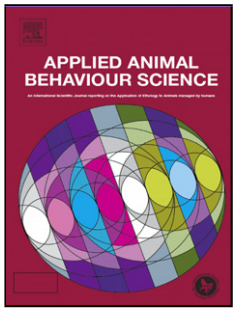


Title: Multi-Operator Qualitative Behavioural Assessment for dogs entering the shelter

Authors: Laura Menchetti, Cecilia Righi, Gabriella Guelfi, Claudia Enas, Livia Moscati, Stefania Mancini, Silvana Diverio

PII: S0168-1591(18)30555-0
DOI: https://doi.org/10.1016/j.applanim.2019.02.008
Reference: APPLAN 4770

To appear in: *APPLAN*

Received date: 3 October 2018
Revised date: 20 January 2019
Accepted date: 6 February 2019

Please cite this article as: Menchetti L, Righi C, Guelfi G, Enas C, Moscati L, Mancini S, Diverio S, Multi-Operator Qualitative Behavioural Assessment for dogs entering the shelter, *Applied Animal Behaviour Science* (2019), https://doi.org/10.1016/j.applanim.2019.02.008

Multi-Operator Qualitative Behavioural Assessment for dogs entering the shelter

Menchetti Laura[a], Righi Cecilia[b], Guelfi Gabriella[a], Enas Claudia[a], Moscati Livia[b], Mancini Stefania[c], Diverio Silvana[a]

[a] Laboratory of Ethology and Animal Welfare (LEBA), Department of Veterinary Medicine, University of Perugia, Via San Costanzo 4, 00126, Perugia, Italy.

[b] Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche "Togo Rosati", Via Salvemini 5, 06126, Perugia, Italy

[c] Public Veterinary Services for Urban Hygiene and Prevention of Stray Dogs, USL Umbria 1, Municipal Rescue Dog Shelter, Strada per Brufa snc, 06148 Collestrada (Perugia), Italy

∗Corresponding author. E-mail addresses: silvana.diverio@unipg.it (S. Diverio)

Highlights

- A fast and straightforward welfare assessment system for shelter dogs was developed.
- Qualitative assessments of Veterinary Officer were functional and consistent
- Dog's behaviour during the capture was the most problematic to define and evaluate
- Behavioural evaluations could improve management and welfare of shelter dogs

Abstract

This study aimed to develop and validate a fast and straightforward welfare assessment system to help shelter staff in decision-making processes. For newly captured dogs entering the shelter, the animal control officer (at

capture time) and the veterinary officer (at entrance examination) compiled a form attributing a qualitative score for the overall *Stress level* and for 5 *Descriptors of dog behavioural traits*. Furthermore, a tester filled out the same form by subjecting the dog to a test battery in his pen, besides performing behavioural observations. The veterinary officer and the tester repeated the evaluation protocol after four weeks of acclimatisation in the shelter. The analysis evaluated inter-observer and test–retest reliability, internal consistency, and construct validity of qualitative scores. Overall, we collected 258 forms regarding 189 dogs. Principal component (PC) extracted by Descriptors showed a good correlation with the *Stress level* score for veterinary and tester confirming the internal consistency of these scales while it was low for animal control officer form. Moreover, qualitative evaluations of the veterinarian and the tester showed congruent correlations with behavioural observations supporting their construct validity. Conversely, the scores expressed by the animal control officer were not consistent with quantitative observations. Then, the veterinary officer and tester forms could be validated and further simplified including only *Stress level* score while the control officer form requires a revision as it does not seem reliable. We did not find agreement between the *Stress level* scores expressed by animal control officer, veterinary and tester suggesting that the three contexts represent different stress stimuli to which the same dog reacts differently. The point of view of the three evaluators can increase the reliability of the assessment. Static but vigilant behaviours prevailed in newly sheltered dogs but activity, interactions and behavioural diversity increased in the second behavioural observation when the dogs were kept in multiple cages and four weeks of adaptation had passed. *Stress level* reduced and PCs tended towards the pole of sociability suggesting a reduction of stress after the period of adaptation to the socio-environmental conditions of the shelter. On the other hand, the considerable inter-individual variability in behaviours reflects differences in coping strategies and or in the manifestation of stress. Our simple tool can not replace a multidisciplinary approach to welfare assessment but could help shelter staff for individual management of dogs complying with their different adaptation skills.

Keywords: shelter dogs; qualitative behaviour assessment; stress; behaviour.

1. Introduction

In some European countries, such as Italy and Austria, the law forbids the euthanasia of free-ranging dogs, unless they are suffering from incurable diseases or are proven to be dangerous to human society ("National Italian Law 281 of 1991," 1991). Despite the numerous positive ethical aspects, this "no-kill policy" has lengthened the duration of dogs staying inside the shelters, increasing the number of dogs housed and public costs. Unfortunately, it has also reduced dogs welfare when poor shelter management and or overcrowding do not guarantee an adequate standard of living, housing and healthcare (Cafazzo et al., 2014; Menchetti et al., 2015). Dogs can thus remain in a poor welfare condition for life. Welfare condition seems also linked to adoptability. A dog well adapted to the environment in which he lives tend to be more interactive with people and other animals, as well as less fearful, aspects generally considered attractive by potential adopters (Menchetti et al., 2015; Taylor and Mills, 2007; Wells and Hepper, 2000). Then, especially in the countries which have implemented a no-kill policy, a welfare assessment system of shelter dogs is essential not only to increase their wellbeing but also their chances of adoption (Barnard et al., 2016; Cafazzo et al., 2014; Taylor and Mills, 2007). Besides, behavioural evaluations of shelter dogs could enhance the safety of both the animals and staff.

Several authors evaluated the welfare in shelter dogs (Cafazzo et al., 2014; Hewison et al., 2014; Hiby et al., 2006; Part et al., 2014; Titulaer et al., 2013; Walker et al., 2016). However, their experimental protocols provided the collection of physiological and behavioural data involving the use of economic, human, and time resources. Welfare assessment systems based on the integration of multiple parameters are more reliable but not affordable and feasible on a large scale. Some authors have assessed the welfare of dogs in kennels using only behavioural assessments (Arena et al., 2017; Protopopova et al., 2014), others have evaluated some tests for predicting behavioural adoptability of shelter dogs (for a review, see Haverbeke et al., 2015). In other studies, training and socialisation programmes have been used as a tool to increase dogs adoption rates (Luescher and Tyson Medlock, 2009; Menchetti et al., 2015; Normando et al., 2009). Then, the attention to the welfare of shelter dogs seems to have grown, but the tools for assessing it are still not routinely applicable. Recently, Barnard et al. (2016) have developed a multifunctional protocol including several indicators concerning shelter management and resources, as well as health and appropriate behaviour of the animals. Their protocol is a useful

tool for objective comparisons between shelter facilities and to direct legislators towards guidelines for shelter animals. Kiddie and Collins (2014) proposed a scoring system for assessing the quality of life of shelter dogs based on qualitative assessment of behaviour. Qualitative assessment seems the most functional approach for use by shelter staff as it is not only flexible on dogs morphological and environmental variations but also easy and timesaving (Wemelsfelder, 2007; Wemelsfelder et al., 2009; Walker et al., 2016). However, as highlighted by Barnard et al. (2016) and Kiddie & Collins (2015), time constraints and limited staff availability, as well as the complexity of the evaluation form, could represent a real obstacle to the systematic use of welfare assessment tools in dogs shelters. Also, a qualitative assessment is a subjective tool that can be valuable only if it provides reliable and valid information (Meagher, 2009).

