouncing	100	*	1
Continuous vocalization	100	1.00	0.84
<b>Calming signal</b>			
Paw lift	100	1.00	0.95
Avert gaze	97.37	*	0.97
Yawn	97.37	0.65(0.01-1.00)	0.92
Lip lick	86.84	0.65(0.38-0.93)	0.50
Shake off	100	*	1
<b>Position</b>			
Start in front of kennel	89.47	0.54(0.15-0.93)	0.74
Move toward front of kennel	100	*	1
Start in back of kennel	97.37	0.84(0.54-1.00)	0.81
Move toward back of kennel	94.74	0.72(0.35-1.00)	0.79
Move from front to back repeatedly	94.74	*	0.95
<b>Other</b>			
Frantic	97.37	0.79(0.38-1.00)	0.87
Play with waterer	100	*	1
Play with feeder	100	*	1

**Table 7: Percent agreement, kappa estimate of inter-rater reliability, and prevalence indices for the behaviors measured by expert raters in the ethogram-style assessment. \* indicates that agreement entirely or almost entirely in 1 category precluded calculation of a kappa statistic for this behavior. a. CI - confidence interval b. PI - prevalence index**

### Study two: Novice pilot testing

#### Methods

To assess the ability of the RYG scoring tool to be reliably used by novices in canine behavior, two female undergraduate student raters with limited formal

training in canine behavior and welfare evaluated the response to approach of 58 dogs from three USDA licensed commercial breeding facilities located in Indiana, USA (Facilities 3, 4, and 5).

	% agreement	Kappa (95% CI <sup>a</sup> )	PI <sup>b</sup>
Maintenance	100	1.00	0.79
Affiliative	94.74	0.64 (0.17-1.00)	0.84
Fearful	86.84	0.73 (0.52-0.94)	0.18
Repetitive	100	1.00	0.74
Calm	89.47	0.74 (0.51-0.98)	0.42
Positional front	89.47	0.54 (0.15-0.93)	0.74
Positional back	97.37	0.91 (0.73-1.00)	0.66
Other	86.84	0.79 (0.38-1.00)	0.87
Overall	94.41	0.88 (0.82-0.93)	0.26

**Table 8: Categorical agreement when assessed by expert raters. a. CI - confidence interval b. PI - prevalence index**

	Rater 1	Rater 2
N. Red (%)	5 (13.16)	4 (10.53)
N. Yellow (%)	4 (10.53)	7 (18.42)
N. Green (%)	29 (76.31)	27 (71.05)
Total	38	38

**Table 9: Summary of dogs with response to approach rated as red, yellow or green by expert raters**

Prior to evaluating the dogs, these raters were trained on how to perform the approach and identify behaviors falling into the RYG categories. As with the expert raters, the rating pair approached the dog with body angled perpendicular to the kennel door. The experimenter closest to the door bent at the waist and extended a hand toward the dog. Each dog's immediate response to approach was recorded as red, yellow, or green. On the first approach, only the raters were present. The raters approached the front of the kennel, noted the selected dog's response to approach and then evaluated the metrics of physical well-being. After each selected dog at the facility was evaluated, the raters waited for 30 minutes, then repeated the test with a caretaker present in order to gauge whether dogs' responses differed with a familiar person. Table 4 details the number of dogs observed at each facility and, for the novice raters, those observed without a caretaker present, with a caretaker present, and both with and without a caretaker present



Percent agreement was calculated for the physical metrics and responses to approach classified as red, yellow or green. As described in the expert pilot testing section, Cohen's kappa was calculated for physical metrics recorded as present or absent as a measure of IRR. A weighted kappa was utilized for the physical metrics recorded on an ordinal scale and the RYG portion of the tool. Additionally, the percentage of dogs with response to approach in each category was calculated for approaches conducted with and without the caretaker present at the individual facility level and for the overall sample. Where the raters agreed upon the dogs' responses to approach a Wilcoxon signed rank test was performed to determine if the presence of a caretaker made a statistically significant difference in the responses to approach using SPSS Statistics for Windows, Version 22.0 (IBM Corp. Released 2013. Armonk, NY: IBM Corp.).

