Using a Delphi method to estimate the relevance of indicators for the assessment of shelter dog welfare

GV Berteselli*[†‡], S Messori[‡], L Arena[‡§], L Smith[#], P Dalla Villa[‡] and F de Massis[‡]

[†] Department of Veterinary Medicine and Animal Science, Università degli Studi di Milano, Via Festa del Perdono 7, 20122 Milano, Italy
[‡] Istituto Zooprofilattico Sperimentale di Abruzzo e Molise 'G Caporale', Campo Boario, 64100 Teramo, Italy
[§] Facoltà di Medicina Veterinaria, Università di Teramo, SP18, 64100 Teramo, Italy
[#] Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK
* Contact for correspondence: g.berteselli@unimi.it

Abstract

The European regulatory framework lacks standardisation as regards the minimum requirements for shelter facilities, making defining welfare standards for dogs challenging. Dog (Canis familiaris) welfare assessments should consist of a comprehensive set of measurements that allow the calculation of an overall 'welfare score.' The Shelter Quality protocol was developed for the purpose of assessing shelter dog welfare. The study aims to establish a standardised system for evaluating shelter dog welfare by obtaining agreement from experts on the weighting of different measures contributing to an overall welfare score. The Delphi technique is a widely used method for establishing consensus among experts. Two Delphi procedures were implemented and we compared their effectiveness in achieving expert consensus by evaluating rounds' numbers required to reach consensus and the response and attrition rates. Expert consensus was achieved in Delphi 1 when the standard deviation in the expert weightings was ≤ 5. This was achieved easily for the welfare score weightings of the four principles: 'Good feeding', 'Good housing', 'Good Health', and 'Appropriate behaviour.' Animal-based measures were found to reach consensus more quickly than resource-based measures. In Delphi 2, we used the coefficient of variation to determine consensus. No statistical differences were found between the two Delphi methods for attrition rate, response rate or number of participants. Continuing rounds until a consensus is reached is recommended as this method balances time and participant fatigue. A standardised scoring system is provided, using a single overall score of welfare that can be used to compare welfare standards between shelters.

Keywords: *animal welfare, companion animal, Delphi methodology, expert consensus, shelter dogs, Shelter Quality protocol*

Introduction

According to Italian law (14/08/1991, n 281), free-roaming dogs (ie, unrestricted, unsupervised dogs) must be captured and placed in the local public sanitary shelter and remain so for a minimum, ten-day quarantine period to reduce risk of disease transmission. After this time, if they are not returned to their owner or adopted, dogs are transferred to long-term shelters where they will remain until their adoption or death. In Italy, long-term shelters are managed by the municipality government, private groups or non-government organisations. Italian national legal framework does not specify minimum standards for the management and keeping of dogs. Instead, these standards are under regional control, leading to high variability in management approaches across the country (Barnard *et al* 2016; Arena *et al* 2019a).

Shelter dogs (*Canis familiaris*) are at risk of poor social, environmental and management conditions that may lead to negative health and welfare states. These conditions include social deprivation, physical restriction, overcrowding, non-stimulating environments, and separation from an attach-ment figure (aversive stimuli). Such conditions may contribute to the onset of behavioural problems including anxiety and fear-related disorders, stress-related aggression, and repetitive behaviours. Shelter dogs may therefore be considered at high risk of stress, which can compromise dog health and welfare (Beerda *et al* 2000; Hennessy *et al* 2001; Wells *et al* 2002; Hiby *et al* 2006; Taylor & Mills 2007; Dalla Villa *et al* 2013; Protopopova 2016).

Stress can be defined as an individual's physiological and behavioural response to perceived uncontrollability and/or unpredictability in the environment. In stressful circumstances, an individual is unable to respond suffi-ciently to meet the environmental demand (Notari 2009; Koolhaas *et al* 2011). The extent to which a stressor impacts an individual is determined by their coping ability and their capacity to recover quickly both physi-cally and socially (ie their resilience) (Colditz & Hine 2016). Coping abilities are defined as behavioural and physiological efforts to respond to the stressor (Moberg 2000; Broom 2001; Notari 2009).