Considering practicality, we have developed a screening protocol that can be easily used and interpreted to help shelter staff in decision-making processes concerning newly captured dogs. To make the system self-sustainable, it had to involve the shelter staff and be fast and straightforward applicable, in order to not hinder their routine activity. Thereby, the shelter animal control and the veterinary officers had to evaluate the dog through qualitative scores but the questionnaire to complete was only one page long. Compared to the systems used previously, our protocol was planned to be simpler, shorter and timesaving for favouring its application on a large scale. Moreover, for the first time, our protocol involved both the animal control officer and the veterinary officer. The evaluation of the animal in multiple contexts and by several operators could increase the reliability of the judgment, in addition to raising the staff's awareness. Finally, our protocol also involved an external tester who compiled the same form besides carrying out behavioural observations of the dogs. Therefore, the protocol validity could be evaluated by testing inter-observer and test–retest reliability, internal consistency, and construct validity of the qualitative scores.

This study aimed to evaluate the applicability and validity of this newly developed protocol as a fast and straightforward qualitative behavioural assessment system in shelter dogs, possibly identifying its critical points and corrective actions. This protocol was applied at dogs entering the shelter, and it was repeated after four weeks. We wanted to evaluate the adaptive responses of the dogs to the real socio-environmental condition of the shelter. Then, no changes were introduced to the routine management procedures of the facilities. This is

applied research, and its final goal is to promote an early evaluation system for the dogs to improve their management in the shelter and, therefore, their welfare.

2. Material and Methods

The study is part of a broader project, named "RandAgiamo®", which aims to increase the adoptability of adult shelter dogs in the Umbria region of Italy (for details, see Menchetti et al., 2015 and Diverio et al., 2016). Moreover, this work was supported by the Italian Ministry of Health ("Evaluation of animal welfare in confined animal population: pilot study for the welfare evaluation model construction in sheltered dogs" – Ricerca Corrente IZSUM RC122014).

2.1. Research design

The protocol provided for the behavioural assessment of dogs by three different evaluators:

- the animal control officer (shelter staff)
- the veterinary officer (shelter staff)
- the external tester (a researcher trained to collect behavioural data in field situations, with a solid scientific background on animal behaviour (in particular dogs) and biological sciences).

The animal control officer carried out a single behavioural assessment for each dog based on the approach during the capture phases. The veterinary officer and the tester evaluated inside the consultation room and the dog's pen, respectively. They evaluated the dog within three days of shelter intake (1st sampling time, T1) and, if the dog was still present in the shelter, they repeated the same evaluation after four weeks of acclimation (2nd sampling time, T2).

For each evaluator, we developed a Qualitative Evaluation Form, and we adapted it to the role the subject played in the shelter. The Qualitative Evaluation Forms were called *Animal Control Officer*, *Veterinary Officer*, and *Tester Evaluation Form* by the respective evaluating subject. Besides, the tester performed also behavioural observations of the dogs.

Only dogs captured by the controller were included in the study. Therefore, dogs brought to the shelter by other people or transferred from other facilities were excluded, including dogs with a behavioural history of aggression. Dogs that showed severe aggression before and during the capture were excluded because they were differently managed following specific dangerous dog protocols, that included safety precautions and low-handling risk measurements.

All the evaluators involved in the protocol were trained before the start of the study. Training included two four-hour class lessons and a practical stage with individual assistance during the first applications of the protocol (the length of the stage ranged from one to three days according to the individual needs of the training evaluator).

2.2. The Qualitative Evaluation Forms

In the Forms for qualitative behavioural assessment of the dog behaviour, the evaluators (animal control officer, veterinary officer, and tester) rated their judgment on animal's behavioural expression, using qualitative descriptors that reflect the animals' emotional state (Wemelsfelder, 2007).

The Qualitative Evaluation Forms developed for the three evaluators had a common structure so they could be parsed and compared. All the Forms contain (Fig. 1):

- ✓ Essential dog demographic and clinical information: breed, size, coat, sex, presumed age, physical conditions, particular signs (the evaluator assigned breed and morphological data, sex and presumed age following specific internal shelter procedures and dispositions on how to categorize and recognize newly entering dogs in the routine practices; physical conditions included malnutrition, the presence of wounds or traumas, lameness and all other health aspects that could be detected from the first external

observation, in addition with the reproductive state; particular signs included both morphological peculiarities and health problems).

- ✓ Two different approaches to qualitative behavioural assessments:

  1- Stress level. An overall evaluation of the dog's reaction mainly referring to the stress level, obtained by attributing a score from 0 (extremely calm and sociable) to 5 (extremely stressed; Overall, 2013). The score was alongside a brief description and a schematic graphic representation of the dog's behaviour, similar to those used in some Pain Scale (Mich et al., 2010).

  2- Descriptors of behavioural traits. An assessment of specific behavioural traits by attributing a score from 0 (absence of the behaviour, i.e. non-sociable) to 5 (extreme presentation of the behaviour, i.e. extremely sociable). The descriptors (sociability, calmness, fear, excitability, aggressiveness) were chosen based on animal temperament and personality studies, in particular in the literature describing behavioural traits of shelter dogs (Dowling-Guyer et al., 2011; Mornement et al., 2014; Svartberg and Forkman, 2002).

For each session, the specific behavioural traits were evaluated in two different phases by the animal control officer (*Animal Control Officer Form*) and by the veterinary (*Veterinary Officer Form;* Table 1). The animal control and veterinary officers involved in the evaluation process were not required to make modifications or interventions other than the ordinary routine procedures implied in their duties. Therefore, the animal control officer evaluated the dog at the capture time in the place where he found the dog, whereas the veterinary personnel evaluated the same dog during the routine clinical evaluation held in the consultation room when entering the shelter.

Instead, the tester attributed the score to each behavioural trait by submitting the dog to a standard behavioural test, consisting in a battery of 8 sub-tests (*Tester form;* modified by Arhant and Troxler (2014) and Bollen and Horowitz (2008); Table 1). The tester conducted the tests in the pen in which the dog was housed, between 9:00 and 12 a.m.. On the first sampling time (within three days of entry into the shelter, T1), the dogs were preventively housed individually, as is the norm for newly captured (average size of a single dog cage was 3 m x 1.5 m). In one case, a bitch was located in the cage with her three puppies, while in another case two dogs

were kept in the same cage since they were caught together (presumable brothers). All pens were provided by a front gate and a doghouse. All pens had a roof to give protection from the sun and general weather conditions. A net, through which neighbouring dogs can be seen, separated the pens of each row. On the second sampling time (after four weeks of staying in the shelter, T2), the dogs were housed in multiple (2-4 dogs) fenced areas (pens), which size ranged between 15 and 40 m^2, according to the internal shelter procedures. The multiple cages had one kennel per dog, part of which was roofed. Dogs were fed once in the morning, and fresh water was available at all times.

Two independent testers (both females) simultaneously evaluated five dogs in succession to determine the inter-rater reliability in identical observable situations of the *Tester Form*. The external testers were two researchers trained to collect behavioural data in field situations, with a solid scientific background on animal behaviour (in particular dogs) and biological sciences.

2.3. Behavioural observations

The tester was positioned in an area outside the pen and did not interact with the dog (non-manipulative method). Each observation session had a duration of 15 minutes. The tester recorded the dogs' behaviour (Table 2) by using the *scan animal sampling* method at 10-second intervals. The Dog Ethogram has been adapted from Beerda et al. (1998). All behaviour was quantified in duration and expressed as a percentage of time within each session.

The diversity of behaviour patterns performed was calculated for each dog within each session for Main Behaviour class using the Shannon Diversity Index (H Index; Shannon and Weaver, 1949) according to the following formula:

$H = -\Sigma (p_i * \ln p_i)$

where p_i is the proportion of time engaged in the i-th behaviour. The value of H increases with the number of behavioural patterns performed and with equality of time spent engaged in each behavioural pattern. Lower values represent less behavioural diversity (Kistler et al., 2009; Part et al., 2014).