Measurement	% Agreement	Kappa (95% CI <sup>a</sup> )	PI <sup>b</sup>
BCS	82.98	-0.04 (-0.11, 0.03)	0.83
Body Cleanliness	67.39	0.37 (0.15, 0.60)	N/A
Nasal Discharge	100	*	1.0
Ocular discharge	90.12	*	0.90
Tear staining	68.08	0.20 (-0.00, 0.40)	0.55
Sneezing	100	*	1.0
Coughing	100	*	1.0
Missing fur or poor coat	100	*	1.0
Wounds, sores or lesions	100	*	1.0
Lameness	100	*	1.0

**Table 10: Percent agreement, kappa estimate of inter-rater reliability, and prevalence indices for the physical health attributed measured by novice raters. \* indicates that agreement entirely or almost entirely in 1 category precluded calculation of a kappa statistic. A prevalence index was not calculated for Body Cleanliness as this was assessed on an ordinal scale with a weighted kappa. a. CI - confidence interval b.PI-prevalence index**

### Results

When the binary indicators of physical well-being were evaluated in aggregate, percent agreement between the novice raters was 96.29% and kappa was 0.26 (95% CI: 0.02, 0.50), fair agreement. The kappa value was strongly influenced by a homogeneous sample as

indicated by the overall PI of 0.95. Table 10 provides the percent agreement, kappa, and PI for each of the indicators of physical well-being.

The novice raters were in agreement as to the type of response to approach of 81.13% of the 53 dogs observed when a caretaker was not present (N=43). When a caretaker was present, the novice raters were in agreement as to the type of response to approach of 85.71% of the 56 dogs observed (N=48). Table 11 displays each rater's classification of responses to approach with and without a caretaker present.

	No caretaker		Caretaker	
	Rater 3	Rater 4	Rater 3	Rater 4
N. Red (%)	15 (28.30)	9 (16.98)	8 (14.28)	6 (10.71)
N. Yellow (%)	4 (7.55)	8 (15.09)	12 (21.43)	15 (26.79)
N. Green (%)	34 (64.15)	36 (67.92)	36 (64.29)	35 (62.50)
Total	53	53	56	56

**Table 11: Summary of dogs with response to approach rated as red, yellow or green with and without caretaker presence by novice raters**

### Interrater reliability

Table 12 displays percent agreement and kappa values for the behavioral portion of the FIDO tool with and without a caretaker present. Because IRR reflects an assessment of the tool, rather than the subjects of observation, it is important to evaluate the raters' responses in aggregate and thus the facility level values are not reported. When all of the observations were considered regardless of caretaker presence, the raters agreed 83.49% of the time (91/109). The weighted kappa value was 0.74 (95% CI: 0.63, 0.85), indicating substantial agreement (Landis and Koch 1977). A PI was not calculated because the weighted kappa was used to take into account all 3 outcome categories and at this time a PI or equivalent has not been developed for the weighted kappa (Burn and Weir 2011).

Without caretaker present		With caretaker present		Overall	
% agreement	K (95% CI) <sup>a</sup>	% agreement	K (95% CI)	% agreement	K (95% CI)
81.13	0.72 (0.56 - 0.87)	85.71	0.76 (0.61 - 0.92)	83.49	0.74 (0.63 - 0.85)

**Table 12: Percent agreement and weighted kappa for the RYG assessment by caretaker presence when performed by novice evaluators. a. CI - confidence interval**

### Effect of caretaker presence

Fifty-one dogs were observed both with and without a caretaker present. The raters agreed on the response to approach of 37 dogs (72.55%). Table 13 displays the percentage of responses to approach classified in each of the RYG categories when the raters agreed upon the dog's response to approach both with and without a caretaker. There was a high proportion of green responses to approach both without a caretaker present (81.08%, N=30) and with a caretaker present (78.38%, N=29). The percentage of yellow responses increased when a caretaker was present (16.21%, N=6 versus 2.70%, N=1). The percentage of red responses decreased when a caretaker was present (5.40%, N=2 versus 16.22%, N=6). Upon analysis of the agreed-upon paired samples from all breeding facilities using a Wilcoxon signed rank test the presence of a caretaker did not produce a statistically significant change in response to approach (T=-0.69; N=37; P=0.490).