Behavioural responses represent efficient mechanisms for an individual to respond to a stressful situation as they can aid in the return to homeostasis. Behavioural responses can be due to the rapid activation of the autonomic nervous system or, alternatively, may be in response to chronic stress (Moberg 2000; Hennessy *et al* 2013). Behavioural stress responses can result in reduced behavioural variability (eg decreased activity, exploration and locomotor behaviours) and an increase in the performance of stereotypic and stress-related behaviours, such as excessive auto-grooming, excessive barking, aggressiveness, crying, panting, trembling, and changes to feeding (Mariti *et al* 2012; Part *et al* 2014). Animal behaviour is strictly related to the quality of the environment and to external stimuli. In order to mitigate the detrimental effects of long-term sheltering, an individual's physical and behavioural health and welfare must be prioritised, regardless of length of stay. The shelter environment must allow opportunities for dogs to display species-specific behaviours. Access to protected indoor-outdoor areas are ideal for dogs, particularly when animals are held long term. Outdoor spaces must be suitably enclosed to protect dogs from adverse weather and prohibit escape. Appropriate management, resources (eg food, water, space and shelter), training activities, socialisation with humans and conspecifics, access to outdoor areas, and exercise can reduce potential stress caused by the shelter environment, and improve the dogs' quality of life and welfare. Approaches that are tailored towards individual dogs are desirable (Coppola *et al* 2006; Taylor & Mills 2007; Newbury *et al* 2010; Braun 2011; Miller & Zawistowski 2015).

In the last few decades, researchers have increasingly focused on animal-based measures of assessing animal welfare (Whay *et al* 2003a). Animal-based measures (eg outcome-measures) have clear advantages in comparison to management- and resource-based measures (eg input-based measures). Animal-based measures allow direct, real-time assessment of an individual's welfare state (European Food Safety Authority [EFSA] 2012; Barnard *et al* 2016), and are therefore particularly useful in assessing shelter dog welfare, as they allow researchers and shelter staff to directly and immediately evaluate the dogs' reaction to the shelter environment (Hiby *et al* 2006; Kiddie & Collins 2014; Barnard *et al* 2016).

The Shelter Quality protocol was developed to standardise the assessment of shelter dog welfare using animal- rather than input-based measures (Barnard *et al* 2016; Berteselli *et al* 2019; Arena *et al* 2019a). The Shelter Quality protocol is based on the Welfare Quality®: principles, criteria and measures (Welfare Quality® 2009). Each of the four Welfare Quality® principles (good feeding, good housing, good health and appropriate behaviour) comprise independent but complementary welfare criteria derived from scientifically evidenced measures of welfare. Previous studies have determined the Shelter Quality protocol's validity, feasibility and repeatability (Barnard *et al* 2016; Berteselli *et al* 2019).

The welfare assessment proposed by the Shelter Quality protocol is characterised by three levels of assessment: (i) the shelter level, including management-based measures only (eg feeding regime, exercise routine); (ii) pen level, including both resource- (eg space allowance, bedding) and animal-based measures (eg behaviour, diarrhoea, emotional state); and (iii) individual level, including animal-based measures only (eg health parameters and human-animal relationship) (Table S1). Although the Shelter Quality protocol has proven to be a valid tool (Barnard *et al* 2016; Berteselli *et al* 2019), the assessment fails to provide an overall welfare score. The development of an overall scoring system is therefore required in order to provide a comprehensive welfare assessment tool.

An overall scoring system can be developed by calculating the total score of the welfare measures, weighted by the importance of each measure in calculating dog welfare. The Delphi technique can assist in obtaining expert consensus in the appropriate weighting of each of the welfare measures. The Delphi technique is an interactive process designed to achieve consensus in a group of experts on specific topics (Linstone & Turoff 1975; Hasson *et al* 2000). The Delphi technique consists of a number of interactions, or rounds, during which the administrator (ie the manager of the process) provides participants with statistical summaries of the answers given by the panel of experts in previous rounds. The panel of experts provide responses in sequential rounds. At the beginning of each round, the experts are provided with feedback comparing their previous response to the overall summary of responses provided in the previous round. Generally, the process continues until a consensus is achieved among experts (Linstone & Turoff 1975; Hasson *et al* 2000; Hsu & Sandford 2007; Yousuf 2007). The Delphi technique has been applied to many scientific fields (eg social science, healthcare and veterinary research), including animal welfare (Whay *et al* 2003b; Bennett *et al* 2004; Webster *et al* 2004). The main steps of the Delphi technique are summarised in Table 2.