2.4. Validity of the Qualitative Evaluation Forms

The validity of the *Qualitative Evaluation Forms* was evaluated by testing for: (i) the agreement between multiple people independently rating the same dog, both in identical observable situations and in different contexts (inter-observer reliability); (ii) the agreement between results conducted at two different times (test–retest reliability); (iii) the associations between the *Stress level* scale and multiple *Descriptors of behavioural traits* (internal consistency); (iv) the associations between qualitative evaluations and quantitative ethogram-based measures (construct validity; Meagher, 2009).

2.4. Data analysis

We analysed the distributions within demographic data by using Chi-square goodness-of-fit tests. We used the Spearman coefficient (ρ) to evaluate correlations between variables, while we used the Cohen's κ and Kendall's τ tests for evaluating inter-observer agreement for nominal and ordinal data, respectively. A coefficient ±0.1 represent a small, ±0.3 a medium and ±0.5 a large effect size (Field, 2009). We analysed the *Descriptors of behavioural traits* included in the *Qualitative Evaluation Forms* by Principal Component Analysis (PCA) with Varimax Rotation. We chose to extract one component to compare the *Descriptors* with the score of *Stress level* and, then, to evaluate the internal consistency (Meagher, 2009). We calculated corresponding PC scores using the Regression method. We evaluated the effect of the number of sampling time (2 levels: T1 and T2) on *Stress level*, PCs scores and H Index by Generalized Linear Model (GLM) procedures. The *Stress level* was treated as an ordinal variable and cumlogit link function and multinomial distribution were used in the GLMs. Normal distribution and identity link function were used for PC and H Index analysis. Results were expressed as odd ratio (OR) with corresponding 95% confidence intervals (CIs) and P values (from Wald statistic).

We first used the descriptive statistics and Chi-square goodness-of-fit test for behavioural observations. Then, to evaluate the changes from the first to the second sampling time, we coded the behavioural variables by using the one-zero method and we analysed them by using the Generalized Estimating Equations procedure with an exchangeable correlation structure and robust standard errors. Binomial and logit were the probability distribution and the link function, respectively (Diverio et al., 2017). The dog and the sampling time were included in the model as the subject and within-subject factors, respectively. The results were expressed as

estimated marginal means ± standard error (SE), but the raw data are shown in the figure. All data were analysed using SPSS Statistics, version 23 (IBM, SPSS Inc., Chicago, IL). Although $P \leq 0.05$ was the accepted level of statistical significance, trends between $P > 0.05$ and $P < 0.1$ have been also presented and discussed.

3. Results

3.1. Characteristics of sample

From March 2016 to May 2017, 258 forms regarding 189 dogs were collected: 106 *Animal Control Officer*, 64 *Veterinary Officer,* and 43 *Tester Forms.* Moreover, the tester performed 43 behavioural observations. Although four shelters were involved in the project, over 70% of the forms were collected by one single shelter. Therefore, since the distribution of our sample among the different shelters was very unbalanced, we have not analysed the "shelter" effect. The tester evaluated 24 dogs (19 of which were also available for the second sampling time) while the veterinary officers evaluated 46 dogs (18 of which were also available for the second sampling time). Table 1SM shows demographic data recorded in the forms. The majority of dogs were males ($\chi^2 = 9.3$, $P < 0.01$), medium size ($\chi^2 = 64.4$, $P < 0.001$) and mixed ($\chi^2 = 714.6$, $P < 0.001$). Concerning the coat colour, higher frequencies were observed for red, black and white, black, gold, and fawn ($\chi^2 = 23.1$, $P < 0.01$). Median presumed age was 24 months (Q1=12 months, Q3= 36 months). We were able to conduct the complete test for each subject, since the tester found no safety impediment to enter inside the cage of the dogs (for instance, in case of aggressive or phobic subjects).

3.2. Stress level and Descriptors of behavioural traits: inter-observed reliability (in different contexts), internal consistency and test-retest reliability

We did not find agreement between the *Stress level* scores assigned by the three evaluators who assessed the dogs in different contexts (inter-observed reliability in different contexts): $\tau = -0.400$ ($P > 0.1$) for Tester-Animal

Control Officer, $\tau = 0.183$ ($P > 0.1$) for Tester-Veterinary Officer, and $\tau = 0.242$ ($P > 0.1$) for Animal Control Officer-Veterinary Officer.

The PCA on the *Descriptors of behavioural traits* extracted the components called Animal Control Officer PC, Veterinary Officer PC, and Tester PC from the Animal Control Officer, Veterinary Officer and Tester Qualitative Evaluation Forms, respectively (Table 3). In Animal Control Officer and Veterinary Officer PCs, we observed positive loadings for behavioural traits indicating Excitability, Aggressiveness and Fear while only the Fear had a positive sign in the Tester PC. In all three components, Calmness and Sociability had negative loadings. We did not find significant correlations between the PC scores of the forms compiled by the various evaluators: $\rho = -0.600$ ($P > 0.1$) for Tester-Animal Control Officer, $\rho = 0.240$ ($P > 0.1$) for Tester-Veterinary Officer, $\rho=0.036$ ($P > 0.1$) for Animal Control Officer- Veterinary Officer.

The PCs extracted from *Descriptors of behavioural traits* were compared with the score of *Stress level* to evaluate the internal consistency. Correlation coefficients indicated a medium association for Animal Control Officer scores, a large association for Veterinary Officer and Tester scores (Table 3).

Finally, changes in qualitative assessments after four weeks of staying in the shelter measured the test-retest reliability. We found, although as a trend, a reduction on both Stress level (OR = 0.417, 95%CI = 0.155-1.124, P = 0.084) and PC score (OR = 0.578, 95%CI = 0.317-1.057, P = 0.075) of Tester Forms from the first to the second sampling time (Figure 2, Panels A and B). However, we found a high variability of these scores, especially for the second sampling time (Figure 2, Panels A and B).

3.3. Quantitative ethogram-based measures: overall behaviour in the shelter, test-retest reliability and inter-observed reliability (in identical observable situations)

Overall, dogs spent most of the time standing (33%; $\chi^2 = 7435.6$, $P<0.001$), followed by lying down watchful (22.5%), sitting (17.8%), walking (8.1%), lying down sleeping (3.0%), grooming (2.4%) and doing in circles or back-behind stereotyped movements (2.4%). The other categories of the class *Main behaviours* recorded percentages lower than 2.0%. Among the *Occasional behaviours*, only the vocalizations (total = 5.5%; mean per session ± standard deviation= 5.3±0.1%) were found in percentages above 2.0%. Most of the time the dogs

assumed a neutral posture (85.7%; $\chi^2 = 4571.6$, $P < 0.001$), followed by low (11.4%) and high posture (2.8%). Half the time the dogs remained at a distance of 1-5 meters from the tester (50.3%; in contact with the fence = 28.0%, within 1 m = 21.7%; $\chi^2 = 506.2$, $P < 0.001$) and in the anterior area of the pen (49.8%; posterior area = 28.1%, inside the kennel = 22.2%; $\chi^2 = 472.9$, $P < 0.001$).

The models that evaluated the effect of the sampling time on the behaviours showed a reduction of the time spent sitting (OR = 0.232, 95%CI = 0.055-0.972, P = 0.046) and moving in circle (OR = 0.090, 95%CI = 0.003-2.739, P = 0.167; Figure 3) at T2. Conversely, on the second sampling time increased the time in which the animal walks (OR = 2.956, 95%CI = 0.895-4.725, P = 0.090), runs (OR = 3.590, 95%CI = 1.527-8.443, P = 0.003), jumps to the walls (OR = 4.026, 95%CI = 1.353-11.978, P = 0.012), interacts with other dogs (OR = 6.713, 95%CI = 1.172-38.434, P = 0.032) and drinks (OR = 3.255, 95%CI = 0.983-10.781, P = 0.053).