## DISCUSSION

The first goal of this study was to develop a field-ready tool for use by laypersons in behavior to quickly and accurately gauge the immediate health and welfare status of kennelled dogs. We also aimed to pilot-test the tool for IRR, and evaluate the effect of caretaker presence in the context of commercial breeding facilities. Importantly, the FIDO tool developed was intended to provide a useful gauge of key aspects of dog welfare status at the time of observation in a field evaluation, as is often needed for routine site inspections or investigation of animal welfare complaints. Under such circumstances, it is important that inspectors or investigators have reliable, standardized, animal-based metrics for use to facilitate

consistent, objective evaluations. The FIDO tool is not intended to be used as a comprehensive assessment of the long-term welfare status of dogs. However, FIDO may provide an important screening tool that permits users to determine if and where more detailed dog welfare assessments are needed.

	Red % (N)	Yellow % (N)	Green % (N)	Total N	Wilcoxon T (P)
<b>Facility 3</b>					
Without caretaker present	33.33 (5)	0.00 (0)	66.67 (10)	15	-0.63 (0.529)
With caretaker present	13.33 (2)	33.33 (5)	53.33 (8)	15	
<b>Facility 4</b>					
Without caretaker present	0.00 (0)	10.00 (1)	90.00 (9)	10	-0.49 (0.627)
With caretaker present	0.00 (0)	0.00 (0)	100.00 (10)	10	
<b>Facility 5</b>					
Without caretaker present	8.33 (1)	0.00 (0)	91.67 (11)	12	-1.51 (0.132)
With caretaker present	0.00 (0)	8.33 (1)	91.67 (11)	12	
<b>All facilities</b>					
Without caretaker present	16.22 (6)	2.70 (1)	81.08 (30)	37	-0.69 (0.490)
With caretaker present	5.40 (2)	16.22 (6)	78.38 (29)	37	

**Table 13: Percentage of types of agreed upon response to approach expressed by dogs at commercial breeding facilities based on caretaker presence.**

While behavioral and welfare assessment tools have been developed with a focus on dogs living with families and housed in animal shelters (Kiddie and Collins, 2014; Belshaw et al. 2015; Barnard et al. 2016) and tools for the evaluation of personality traits in working dogs have also been developed (Sinn et al. 2010) this tool is unique. First, it was developed for those who need to quickly and accurately evaluate the physical and behavioral status of breeding dogs in the field without the use of video-recordings or other methodologies and resources that normally permit detailed analyses of long term behavioral patterns (Beerda et al. 1999; Hepper and Wells 2000). Second, the tool was developed to be particularly useful and applicable for evaluating dogs housed in commercial

breeding facilities. Further, it is the only tool to our knowledge reported to have been tested on-site with this particular population of dogs. The experiences of these dogs are likely to be very different from those housed in shelters as they may have limited exposure to environments and people outside of the breeding kennel. This has implications for those kennels that rehome their dogs, as dogs rehomed from commercial breeding operations have been reported to be more likely to exhibit behaviors associated with fear than companion dogs raised in other settings (McMillan et al. 2011). Thus, the dogs' responses to approach as a measure of fear (and a surrogate measure of socialization) were evaluated. Percent agreement, IRR, and PIs were calculated as part of the development of the tool. As part of pilot testing the RYG behavioral assessment portion of the tool, IRR was determined and the effect of caretaker presence was evaluated. Third, the tool not only allows an evaluator to indirectly assess the dog's socialization to strangers, it also may provide insight into the quality of the relationship with the dog's caretakers when scoring is done with and without their presence.

The expert evaluators had near perfect agreement on the physical metrics of the dogs evaluated. The novices likewise had high levels of agreement here, suggesting that the tool has high IRR on these components. Additionally, the experts had very high levels of percent agreement and kappa values that indicated almost perfect agreement when the data were coded as individual behaviors. However, the high or low prevalence of certain individual behaviors prohibited the calculation of Cohen's kappa for many of the behaviors. When investigating a construct, such as fear of strangers, a rating system such as the RYG system used in this study may be a better choice than coding individual behaviors (Kubinyi et al. 2015). In a comparison of predictive validity between behavioral and trait (or construct) ratings for success in a working dog training program, only slight differences in the predictive validity between methods were detected (Wilsson and Sinn 2012). Similarly, in this study, despite moving from behavioral coding to construct ratings, almost perfect inter-rater reliability was maintained as indicated by kappa values greater than 0.8 (Landis and Koch 1977).