There is currently no recognised standard procedure for the use of the Delphi technique, leading to variation in methods between studies. We apply two methods to determine an appropriate threshold for expert consensus. For example, studies vary by either continuing the rounds until a consensus is reached or using a pre-determined number of rounds. Studies may also vary in whether they use standard deviation (± SD) vs coefficient of variation (CV) to determine the threshold of consensus (Hasson *et al* 2000; Powell 2003; Hsu & Sandford 2007; Diamond *et al* 2014).

The aim of this study was to obtain agreement from experts in the field on the weighting of different welfare measures in the Shelter Quality protocol in relation to the overall welfare score. This allows the development of a standardised system of evaluating shelter dog welfare and evaluate shelters overall. We use two variations of the Delphi technique, Delphi 1 and 2, and compare these methods in terms of the number of respondents, number of rounds, response rate (RR), attrition rate (AR: rate of participants withdrawing) to determine the most effective and efficient method in achieving expert consensus.

Table 2   Steps of the Delphi procedure (Listone & Turoff 1975; Hsu & Sandford 2007; Yusuf 2007).

Delphi procedure	
1	Construction of the questionnaire
2	Formation of the expert panel
3	Submission of the first questionnaire to the panel
4	Calculation of the average score of responses
5	Submission of second questionnaire in which experts are asked to provide a second assessment, whilst taking into consideration the average score of responses and their answers to motivate their position
6	Calculation of new expert average score and inclusion into the questionnaire of reasoning and comments
7	Iteration of steps 3–6 until the desired degree of consensus is achieved

Materials and methods

Delphi procedure

Two Delphi techniques were implemented, Delphi 1 and 2, consisting of four and three rounds of structured questionnaires, respectively. In Delphi 1, we continued the rounds until a consensus was reached. In Delphi 2 we used a predetermined number of rounds and compared the (i) SD vs (ii) CV to determine a threshold for consensus. We implemented Delphi 2 after Delphi 1 concluded.

Traditionally, the questionnaire in the first round of the Delphi method consists only of open questions but, in this study, we also included closed questions (ie the modified Delphi) (Hasson *et al* 2000; Keeney *et al* 2006; Hsu & Sandford 2007). The use of the modified Delphi procedure is appropriate when the "basic information concerning the target issue is available and usable" (Kerlinger 1973). In this case, the Shelter Quality protocol measures and procedures for assessing shelter dog welfare were based on previous expert opinions, an extensive literature search, and in-field validation (Barnard *et al* 2016; Berteselli *et al* 2019). The Delphi surveys were distributed and respondent data managed using an online survey system (Surveymonkey®, SurveyMonkey Inc, San Mateo, CA, USA). The experts received an initial email inviting them to participate in the Delphi questionnaire. We allowed an approximately three-week time-limit for responses in each round. One week prior to the deadline, email reminders were sent to experts who had not responded. Delphi 1 was carried out between May 2016 and October 2016; and Delphi 2 between July 2017 and September 2017.

Delphi 1

The Delphi 1 procedure aimed to provide an overall animal welfare score for assessing shelter dog welfare by weighting the relative importance of the principles, criteria and measures of the Shelter Quality protocol. In the first round, experts were sent: (i) the structured questionnaire; (ii) a description of the Shelter Quality protocol; (iii) a description of the Delphi technique; and (iv) the aim of the present study. The experts were asked to weight each shelter dog welfare measure and the corresponding welfare criteria for the principles: 'Good feeding', 'Good housing', 'Good health', and 'Appropriate behaviour.' These principles specifically relate to long-term shelter dog welfare. The experts were asked to use a 0 to 100 scale, taking into account that the sum for all measures, criteria and principles must equal 100. At this stage of the Delphi procedure, the experts were not able to justify their scoring or provide comments.

In the following rounds (two, three and four), each expert who participated in the previous round received a second questionnaire including their answers alongside the average weight for the group (mean [± SD]). This allowed experts to compare their responses to the group average. In order to help the experts reconsider their choice and assist in achieving a group consensus, each expert was provided with details of whether their weighting was lower or higher than the group average. Experts were then asked whether they would like to change or confirm their previous answers. The experts were prompted to justify their decision, particularly if their scoring was not in agreement with the group average.