No difference was found for the dog's vocalizations (4.0±1.2% and 7.7±3.5% during T1 and T2, respectively; OR = 2.011, 95%CI = 0.549-7.365, P = 0.292).

As regards the posture (Figure 2SM), during second sampling time increased the neutral (76.6±7.0% and 97.8±1.9% during T1 and T2, respectively; OR = 13.568, 95%CI = 2.196-83.834, P = 0.005) while reduced the high (4.7±1.8% and 1.4±1.2% during T1 and T2, respectively; OR = 0.296, 95%CI = 0.053-1.656, P = 0.166) and low posture (17.4±6.9% and 1.1±1.7% during T1 and T2, respectively; OR = 0.052, 95%CI = 0.002-1.561, P = 0.089).

Moreover, during second sampling time the dogs spent more time in the posterior area of the pen (20.1±5.3% and 37.7±7.6% during T1 and T2, respectively; OR = 2.398, 95%CI = 0.943-6.093, P = 0.066) but less time inside the kennel (29.7±8.1% and 11.9±4.8% during T1 and T2, respectively; OR = 0.288, 95%CI = 0.069-1.205, P = 0.088).

Finally, the differences were not significant for the variables included in the "Distance to tester" class.

Dogs also showed a greater diversity of behaviours (H Index) at T2 compared with T1 (0.92±0.1 and 1.20±0.1 during T1 and T2, respectively; OR = 1.321, 95%CI = 1.123-1.554, P = 0.001; Figure 3SM).

Statistical tests to evaluate the inter-observer agreement between the two testers indicated a fair agreement for both behavioural observations ($\kappa = 0.559$; $P < 0.001$) and qualitative scores ($\tau = 0.490$ $P < 0.001$).

3.4 Qualitative evaluations and behavioural observations: construct validity

We found several correlations between the behavioural observations and the qualitative evaluations of the tester (Table 4). Stress level expressed by the tester was negatively correlated with Grooming ($P<0.05$), Interactions with other dogs ($P<0.01$), Eating ($P<0.01$), Neutral posture ($P<0.05$), and H Index ($P<0.05$), while it was positively correlated with In circles or back-behind stereotyped movements ($P<0.05$) and Low posture ($P<0.05$). PCs extracted from *Descriptors of behavioural traits* of the tester also showed positive correlation with Inside the kennel variable ($P<0.05$). Qualitative evaluations of the Veterinary Officers showed positive correlations with Low posture and Lying down watchful ($P<0.05$), negative correlations with Standing ($P<0.01$). Scores positive expressed by the Animal Control officers were associated with Walking, Standing, and Grooming ($P<0.05$).

4. Discussion

The primary goal of this study was to set up and validate a sustainable welfare assessment system for newly entered shelter dogs. In this regard, we have developed a *Qualitative Evaluation Forms* for three professional figures who could interact with the dog during his capture and shelter experience: the animal control officer, the veterinary officer and a behavioural assessor (tester). By using similarly structured Qualitative Evaluation Forms, the three professional figures had to attribute a score for the general level of the animal stress (*Stress level*) and scores for specific *Descriptors of behavioural traits*. Inter-observer and test–retest reliability, internal consistency, and construct validity of qualitative scores were evaluated.

The Principal Component Analysis on the *Descriptors of behavioural traits* extracted components characterised by positive loadings for behavioural traits that indicate excitability, aggressiveness and fear while negative loadings for calmness and sociability. The component extracted from the analysis of the Qualitative Evaluation Forms compiled by the three professional figures resembled each other and can be described as an *Aggressiveness-Sociability axis*. A similar component has been identified in the dog by other authors (Hennessy et al., 2001; Netto and Planta, 1997) although more often Sociability, Aggressiveness and Fear emerged in

separate factors (Hsu and Serpell, 2003; Jones and Gosling, 2005; Menchetti et al., 2018; Svartberg and Forkman, 2002). However, our work aimed to evaluate a fast and straightforward system to identify only the dog's main behaviour. The multiple descriptors were used primarily to assess the internal consistency of the Stress level scale (Meagher, 2009). For this reason, we carried out a selection of a few behavioural descriptors producing only one broad dimension. Findings showed a good correlation of Descriptors PCs extracted by the Veterinary Officer and the Tester Forms with the respective scores of *Stress level*. Moreover, the behavioural observations supported the qualitative evaluations of the veterinary officers and the tester. Indeed, low *Stress level* scores were correlated with dog social behaviours while high scores with behaviours indicative of stress, such as stereotypical movements and low postures (Beerda et al., 1999a; Cafazzo et al., 2014; Dowling-Guyer et al., 2011). Then, the Veterinary Officer and Tester Forms showed good internal consistency, and thereby these could be further simplified including only the Stress level score. Conversely, the correlation between the Descriptor's PC and *Stress level* was lower for the Animal Control Officer Forms questioning their internal consistency. Moreover, the scores expressed by the animal control officer during the capture phase were not consistent with the behavioural observations into the pen suggesting a low construct validity. Although all the evaluators had been trained to score the dogs qualitatively, it might be possible that the Animal Control Officers may be less familiar with behavioural assessments than the veterinarians or the testers. Furthermore, the behaviour of the dog during the capture phase could be more complicated to define. As far as we know, there is no study on the behaviour and welfare at the capture time of stray dogs. However, as recognised by the World Organisation for Animal Health (OIE, 2015) and by the Food and Agriculture Organization of the United Nations (FAO, 2014), the process of catching is very stressful and can have a significant impact on dogs' welfare. Therefore, although the current procedures try to minimise stress, the evaluation of the dog behaviour during capture requires the development of a specific detection system. Our findings suggest that the *Qualitative Evaluation Forms* for the Animal Control Officer should be reviewed, for example including more objective behavioural descriptors. Besides, this Form would probably be more beneficial being diversified from that of the Veterinary Officer and Tester.

In our study, we did not find inter-observed reliability between the qualitative evaluations expressed by the three professional figures in different contexts. The agreement between people in scoring stress is not obvious, being

influenced by their training, personal experience, subjective interpretations, and duration of interactions with the animal (Lind et al., 2017; Meagher, 2009). However, we found a fair inter-observer agreement between the two testers who evaluated the same observational situation as well as congruent associations between qualitative scores and ethogram-based measures of the tester. This suggests that this qualitative behavioural assessment had an acceptable accuracy and construct validity (Meagher, 2009; Walker et al., 2016; Wemelsfelder et al., 2009). Although further training of the evaluators could improve the use of our forms and enhance the inter-observer reliability, further hypotheses could be evaluated. The three evaluations were carried out in contexts that represent different psychological stressors for the dog. Besides, each dog may react differently to the approach and containment in the place of capture, inside the veterinary clinic and the pen. The stress, as the pain (Sharkey, 2013), is subjective, dynamic, and multidimensional. The behaviour is likely to be influenced not only by the personality and the breed of the dog but also by the environment and his past experience, as well as his current welfare status (Hiby et al., 2006; Jones and Gosling, 2005; Protopopova, 2016; Rooney et al., 2007; Walker et al., 2016). It is therefore plausible that the same dog can react differently to different stress stimuli. Thus, the judgment on his reaction must be modulated by the context in which it is evaluated. For example, several studies have shown that dogs in a veterinary clinic exhibit different physiological changes, such as heart rate and temperature (Bragg et al., 2015), and behavioural reactions from those shown in other environmental contexts (Döring et al., 2009; Lind et al., 2017; Mariti et al., 2015). Future studies could provide behavioural observations also during the capture phase and veterinary visit. In any case, the point of view of the three evaluators can increase the reliability of dog assessment because they test the behavioural reactions in different environmental contexts.