When the RYG rating system was used by the novice raters, percent agreement was also high while the kappa statistic indicated substantial agreement (between 0.61 and 0.80) (Landis and Koch 1977). Although this tool is intended for use by raters without expertise in canine behavior, our findings suggest that additional training for raters in both behavior and the use of the tool may further improve IRR. As the majority of dogs evaluated by the novice raters demonstrated a green response to approach, it is possible that homogeneity of responses influenced the kappa statistic. Further evaluation of the tool in kennels of different types and management styles may elicit more variation in responses, which may help to confirm the strong reliability of the tool.

The prominence of green responses to approach may be challenging analytically, but it also provides insight into the behavior of commercial breeding dogs at their home facilities. The high proportion of green responses to approach contrasts with previous reports of high levels of fear in rehomed breeding dogs (McMillan et al. 2011). The discrepancies may be due to the fact that the approach tests were performed on-site at the dogs' home pens and they were therefore unlikely to have experienced the stressors that cumulatively might result in the global fear observed in other studies focusing on owner reports or observations made after leaving the commercial breeding environment. The owners of the breeding facilities chose to participate in the study, and thus the facilities included here may not be representative of all commercial breeding facilities. Additionally, this test has not yet been validated against long term metrics of dog well-being and therefore might only reflect the immediate welfare status. Further evaluation is needed to determine if this is indeed the case.

The findings of this study may also potentially provide insight into facility management and the quality (and perhaps frequency) of interactions dogs are having with people they regularly encounter. Overall, the presence of a caretaker in this study resulted in a decrease in the proportion of responses to approach classified as red. Although no statistically significant difference was detected across facilities when a caretaker was present versus absent, different patterns in behavioral responses to approach by strangers with

and without caretaker presence emerged at the individual facilities. If indeed response to approach is a useful surrogate measure for socialization, such changes may indicate the effectiveness of different management and socialization programs on dogs' responses to both caretakers and strangers. For example, in facilities where the proportion of red responses to approach decreased when the primary caretaker was present, it is likely that social buffering of stress induced by the presence of strangers occurred. Such patterns have been reported in other species (Boissy et al. 1998; Hennessy et al. 2000; Ruis et al. 2001; Hennessy et al. 2009; Croney 2014). This would suggest that the dogs were well socialized to their caretakers, and likely had positive interactions with them. However, keeping in mind breed variation relative to selection for or against affinity for strangers, having a high percentage of dogs showing red responses to strangers approaching in a non-threatening manner may indicate a need for improved socialization to people. Likewise, in facilities where the proportion of green responses to approach decreased when the primary caretaker was present, poor or insufficient interactions with caretakers may be indicated. Thus, there is potential for this tool to help identify facilities and individual dogs within facilities that would benefit from improved socialization efforts. There is also potential to monitor and potentially improve the quality of dog-caretaker interactions.

Overall, the assessment tool evaluated in this study appears to be a useful means of quickly gaining some insight into the immediate welfare status of kennelled dogs. There has been debate about the validity of behavioral assessment tools in predicting aggressive behavior in dogs (Kis et al. 2014; Patronek and Bradley 2016). Although the tool studied in this paper focused on capturing fearful behaviors and fear may manifest in aggression (Lindsay 2001; Sonntag and Overall 2014), additional testing of its validity is needed before anything might be concluded about its predictive power. Instead, our initial testing focused on reliability given the need for consistency of behavior identification in those tasked with on-site evaluations of dogs. As further evidence of the need for such consistency, in a study of 60 observers, the groups of dog owners, veterinarians, dog trainers, and non-dog owners did not differ in ability to identify behavior, but

there were differences between individuals within and between all categories (Tami and Gallagher 2009). Comparing agreement between expert and novice raters is a necessary next step in the evaluation of reliability and a crucial step in determining that the novice raters are interpreting responses to approach in the same manner as experts.