Any sets of welfare measures that achieved an acceptable level of agreement (SD ≤ 5) during the first round, were not included in the following rounds. In the absence of a gold standard in determining the consensus, we decided *a priori* that, in Delphi 1, an acceptable level of agreement was reached when the SD among weightings was equal to or less than 5 for all sets of measures used in the Shelter Quality Protocol welfare assessment (Sharma *et al* 2003). Since the mean represents the group opinion and the SD represents the level of agreement (Greatorex & Dexter 2000), a decrease in SD between rounds indicated an increased level of agreement (Jairath & Weinstein 1994; Yousuf 2007). The rounds continued until the consensus was achieved.

Delphi 2

The Shelter Quality Protocol welfare measures included various indicators of welfare. Categorical variables were used to define each of these welfare measures. For example, body condition could be evaluated as: adequate, too thin, and too heavy. In order to obtain an overall welfare score, we defined the weight of each Shelter Quality Protocol welfare

measure in relation to long-term shelter dog welfare. When there were two or more categorical welfare measures, and when scientific knowledge on the issue was inadequate to attribute an unequivocally correct scoring, we resorted to expert opinion using the second Delphi procedure.

In Delphi 2, the experts were asked to assign a weight to each of the categorical variables using a 0 to 100 scale. The scale varied according to the specific welfare measure under consideration (eg for body condition, 0 corresponds to the worst and 100 the best; for reaction toward humans, 0 corresponds to unacceptable and 100 to acceptable). Any value between 0 and 100 was permitted and the values did not need to sum to an overall score of 100, as each variable was considered independently. Three rounds were performed in Delphi 2 (determined *a priori*).

In Delphi 2, a consensus was determined using the CV. The authors decided to use this method in light of the Delphi 1 results, to allow a consensus to be achieved and to assess factors of the Delphi method (eg number of respondents, RR, AR and dropping out of the procedure). CV is a standardised measure of dispersion and is useful for comparing distributions. In terms of the Delphi method, CV allows a direct comparison of statements from succeeding rounds. CV was calculated in this study, using Equation 1 (which shows the calculation of coefficient of variation) as a measure of consensus. We followed the procedure for interpretation of CV using methods described by English and Kernan (1976), where: $0 \leq V \leq 0.5$ = good degree of consensus, no need for additional round; $0.5 \leq V \leq 0.8$ = less than satisfactory degree of consensus, possible need for additional round; $V > 0.8$ = poor degree of consensus, definite need for additional round.

$$\text{Coefficient of variation} = \frac{\text{Standard deviation}}{\text{Mean}}$$

Data were analysed for normality. The student's *t*-test was used to analyse the difference between Delphi 1 and Delphi 2 in number of participants, RR and AR. We determined a statistically significant difference if *P*-values < 0.05. Statistical analyses were carried out using R V.2.15.3.

The procedure for submitting the questionnaires was the same as in Delphi 1.

Expert panel

The experts were selected on the basis of authorship of relevant papers, experience and skill in animal welfare, canine ethology and behaviour, animal health, shelter medicine and shelter management. The panel included experts from a range of countries, including practitioners, specialists in veterinary behavioural medicine, specialists in shelter medicine, biologists, ethologists and experts in the field of animal welfare. One hundred and ninety-five international experts were invited to take part in Delphi 1 and 259 in Delphi 2. We increased the number of invited experts in Delphi 2 to ensure sufficient number of participants completed the full Delphi procedure.

Ethical statement

All participants were informed clearly about Delphi's procedure and its aim via a letter of invitation and no sensitive data were collected. Participation was voluntary and anonymous with participants able to leave the Delphi procedure at any time.

Results

Results of Delphi 1

In the first round, 60 experts completed the questionnaire. During this round, no set of measures (welfare principles, criteria and measures) reached agreement. In the second round, 40 of the 60 experts (who completed the first round) responded (RR: 67% and AR: –33%). All welfare principles and the criterion, 'Absence of prolonged hunger' and the criterion, 'Absence of prolonged thirst' reached agreement within the established threshold (SD ≤ 5). In the third round, 36 of the 40 experts (who completed round 2) responded (RR: 90% and AR: –10%) and the criteria 'Comfort around resting', 'Thermal comfort' and 'Ease of movement' reached agreement. The remaining measures were presented in a fourth round. In this round, 21 of the 36 experts (that completed round 3) responded (RR: 58% and AR: –42%), and all reached consensus (Table S3).