Our protocol provides behavioural observations within the pen where, generally, static but vigilant behaviours prevailed. In fact, dogs spent most of the time standing, lying down watchful and sitting. The replication of the evaluations after four weeks of acclimatisation allows evaluating the effect of socio-environmental context of the shelter on dog welfare. Indeed, compared with the first, during the second sampling time, there was a reduction in time spent sitting and of stereotyped movements, while the time in which the dogs walk, run, jump to the walls, and drink increased. During the second sampling time, the dogs increased not only their activity levels but also their interactions and the time spent eating. They showed a more diversified repertoire of

behaviours, as well as a reduction of time spent in a low or high posture and inside the pen. These findings are likely to reflect an improved welfare state over time in the shelter (Beerda et al., 1999b; Kiddie and Collins, 2015; Kistler et al., 2009; Part et al., 2014). Previous studies report contrasting effects of the length of time spent in the shelter on the activity level (for a review, see Protopopova, 2016). Besides, both inactivity and increased movement have been connected with stress (Beerda et al., 2000; Hiby et al., 2006; Kiddie and Collins, 2015). Protopopova et al. (2014) did not find consistent behavioural changes due to time spent at the shelter. Part et al. (2014) found that dogs held in shelter reduced the time lying down, increased time travelling and Shannon Diversity Index compared to the same dogs in a home environment. Other studies observed a reduction in activity levels in relation with time spent in the shelter. This change suggests a reduced motivation to interact with the surroundings, thereby as a potential indicator of reduced welfare (Rooney et al., 2007; Titulaer et al., 2013; Wells and Hepper, 2000). Conversely, Walker et al. (2016) found reduced standing in dogs after 30 days in shelter suggesting an on-going adjustment.

In our study, the two visits were performed in a different housing. Therefore, behavioural changes are not due only to the "time" factor but rather to the summation of time and environment. This is a critical fact from the experimental point of view because we can not separate the effect of these two factors. However, our research was applied to the real management of the shelter. Standard shelter procedures require dogs, after an initial period of isolation, to be moved into multiple pens and we had not introduced artifices. Physical contact with the other dogs stimulated the activities and reduced the apathy that characterises the first days of detention in the shelter. Our results confirm that group housing is motivating for the animals and crucial in the adaptation processes to the shelter suggesting that social isolation should be limited, as much as possible, over time (Barnard et al., 2016; Beerda et al., 2000; Taylor and Mills, 2007). Also, qualitative assessments suggested an increase in dog's sociability during the second sampling time. The environment should have less influence on qualitative assessment than quantitative observations. As suggested by Wemelsfelder (2007), a qualitative judgment "*tends not to depend on quantity — it is not necessarily true that the more there is of something the better it gets. For example, if animals are provided with greater levels of stimulation in their environment, leading to higher levels of activity, this does not necessarily mean that their welfare improves. What matters to quality is how things are done, their style; quality is a dynamic notion*". Furthermore, the convergence of the

behavioural observations, although with a possible bias, and the qualitative scores support the idea of a reduction of stress after four weeks of acclimation in the shelter.

However, it is essential to emphasise the high inter-individual variability in behavioural patterns, especially during the second sampling time. This aspect emerges in several studies carried out inside the shelter as a reflection of individual coping strategies and/or differences in the behavioural manifestation of stress (Part et al., 2014; Rooney et al., 2007; Titulaer et al., 2013). Hiby et al. (2006) evaluated the physiological and behavioural changes of dogs entering re-homing shelters finding different cortisol responses in the long-term, as well as large between-dog variation. Titulaer et al. (2013) suggest that the welfare of the dog is not so much determined by the time in shelters but rather by the individual responses to the new environment. The novel kennel environment, subjecting the dog to different experiences of socialisation, can determine individual answers. For example, breed, previous socialisation experiences and hierarchies influence the adaptation skill in group life (Taylor and Mills, 2007). In a shelter, unlike the life of a stray, a relationship with humans is inevitable, and not all stray dogs can find benefits from this. Indeed, dog-human interactions is considered advantageous from a welfare point of view (Kiddie and Collins, 2015; Wells and Hepper, 2000) but excessive contact or visual access to humans may also be detrimental for fearful dogs (Conley et al., 2014; Hewison et al., 2014; Taylor and Mills, 2007). Effect of contacts with humans in shelters should be further explored in order to improve the management of dogs within the facility.

At that, an early assessment of the intra-and inter-specific socialisation of the dog could help staff in the relevant formation of groups or logistic organisation within the facility (i.e., shelter area more isolated and less noisy rather than those more exposed to the contact with humans). The *Stress level* score that we propose does not replace quantitative behavioural evaluation. Besides, behavioural assessments should be integrated with physiological and pathological measurements to determine the animal's welfare. However, a simple quantitative assessment responds to the needs of the economy and timesaving representing a compromise to the void that exists, also in terms of sensitivity, in many shelters nowadays. The use of the protocol could also promote the shelter staff to pay more attention to the psycho-physical comfort of the animals, expecially in critical situations such as the capture and the veterinary visit. In practice, our protocol could be used in order to identify: (i) the most problematic dogs that should be helped to adapt to the shelter environment and/or follow specific training;

(ii) aggressive dog that could be handled carefully by the shelter staff and require the intervention of animal behaviourist; (iii) the most sociable dogs that could follow rapid paths of adoption; and (iv) the overall level of dogs' welfare within the shelter to carry out any corrective actions in the management. For these purposes, our *Stress level* could be inserted in the individual file besides the other clinical information of the dog while the medium *Stress level* of the animals kept in the shelter could represent a general index of wellbeing within the facility.

5. Conclusion

We have developed a Multi-operator Qualitative Behavioural Assessment System for dogs entering shelter involving animal control officers, veterinary officers and testers. The collecting protocol has been designed to be of easy and straightforward use for the staff, thus not hinder their routine activities. Veterinary Officer and Tester Forms showed good internal consistency and construct validity suggesting that they can be validated and further simplified including only a *Stress level* score. Conversely, an implementation seems necessary for the Animal Control Officer Form, including specific parameters and scores to evaluate the dog capture phase. There was no inter-observer agreement between the qualitative evaluations expressed by the three professional figures. Further training of shelter staff can improve the use of qualitative behaviour assessment. However, different contexts could represent different psychological stressors for the dog, which are able to trigger diversified adaptive responses. Both qualitative and quantitative assessments suggested an improvement in dog welfare in the second behavioural observation, when the dogs were kept in multiple cages and four weeks of adaptation had passed. However, also increased the inter-individual variability in their behaviour. The *Stress level* score does not claim to replace multifunctional approaches in the assessment of welfare. However, it could stimulate more careful management of dogs entering the shelter. Moreover, periodic planning of assessments could monitor the dogs' condition within the shelter highlighting critical issues, leading to significant benefits for both staff and key players.

Conflict if interests

None.

Acknowledgements

The authors wish to thank Giovanni Giovannini, Anna Rita Flamini, Roberto Giannelli, Aliuscia Pelucca, and Maria Donata Giaimo of the Umbria Region for promoting and supporting the RandAgiamo® project. We would like gratefully acknowledge all the people who helped and supported this study, in particular Claudia Petroni who helped to collect data, all the personnel of the Canile Sanitario of Collestrada, Città di Castello, Todi e Assisi (Perugia). And not at least, we would like to thank all the RandAgiamo® staff (Maria Chiara Catalani, Beatrice Boccini, Martina Iaboni, Silvia Amato, Roberta Gobbi) and all volunteers who greatly and generously dedicate their time and commitment for improving the quality of life and chance of adoption of shelter dogs.