In addition to being reliable, robust welfare assessment tools must be valid (Meager 2009). This pilot study's major limitation is that the instantaneous metrics of well-being used have not yet been validated against those reflective of long-term welfare. However, this tool is not intended to replace the in-depth, detailed or longer term assessments of dog welfare which are needed to fully understand dogs' states relative to their attempts to cope with their environments (Broom 1991). Nor is it intended to stand alone as a predictor of a dog's ability to be an acceptable companion. To accomplish these goals, the validity of the metrics used would need to be investigated in various ways. For example, comparing the results of this tool to validated tools as has been done with other species (Powell et al. 2016), or to physiologic measures such as cortisol levels would be helpful in confirming that the response to approach by a stranger is correlated with overall welfare status. Further, scoring dogs outside of their home pens and in response to different stimuli would lend additional insight into both the behavioral well-being of dogs and the validity of the tool.

These limitations notwithstanding, the assessment does appear to offer good reliability and is potentially useful in identifying individual dogs potentially in need of interventions who appear at least initially to be fearful upon approach. Given this, while this study focused on the commercial dog breeding industry, the tool has the potential to be of benefit to people working with dogs housed in other types of kennels in other types of industries, including animal shelters and research facilities. Future studies will focus on evaluating the effects of additional training for novices, applying the FIDO tool to more diverse groups of dogs in different types of facilities, and testing the validity of the tool through comparisons between novice and expert raters and through comparisons with other methods of evaluating welfare in dogs housed in kennels.



## CONCLUSION

This study aimed to develop and assess the reliability of the FIDO tool, a non-invasive, instantaneous field ready evaluation of the welfare status of dogs housed in commercial breeding facilities. This tool is intended to be useful to kennel inspectors, caretakers and others who require the ability to quickly gauge dogs' welfare states. Overall, the tool provides a useful means to assess individual dogs, identify those having difficulty coping with their environment, and inform recommendations to improve dog welfare in kennel settings. It may therefore be helpful to commercial breeding facilities aiming to promote the physical, emotional, and behavioral well-being of their breeding dogs and puppies. Further, although refinement of training methods and further testing in more diverse environments are needed, this tool has the potential to be useful in other types of facilities where dogs are kenneled.

## Acknowledgements

Funding for this project was provided by the World Pet Association and the Pet Food Institute. Mary Jordan's participation was supported by the Merial Veterinary Scholars Summer Research Program. Monica Colon's participation was supported by the Purdue University Summer Research Opportunities Program. The authors would like to thank Dr. Judith Stella for her critical review of the manuscript.

## 7. REFERENCES

- Barnard, S., Pedernera, C., Candeloro, L., Ferri, N., Velarde A., and Dalla Villa, P. 2016. Development of a new welfare assessment protocol for practical application in long-term dog shelters. *Veterinary Record* 178: 18, accessed February 14, 2017, doi: 10.1136/vr.103336.
- Beerda, B., Schilder, M. B. H., van Hooff, J. A. R. A. M., and de Vries, H. W., 1997. Manifestations of chronic and acute stress in dogs. *Applied Animal Behaviour Science* 52: 307-319.
- Beerda, B., Schilder, M. B. H., van Hooff, J. A. R. A. M., de Vries, H. W., and Mol, J. A. 1999. Chronic stress in dogs subjected to social and spatial restriction I. Behavioral responses. *Physiology & Behavior* 92: 375-397.
- Belshaw, Z., Asher, L., Harvey, N. D., and Dean, R. S. 2015. Quality of life assessment in domestic dogs: An evidence-based rapid review. *The Veterinary Journal* 206: 203-212, accessed August 18, 2016, doi: 10.1016/j.tvjl.2015.07.016.
- Bir, C., Croney, C., and Widmar, N. O. 2016. Public perceptions of dog welfare, sourcing and breeding regulation. Purdue University, accessed March 6, 2017, <https://www.vet.purdue.edu/CAWS/files/documents/201602-public-perceptions-ofdog-welfare-sourcing-and-breeding-regulation.pdf>.
- Boissy, A., Terlouw, C., Le Neindre, P. 1998. Presence of cues from stressed conspecifics increases reactivity to aversive events in cattle: evidence for the existence of alarm substances in urine. *Physiology and Behavior* 63: 489-495.
- Boissy, A., Manteuffel, G., Jensen, M. B., Moe, R. O., Spruijt, B., Keeling, L. J., Winckler, C., Forkman, B., Dimitrov, I., Langbien, J., Bakken, M., Veissier, I. and Aubert, A. 2007. Assessment of positive emotions in animals to improve their welfare. *Physiology and Behavior* 92: 375-397.
- Broom, D. M. 1991. Animal welfare: concepts and measurement. *Journal of Animal Science* 69: 4167-4175, accessed March 6, 2017, doi: 10.2527/1991.691041676x.
- Burn, C. C. and Weir, A. A. S. 2011. Using prevalence indices to aid interpretation and comparison of agreement ratings between two or more observers. *The Veterinary Journal* 188:166-170, accessed March 7, 2017, doi: 10.1016/j.tvjl.2010.04.021.
- Croney, C. C. 2014. Let's stay together: Implications of social housing for laboratory pig welfare and management. *The Enrichment Record* 19: 14-19, accessed September 7, 2017, <http://enrichmentrecord.com/wp-content/uploads/2014/04/LETS-STAY-TOGETHER.pdf>.
- Hennessy, M. B., Maken, D. S., and Graves, F. C. 2000. Consequences of the presence of the mother or