Results of Delphi 2

In Delphi 2, 54 experts took part in the first round and completed the questionnaire. In the second round, 37 experts responded (RR 68% and AR: –32%). In the third round, 21 experts responded (RR 58% and AR: –42%). All variables achieved consensus at the last round, except 'No or not regular access to outdoor fenced area' (CV 1.32), and 'No or not regular walking on leash' (CV 1.39) (categorical variables of the welfare measure, 'Exercise'). 'Only fear' (one of the possible categorical variables for the welfare measure 'Reaction toward people') reached a CV of 0.72, but this was considered a sufficient consensus by the authors (English & Keran 1976). The results of Delphi 2 are summarised in Table S4.

Tables S1, S3 and S4 can be found in Supplementary Material online.

The characteristics of experts who participated to Delphi 1 and Delphi 2, and the main areas of expertise are reported in Tables 5 and 6, respectively.

No statistically significant differences between the two Delphi procedures were found using the Student's *t*-test (AR: *P* = 0.5201, *t* = 0.6914; RR: *P* = 0.5110, *t* = 0.7072; n of participants; *P* = 0.8833, *t* = 0.1545).

Discussion

Delphi procedure

The interactive Delphi technique involves assembling a panel of experts, creating a questionnaire, synthesising feedback, and guiding a group of experts through a variety of values and judgements. This technique allows accurate

Table 5   Summary of experts (%), including associated institute, gender and nationality.

Characteristics		Distribution (%) Delphi 1				Distribution (%) Delphi 2		
		Round 1 (n = 60)	Round 2 (n = 40)	Round 3 (n = 36)	Round 4 (n = 21)	Round 1 (n = 54)	Round 2 (n = 37)	Round 3 (n = 21)
Associated institute	Research centres	58	58	54	48	57	57	52
	Governmental institution/policy advising	14	20	23	24	17	16	24
	Private consulting	25	22	23	28	23	24	19
	NGO	3	0	0	0	3	3	5
Gender	Male	29	22	25	19	28	14	14
	Female	71	78	75	81	72	86	86
Nationality	Europe	85	88	89	95	93	92	100
	Extra-Europe	15	12	11	5	7	8	0

Table 6   Sample characteristics. Main (M) and additional (A) expertise of the experts participating in the Delphi. Each expert has one main expertise and could have one or more additional areas of expertise.

Expertise	Delphi 1								Delphi 2					
	Round 1 (n = 60)		Round 2 (n = 40)		Round 3 (n = 36)		Round 4 (n = 21)		Round 1 (n = 54)		Round 2 (n = 37)		Round 3 (n = 21)	
	M	A	M	A	M	A	M	A	M	A	M	A	M	A
Animal welfare	17	15	11	10	10	10	6	8	19	14	17	14	11	7
Ethology	12	10	9	7	7	6	4	4	9	12	15	6	8	4
Behavioural medicine	19	3	14	3	13	2	10	1	16	2	2	4	2	2
Animal health and care	3	9	1	6	1	6	0	3	3	10	3	1	0	1
Shelter medicine	9	4	5	3	5	3	1	1	7	2	0	1	0	1

decision-making, as multiple opinions are incorporated into the final consensus (Powell 2003).

To our knowledge, the present study is the first to use the Delphi approach to weight shelter dog welfare measures, allowing the development of a standardised shelter dog welfare scoring system. The Delphi technique is a useful tool that allows us to obtain consensus on multiple, inter-acting, measures of welfare. There is potential for researchers in the wider field of animal welfare to gain consensus on other complex topics, where current literature is also lacking, by applying the Delphi technique (Hasson *et al* 2000; Whay *et al* 2003b; Buckland *et al* 2014).

The involvement of experts from multiple nationalities and expertise allows a more complete overview of the subject of interest (Walker & Selfe 1996; Buckland *et al* 2014). This variety of expertise, coupled with the anonymity of partici-pants, allowed more representative results by preventing group dynamics effects (ie 'bandwagon effect') and leader influence, which can occur during in-person group discus-sions. In particular, anonymity permits a greater willingness to express or change opinion during the Delphi process (Listone & Turoff 1975; Yousuf 2007). This approach also allowed us to moderate individual opinions by comparing them against experts in different disciplines (Buckland *et al* 2014). The validity of results from the Delphi method can be determined by an increase in response rate (RR) and decrease of attrition rate (AR) during the rounds (Evans 1997). There are no guidelines in the literature suggesting appropriate reference values for these factors. Mullen (2003) considers an RR less than 8% unacceptable, while for Walker and Selfe (1996) a 70% minimum response rate should be achieved to maintain rigour, although they offer little support for this claim. In other studies, AR between 20 to 25% and 45 to 50% are considered acceptable (Donohoe & Needham 2009). To reduce response bias due to high attrition, the minimum recommended Delphi RR for each

round is 40–50% (Atkinson & Gold 2001). According to these recommendations, all Delphi rounds described in this study provided valid results.