We also thank the anonymous reviewers, the comments from whom enabled us to substantially improve the original manuscript.

This work was supported by the Italian Ministry of Health ("Evaluation of animal welfare in confined animal population: pilot study for the welfare evaluation model construction in sheltered dogs" – Ricerca Corrente IZSUM RC122014).

References

Arena, L., Wemelsfelder, F., Messori, S., Ferri, N., Barnard, S., 2017. Application of Free Choice Profiling to assess the emotional state of dogs housed in shelter environments. Appl. Anim. Behav. Sci. 195, 72–79. https://doi.org/10.1016/j.applanim.2017.06.005

Arhant, C., Troxler, J., 2014. Approach behaviour of shelter dogs and its relationships with the attitudes of shelter staff to dogs. Appl. Anim. Behav. Sci. 160, 116–126. https://doi.org/10.1016/j.applanim.2014.08.013

Barnard, S., Pedernera, C., Candeloro, L., Ferri, N., Velarde, A., Villa, P.D., 2016. Development of a new welfare assessment protocol for practical application in long-term dog shelters. Vet. Rec. 178, 18. https://doi.org/10.1136/vr.103336

Beerda, B., Hooff, J.A.N.A.R.A.M.V.A.N., Vries, H.W.D.E., Mol, J.A.N.A., 1999a. Chronic Stress in Dogs Subjected to Social and Spatial Restriction . I . Behavioral Responses 66, 233–242.

Beerda, B., Schilder, M.B.H., Bernadina, W., Hooff, J.A.N.A.R.A.M.V.A.N., Vries, H.W.D.E., Mol, J.A.N.A., 1999b. Chronic Stress in Dogs Subjected to Social and Spatial Restriction . I Behavioural Resposes 66, 243–254.

Beerda, B., Schilder, M.B.H., Van Hooff, J.A.R.A.M., De Vries, H.W., Mol, J.A., 2000. Behavioural and hormonal indicators of enduring environmental stress in dogs. Anim. Welf. 9, 49–62. https://doi.org/urn:nbn:nl:ui:32-313407

Beerda, B., Schilder, M.B.H., Van Hooff, J.A.R.A.M., De Vries, H.W., Mol, J.A., 1998. Behavioural, saliva cortisol and heart rate responses to different types of stimuli in dogs. Appl. Anim. Behav. Sci. 58, 365–381. https://doi.org/10.1016/S0168-1591(97)00145-7

Bollen, K.S., Horowitz, J., 2008. Behavioral evaluation and demographic information in the assessment of aggressiveness in shelter dogs. Appl. Anim. Behav. Sci. 112, 120–135. https://doi.org/10.1016/j.applanim.2007.07.007

Bragg, R., Bennet, J., Cummings, A., Quimby, J., 2015. Evaluation of the effects of hospital visit stress on

physiologic variables in dogs. Javma 246, 212–215. https://doi.org/10.2460/javma.246.2.212.

Cafazzo, S., Maragliano, L., Bonanni, R., Scholl, F., Guarducci, M., Scarcella, R., Di Paolo, M., Pontier, D., Lai, O., Carlevaro, F., Bucci, E., Cerini, N., Carlevaro, L., Alfieri, L., Fantini, C., Natoli, E., 2014. Behavioural and physiological indicators of shelter dogs' welfare: Reflections on the no-kill policy on free-ranging dogs in Italy revisited on the basis of 15years of implementation. Physiol. Behav. 133, 223–229. https://doi.org/10.1016/j.physbeh.2014.05.046

Conley, M.J., Fisher, A.D., Hemsworth, P.H., 2014. Effects of human contact and toys on the fear responses to humans of shelter-housed dogs. Appl. Anim. Behav. Sci. https://doi.org/10.1016/j.applanim.2014.03.008

Diverio, S., Boccini, B., Menchetti, L., Bennett, P.C., 2016. The Italian perception of the "ideal companion dog." J. Vet. Behav. Clin. Appl. Res. 1–9. https://doi.org/10.1016/j.jveb.2016.02.004

Diverio, S., Menchetti, L., Riggio, G., Azzari, C., Iaboni, M., Zasso, R., Di Mari, W., Santoro, M.M., 2017. Dogs' coping styles and dog-handler relationships influence avalanche search team performance. Appl. Anim. Behav. Sci. https://doi.org/10.1016/j.applanim.2017.02.005

Döring, D., Roscher, A., Scheipl, F., Küchenhoff, H., Erhard, M.H., 2009. Fear-related behaviour of dogs in veterinary practice. Vet. J. 182, 38–43. https://doi.org/10.1016/j.tvjl.2008.05.006

Dowling-Guyer, S., Marder, A., D'Arpino, S., 2011. Behavioral traits detected in shelter dogs by a behavior evaluation. Appl. Anim. Behav. Sci. 130, 107–114. https://doi.org/10.1016/j.applanim.2010.12.004

FAO, 2014. Dog Population Management. Report of the FAO/WSPA/IZSAM expert meeting - Banna, Italy, 14-19 March 2011.

Field, A., 2009. Discovering statistics using SPSS, SAGE Publi. ed. London. UK.

Haverbeke, A., Pluijmakers, J., Diederich, C., 2015. Behavioral evaluations of shelter dogs: Literature review, perspectives, and follow-up within the European member states's legislation with emphasis on the Belgian situation. J. Vet. Behav. Clin. Appl. Res. 10, 5–11. https://doi.org/10.1016/j.jveb.2014.07.004

Hennessy, M.B., Voith, V.L., Mazzei, S.J., Buttram, J., Miller, D.D., Linden, F., 2001. Behavior and cortisol

levels of dogs in a public animal shelter, and an exploration of the ability of these measures to predict problem behavior after adoption. Appl. Anim. Behav. Sci. 73, 217–233. https://doi.org/10.1016/S0168-1591(01)00139-3

Hewison, L.F., Wright, H.F., Zulch, H.E., Ellis, S.L.H., 2014. Short term consequences of preventing visitor access to kennels on noise and the behaviour and physiology of dogs housed in a rescue shelter. Physiol. Behav. https://doi.org/10.1016/j.physbeh.2014.04.045

Hiby, E.F., Rooney, N.J., Bradshaw, J.W.S., 2006. Behavioural and physiological responses of dogs entering re-homing kennels. Physiol. Behav. 89, 385–391. https://doi.org/10.1016/j.physbeh.2006.07.012

Hsu, Y., Serpell, J. a, 2003. Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. J. Am. Vet. Med. Assoc. 223, 1293–1300. https://doi.org/10.2460/javma.2003.223.1293

Jones, A.C., Gosling, S.D., 2005. Temperament and personality in dogs (Canis familiaris): A review and evaluation of past research. Appl. Anim. Behav. Sci. 95, 1–53. https://doi.org/10.1016/j.applanim.2005.04.008

Kiddie, J., Collins, L., 2015. Identifying environmental and management factors that may be associated with the quality of life of kennelled dogs (Canis familiaris). Appl. Anim. Behav. Sci. 167, 43–55. https://doi.org/10.1016/j.applanim.2015.03.007

Kiddie, J.L.., Collins, L.M.., 2014. Development and validation of a quality of life assessment tool for use in kennelled dogs (Canis familiaris). Appl. Anim. Behav. Sci. 158, 57–68. https://doi.org/10.1016/j.applanim.2014.05.008

Kistler, C., Hegglin, D., Würbel, H., König, B., 2009. Feeding enrichment in an opportunistic carnivore: The red fox. Appl. Anim. Behav. Sci. 116, 260–265. https://doi.org/10.1016/j.applanim.2008.09.004

Lind, A.K., Hydbring-Sandberg, E., Forkman, B., Keeling, L.J., 2017. Assessing stress in dogs during a visit to the veterinary clinic: Correlations between dog behavior in standardized tests and assessments by veterinary staff and owners. J. Vet. Behav. Clin. Appl. Res. 17, 24–31.