- unfamiliar adult female on cortisol, ACTH, testosterone and behavioral responses of periadolescent guinea pigs during exposure to novelty. *Psychoneuroendocrinology* 25: 619-632. doi: 10.1016/S0306-4530(00)00014-7.
- Hennessy, M. B., Kaiser, S., & Sachser, N. 2009. Social buffering of the stress response: diversity, mechanisms, and functions. *Frontiers in neuroendocrinology* 30: 470-482. doi: 10.1016/j.yfrne.2009.06.001
- Hepper, P. G. and Wells, D. L. 2000. The influence of environmental change on the behavior of sheltered dogs. *Applied Animal Behaviour Science* 68: 151-162. doi: 10.1016/S0168-1591(00)00100-3.
- Hurt, M. 2016. *Evaluating the physical welfare of dogs in commercial breeding facilities in the United States*. M.S. thesis, Purdue University, USA.
- Kiddie, J. L. and Collins, L. M. 2014. Development and validation of a quality of life assessment tool for use in kennelled dogs (*Canis familiaris*). *Applied Animal Behavior Science* 158: 57-68, accessed August 12, 2016, doi: 10.1016/j.applanim.2014.05.008.
- Kis, A., Klausz, B., Persa, E., Miklosi, A., and Gacsi, M. 2014. Timing and presence of an attachment person affect sensitivity of aggression tests in shelter dogs. *Veterinary Record* 174, accessed November 1, 2016 doi: 10.1136/vr.101955.
- Kubinyi, E., Gosling, S. D., and Miklosi, A. 2015. A comparison of rating and coding behavioural traits in dogs. *Acta Biologica Hungarica* 66: 27-40, accessed March 3, 2017, doi: 10.1556/ABiol.66.2015.1.3.
- Landis, J. R. and Koch, G. G. 1977. The measure of observer agreement for categorical data. *Biometrics* 33: 159-174.
- Lindsay, S. R. 2000. *Handbook of Applied Dog Behavior and Training Volume 1, Adaptation and Learning*. Ames, Iowa: Iowa State Press.
- Lindsay, S. R., 2001. *Handbook of Applied Dog Behavior and Training. Volume 2, Etiology and assessment of behavior problems*. Ames, Iowa: Iowa State Press.
- McMillan, F. D., Duffy, D. L., and Serpell, J. A. 2011. Mental health of dogs formerly used as 'breeding stock' in commercial breeding establishments. *Applied Animal Behaviour Science* 135: 86-94, accessed February 14, 2017, doi: 10.1016/j.applanim.2011.09.006.
- McMillan, F. D., 2017. Behavioral and psychological outcomes for dogs sold as puppies through pet stores and/or born in commercial breeding establishments: Current knowledge and putative causes. *Journal of Veterinary Behavior* 19: 14-26, accessed April 11, 2017, doi: 10.1016/j.jveb.2017.01.001.
- Meagher, R. K., 2009. Observer ratings: Validity and value as a tool for animal welfare research. *Applied Animal Behaviour Science* 119: 1-14, accessed August 17, 2016. doi: 10.1016/j.applanim.2009.02.026.
- Muri, K., Stubsjoen, S. M., and Valle, P. S. 2013. Development and testing of an on-farm welfare assessment protocol for dairy goats. *Animal Welfare* 22: 385-400, accessed March 16, 2017 doi: 10.7120/0962786.22.3.385.
- Patronek, G. J. and Bradley, J. 2016. No better than flipping a coin: Reconsidering canine behavior evaluations in animal shelters. *Journal of Veterinary Behavior: Clinical Applications and Research* 15: 66-77, accessed March 10, 2017 doi: 10.1016/j.jveb.2016.08.001.
- Powell, C., Hemsworth, L. M., Rice, M., and Hemsworth, P. H. 2016. Comparison of methods to assess fear of humans in commercial breeding gilts and sows. *Applied Animal Behaviour Science* 181: 70-75, accessed March 16, 2017 doi: 10.1016/j.applanim.2016.05.027.
- Protopopova, A., Mehrkam, L. R., Boggess, M. M., and Wynne, C. D. L. 2014. In-kennel behavior predicts length of stay in shelter dogs. *PLOS One* 9: e114319, accessed November 1, 2016 doi: 10.1371/journal.pone.0114319.
- Protopopova, A., 2016. Effects of sheltering on physiology, immune function, behavior, and the welfare of dogs. *Physiology & Behavior* 159: 95-103, accessed November 1, 2016