https://doi.org/10.1016/j.jveb.2016.10.003

Luescher, A.U., Tyson Medlock, R., 2009. The effects of training and environmental alterations on adoption success of shelter dogs. Appl. Anim. Behav. Sci. https://doi.org/10.1016/j.applanim.2008.11.001

Mariti, C., Raspanti, E., Zilocchi, M., Carlone, B., Gazzano, A., 2015. The assessment of dog welfare in the waiting room of a veterinary clinic. Anim. Welf. 24, 299–305. https://doi.org/10.7120/09627286.24.3.299

Meagher, R.K., 2009. Observer ratings: Validity and value as a tool for animal welfare research. Appl. Anim. Behav. Sci. 119, 1–14. https://doi.org/10.1016/j.applanim.2009.02.026

Menchetti, L., Calipri, S., Guelfi, G., Catanzaro, A., Diverio, S., 2018. My Dog Is Not My Cat: Owner Perception of the Same Household. Animals 8, 80. https://doi.org/10.3390/ani8060080

Menchetti, L., Mancini, S., Catalani, M., Boccini, B., Diverio, S., 2015. RandAgiamo™, a Pilot Project Increasing Adoptability of Shelter Dogs in the Umbria Region (Italy). Animals 5, 774–792. https://doi.org/10.3390/ani5030383

Mich, P.M., Hellyer, P.W., Kogan, L., Schoenfeld-Tacher, R., 2010. Effects of a Pilot Training Program on Veterinary Students' Pain Knowledge, Attitude, and Assessment Skills. J. Vet. Med. Educ. 37, 358–368. https://doi.org/10.3138/jvme.37.4.358

Mornement, K.M., Coleman, G.J., Toukhsati, S., Bennett, P.C., 2014. Development of the behavioural assessment for re-homing K9's (B.A.R.K.) protocol. Appl. Anim. Behav. Sci. 151, 75–83. https://doi.org/10.1016/j.applanim.2013.11.008

National Italian Law 281 of 1991, 1991. . Gazz. Uff. 203. https://doi.org/10.1017/CBO9781107415324.004

Netto, W.J., Planta, D.J.U., 1997. Behavioural testing for aggression in the domestic dog. Appl. Anim. Behav. Sci. 52, 243–263. https://doi.org/10.1016/S0168-1591(96)01126-4

Normando, S., Corain, L., Salvadoretti, M., Meers, L., Valsecchi, P., 2009. Effects of an Enhanced Human Interaction Program on shelter dogs' behaviour analysed using a novel nonparametric test. Appl. Anim. Behav. Sci. 116, 211–219. https://doi.org/10.1016/j.applanim.2008.10.005

OIE, 2015. Stray dog population control. Terr. Anim. Heal. Code 1–17.

Overall, K.., 2013. Manual of Clinical Behavioral Medicine for Dogs and Cats. Elsevier- Mosby, Saint Louis MO.

Part, C.E., Kiddie, J.L., Hayes, W.A., Mills, D.S., Neville, R.F., Morton, D.B., Collins, L.M., 2014. Physiological, physical and behavioural changes in dogs (Canis familiaris) when kennelled: Testing the validity of stress parameters. Physiol. Behav. https://doi.org/10.1016/j.physbeh.2014.05.018

Protopopova, A., 2016. Effects of sheltering on physiology, immune function, behavior, and the welfare of dogs. Physiol. Behav. 159, 95–103. https://doi.org/10.1016/j.physbeh.2016.03.020

Protopopova, A., Mehrkam, L.R., Boggess, M.M., Wynne, C.D.L., 2014. In-kennel behavior predicts length of stay in shelter dogs. PLoS One 9, e114319. https://doi.org/10.1371/journal.pone.0114319

Rooney, N.J., Gaines, S.A., Bradshaw, J.W.S., 2007. Behavioural and glucocorticoid responses of dogs (Canis familiaris) to kennelling: Investigating mitigation of stress by prior habituation. Physiol. Behav. 92, 847–854. https://doi.org/10.1016/j.physbeh.2007.06.011

Shannon, C., Weaver, W., 1949. The mathematical theory of communication. University of Illinois Press.

Sharkey, M., 2013. The Challenges of Assessing Osteoarthritis and Postoperative Pain in Dogs. AAPS J. 15, 598–607. https://doi.org/10.1208/s12248-013-9467-5

Svartberg, K., Forkman, B., 2002. Personality traits in the domestic dog (Canis familiaris). Appl. Anim. Behav. Sci. 79, 133–155. https://doi.org/10.1016/S0168-1591(02)00121-1

Taylor, K.D., Mills, D.S., 2007. The effect of the kennel environment on canine welfare: A critical review of experimental studies. Anim. Welf. 16, 435–447.

Titulaer, M., Blackwell, E.J., Mendl, M., Casey, R.A., 2013. Cross sectional study comparing behavioural, cognitive and physiological indicators of welfare between short and long term kennelled domestic dogs. Appl. Anim. Behav. Sci. 147, 149–158. https://doi.org/10.1016/j.applanim.2013.05.001

Walker, J.K., Dale, A.R., D'Eath, R.B., Wemelsfelder, F., 2016. Qualitative Behaviour Assessment of dogs in

the shelter and home environment and relationship with quantitative behaviour assessment and physiological responses. Appl. Anim. Behav. Sci. 184, 97–108. https://doi.org/10.1016/j.applanim.2016.08.012

Wells, D.L., Hepper, P.G., 2000. The influence of environmental change on the behaviour of sheltered dogs. Appl. Anim. Behav. Sci. https://doi.org/10.1016/S0168-1591(00)00100-3

Wemelsfelder, F., 2007. How Animals Communicate Quality of Life : The Qualitative Assessment of Behaviour How Animals Communicate Quality of Life : The Qualitative Assessment of Behaviour. Anim. Welf. 16, 25–31. https://doi.org/Article

Wemelsfelder, F., Nevison, I., Lawrence, A.B., 2009. The effect of perceived environmental background on qualitative assessments of pig behaviour. Anim. Behav. 78, 477–484. https://doi.org/10.1016/j.anbehav.2009.06.005

Table 1. Phases evaluated by the Animal Control and by the Veterinary officers, and behavioural tests used by the Tester. For each phase / test, a score from 0 (absence of the behaviour) to 5 (extreme presentation of the behaviour) was assigned to the following *Descriptors of the behavioural traits*: sociability, calmness, fear, excitability, aggressiveness.

Form	Phases / tests	Description
Animal Control officer	First approach/Capture operation	Routine procedures
	Containment and handling	
Veterinary officer	Medications, sutures, manipulation	Routine procedures
	Therapies (for OS, IM, SC, IV)* and microchip application	
Tester	Pen presentation	The tester walks with slow and smooth movements approaching the pen at an angle of 90° to the pen's door (within 30s)
	Indifferent approach outside the cage	The tester remains about 30 cm from the pen's door without making any physical or vocal interaction with the dog for 10s
	Friendly approach outside the cage	The tester speaks to the dog in a friendly voice and slowly brings his hand close to the door of the cage for 10s
	Indifferent approach inside the cage	The tester enters the cage but does not interact with the dog for 10s
	Friendly approach inside the cage	The tester bends over his legs and speaks softly to the dog for 10s
	Handling inside the cage	The tester tries to stroke the dog's back, legs and ears for 10s
	Approach with a game (ball)	The tester attracts the dog's attention then throws a ball across the cage (time allowed 2 min)
	Approach with food	The tester offers the dog food, first handing it with his hand and then placing it into a bowl (time allowed 2 min)

** Route of administration: OS= oral, IM= intramuscular, SC= subcutaneous, IV= intravenous.*

Table 2. Dog Ethogram

Behavioural class	Behavioural variable
Main behaviour[1]	Walking
	Lying down watchful
	Lying down sleeping
	Sitting
	Standing
	Trotting
	Jumping to the tester
	Jumping to the walls / fence
	Smelling the tester (within 0.5 m)
	Exploring with low head
	Grooming
	In circles or back-behind stereotyped movements
	Interactions with other dogs[3]
	Eating
	Drinking
Occasional behaviour	Shaking
	Panting
	Threatening (growling, showing the teeth)
	Aggressive (trying to bite)
	Vocalizations
	Defecating or urinating
Posture[1]	High (body hugely erect, head and tail high, ears pointed forwards)
	Neutral (posture shown by dogs under neutral conditions)
	Low (tail lowered, ears positioned backwards, legs bent)
Pen area[1,2]	1 Posterior
	2
	3 Anterior
	4
	5 Inside the kennel
Distance to tester[1]	0 m (in contact with the fence)
	within 1 m
	1-5 m

[1] The variables included in this class were mutually exclusive.

[2] The pen was virtually divided into five areas: two in the front, opposite the gate and the tester, two in the back, farther from the tester, and the fifth was the kennel (Fig. 1SM).

[3] Including attempts to interaction through the net of adjacent cages.

Table 3. Loadings of Descriptors of behavioural traits of Animal control officer, Veterinary and Tester PCs extracted with the principal component analysis and correlation of PCs with the respective Stress level score.

Qualitative Evaluation Form/PC	**Variable** **Phases / tests**	**Descriptors of behavioural traits**	**PC Loading**	**Correlation with the respective *Stress level* score**[#]
Animal Control Officer	First approach/Capture Time	Excitability	0.817	0.377***
	First approach/Capture Time	Aggressiveness	0.782	
	First approach/Capture Time	Fear	0.752	
	Containment and handling	Excitability	0.655	
	First approach/Capture Time	Calmness	-0.585	
	First approach/Capture Time	Sociability	-0.571	
Veterinary Officer	Medications, sutures, manipulation	Excitability	0.774	0.674***
	Medications, sutures, manipulation	Fear	0.756	
	Medications, sutures, manipulation	Aggressiveness	0.651	
	Medications, sutures, manipulation	Sociability	-0.698	
	Medications, sutures, manipulation	Calmness	-0.655	
Tester	Friendly approach outside the cage	Fear	0.851	0.815***
	Indifferent approach outside the cage	Fear	0.847	
	Approach with food	Fear	0.714	
	First approach	Calmness	-0.716	
	Indifferent approach outside the cage	Sociability	-0.827	
	First approach	Sociability	-0.869	

Only loadings of ≥0.40 or ≤−0.40 are shown.

[#] Spearman ρ coefficient. ***P<0.001.

Table 4. Correlation between quantitative behavioural variables and qualitative behavioural assessments (Stress level and Principal Components extracted from the Descriptors of behavioural traits (Spearman coefficient; **P<0.001, *P<0.05). For simplicity, we have shown only the variables with at least one significant association.

	QUALITATIVE EVALUATION FORM					
	Stress level[1]			**PCs extracted from *Descriptors of behavioural traits***[2]		
QUANTITATIVE BEHAVIOURAL VARIABLE	**Animal Control officer**	**Veterinary officer**	**Tester**	**Animal Control officer**	**Veterinary officer**	**Tester**
Walking				0.900*		-0.397*
Lying down watchful					0.503*	
Standing		-0.584*		0.900*	-0.782**	
Grooming	0.816*		-0.370*			
In circles or back-behind stereotyped movements			0.330*			
Interactions with other dogs			-0.592**			-0.447**
Eating			-0.400**			-0.403*
Posture High						0.339*
Posture Neutral			-0.360*			-0.487**
Posture Low		0.552*	0.315*			0.378*
Inside the kennel						0.406*
H Index			-0.334*			

[1] Score from 0 (extremely calm and sociable) to 5 (extremely stressed)

2 Low loadings for Calmness and Sociability, high loadings Excitability, Aggressiveness and Fear

Figure caption

Figure 1. Basic structure of the Qualitative Evaluation Forms

BASIC STRUCTURE OF THE QUALITATIVE EVALUATION FORMS

Essential dog demographic and clinical information

STRESS LEVEL	Schematic graphic representation of the dog's behaviour	Brief description of the dog's behaviour
0		Extremely calm and sociable (he/she spontaneously approaches the operator, looks for physical contact, all body can be easily manipulated, easy and calm when captured, driven to the van and transported)
1		Calm and sociable (he/she approaches the operator if called, all body can be easily manipulated, easy to be captured, driven to the van and transported)
2		Calm and sociable, he/she shows few signs of stress and alert reactions to some procedures (he/she approaches the operator within three attempts, some parts of the body have to be manipulated with precaution, he/she shows sign of stress and resistance when captured, driven to the van and transported)
3		Sociable but excitable, signs of stress, excitability and alert reactions to some procedures (he/she is captured by the operator within five attempts, he/she shows sign of stress and or fear and or excitability when manipulated in some parts of the body, he/she shows sign of stress and or fear and or excitability when captured, driven to the van, and transported)
4		Stressed, fearful and unsociable, he/she shows sign of stress and fearful reactions to some procedures (he/she is captured after five attempts, he/she shows sign of fear when manipulated, captured, driven to the van, and transported; he/she can shake, whine, pant, show teeth, growl, defecate, urinate)
5		Extremely stressed, aggressive and unsociable, he/she shows sign of stress and aggressive reactions to some procedures (he/she is captured after five attempts, he/she tries to escape or hide, barks or whines; he/she shows sign of stress and aggression when manipulated, captured, driven to the van, and transported; he/she can try to bite, defecate or urinate)

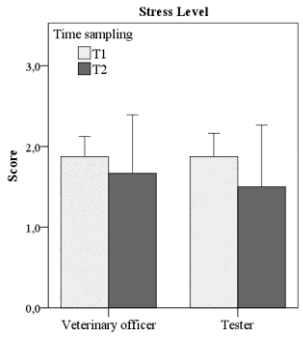
Score for descriptors of the behavioural traits for several phases or tests

Phase / test _____						
Descriptor	**Score**					
Sociability	0	1	2	3	4	5
Calmness	0	1	2	3	4	5
Fear	0	1	2	3	4	5
Excitability	0	1	2	3	4	5
Aggressiveness	0	1	2	3	4	5

Phase / test _____						
Descriptor	**Score**					
Sociability	0	1	2	3	4	5
Calmness	0	1	2	3	4	5
Fear	0	1	2	3	4	5
Excitability	0	1	2	3	4	5
Aggressiveness	0	1	2	3	4	5

Notes _____

Figure 2. *Stress level* (Panel A) and PC scores (Panel B) from Veterinary Officer and Tester Forms for the first (T1, within three days of shelter intake) and second (T2, after 4 weeks of acclimation) time sampling. Values are mean ± standard error.



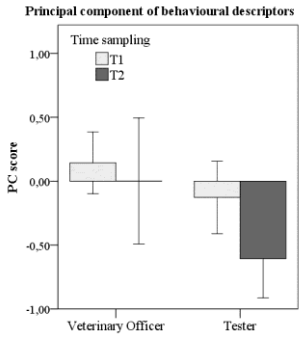


Figure 3. Rate of behaviours included in the Main behaviours class according to sampling time (T1=within three days of shelter intake, T2= after 4 weeks of acclimation in the kennel). #P≤0.1, *P<0.05. The variables "*Jumping to the tester*" and "*Smelling the tester*" are not showed because in percentages lower than 1% for both the first and second sampling time.