- Pullen, A. J., Merrill, R. J. N., and Bradshaw, J. W. S. 2012. The effect of familiarity on behaviour of kennel housed dogs during interactions with humans. *Applied Animal Behaviour Science* 137: 66-73, accessed March 3, 2017, doi: 10.1016/j.applanim.2011.12.009.
- Ruis, M.A.W., te Brake, J.H.A., Engel B., Buist, W. G., Blokhuis, H.J., Koolhaa, J. M., 2001. Adaptation to social isolation: Acute and long-term stress responses of growing gilts with different coping characteristics. *Physiology & Behavior* 73: 541– 551.
- Serpell, J. and Jagoe, J. A. 1995. Early experience and the development of behaviour. In *The Domestic Dog: Its Evolution, Behaviour and Interactions with People*, 79-102, ed. J. Serpell. Cambridge, UK: Cambridge University Press.
- Sinn, D. L., Gosling, S. D., and Hilliard, S. 2010. Personality and performance in military working dogs: Reliability and predictive validity of behavioral tests. *Applied Animal Behaviour Science*, 127, 51-65, accessed August 12, 2016, doi: 10.1016/j.applanim.2010.08.007.
- Sonntag, Q. and Overall, K. L. 2014. Key determinants of dog and cat welfare: behaviour, breeding and household lifestyle. *Revue Scientifique et Technique (International Office of Epizootics)* 33: 213-220.
- Tami, G. and Gallagher, A. 2009. Description of the behaviour of domestic dog (*Canis familiaris*) by experienced and inexperienced people. *Applied Animal Behaviour Science*, 120: 159-169, accessed March 6, 2017, doi: 10.1016/j.applanim.2009.06.009.
- USDA Animal Care 2017. Dog Breeder Resource Guide. <https://www.aphis.usda.gov/aphis/ourfocus/animalwelfare/caw>. Accessed 15 September 2017
- Vas, J., Topal, J., Gyori, B., and Miklosi, A. 2008. Consistency of dogs' reactions to threatening cues of an unfamiliar person. *Applied Animal Behaviour Science* 112: 331-334, accessed March 6, 2017 doi: 10.1016/j.applanim.2007.09.002.
- Vinke, C. M., van Leeuwen, J., and Spruijt, B. M. 2005. Juvenile farmed mink (*Mustela vison*) with additional access to swimming water play more frequently than animals housed with a cylinder and platform, but without swimming water. *Animal Welfare* 16: 435-447.
- Wilsson, E. and Sinn, D. L. 2012. Are there difference between behavioral measurement methods? A comparison of the predictive validity of two ratings methods in a working dog program. *Applied Animal Behavior Science* 141: 158-172, accessed March 16, 2017, doi: 10.1016/j.jveb.2016.10.002.



This paper has been published by  
Pet Behaviour Science  
under a Creative Commons license  
4.0 Non-commercial - Share Alike - Attribution

As an open access journal, it is free of charges for  
both authors and readers

[www.petbehaviourscience.org](http://www.petbehaviourscience.org)