A decline in responses may relate to the total number of rounds, the length of commitment required (eg weeks or months of interaction), distraction between rounds or disillusions with the process (Donohoe & Needham 2009). The size of the Delphi panel may also influence the RR and AR, but there is no defined minimum number of experts required (Keeney *et al* 2006). The panels can vary from a dozen to a few thousand participants, without justification of the sample size. Smaller panels have lower reported AR (Williams & Webb 1994). To mitigate a decline in responses in the present study, the administrator encouraged the experts to take part and complete all rounds by highlighting the importance of their contribution and knowledge in creating a standardised scoring system for the assessment of dogs housed in long-term shelters. Reminders were sent via email by the online survey system motivating each expert to complete the rounds. The addition of new members to replace withdrawals was avoided in this study, as the inclusion of new individuals may negatively affect the Delphi procedure and the reliability of results (Donohoe & Needham 2009).

The two Delphi techniques used (defined vs no defined number of rounds determined *a priori;* SD vs CV for achieving the consensus) had similar results, there were no significant differences in RR, AR, number of respondents and number of abandonments. In the Delphi 2 (characterised by three rounds defined *a priori* and by CV as a method for achieving consensus), two categorical variables did not reach a consensus, possibly relating to the smaller number of rounds. Although, theoretically, three rounds are sufficient to collect the required information and reach a consensus, this was not the case in our study (Powell 2003; Hsu & Sandford 2007).

Welfare principles

During the Delphi procedure, the panel of experts participated in a lively discussion of the topics, allowing a productive debate that assisted in the experts reaching consensus on the weightings of the welfare measures. As expected, in both Delphi 1 and Delphi 2, no set of measures reached a consensus in the first round (Listone & Turoff 1975). Agreement was first achieved in the second round of Delphi 1 on the criteria, 'Absence of prolonged hunger' and 'Absence of prolonged thirst' (welfare principle of 'Good feeding'). The four principles ('Good feeding', 'Good housing', 'Good health', 'Appropriate behaviour') are considered key points in animal welfare assessment. The experts allocated weight among the principles almost equally.

Principle of 'Good feeding'

Within the principle of 'Good feeding', the experts agreed that more weight should be given to absence of prolonged thirst than absence of prolonged hunger. Animals require access to sufficient food and water to maintain normal health and prevent prolonged hunger, thirst, malnutrition or dehydration (Fraser *et al* 2013). The scientific community

agrees that the provision of adequate food and water is fundamental in providing good care to animals (Miele *et al* 2011), a concept that is also agreed by citizens and farmers (Te Velde *et al* 2002). The concept of water and food as a primary animal need is clearly reaffirmed in the first of the Five Freedoms ('Freedom from hunger and thirst') (Farm Animal Welfare Council [FAWC] 1992; Bracke *et al* 1999; Capdeville & Veissier 2001).

Within the welfare criteria, 'Absence of prolonged thirst', consensus on the weighting of 'Availability of water' was achieved after a lively discussion among the experts. Some of the experts gave a high weighting to this measure because, in their opinion, water is a fundamental resource that is essential for life. Although the experts agreed on the importance of water availability, some considered it just as important as the other welfare criteria, 'Absence of prolonged thirst', which included the welfare sub-measures, 'Type of drinkers', 'Cleanliness of water' and 'Safety of drinkers.' Based on the experts' experience, unsafe drinkers and dirty water is a risk to dog health and welfare. The experts only reached agreement in the last round of first Delphi on the increased weighting of the welfare measure, 'Availability of water' in long-term shelter dog welfare. These findings are in line with those of Vanhonacker *et al* (2008), who demonstrated that, for both citizens and farmers, the availability of water was perceived as an important item required for acceptable levels of farm animal welfare.

In Delphi 2, the welfare measure 'Feeding' (criterion 'Absence of prolonged hunger'), the experts determined that feeding dogs twice per day was the most appropriate option for shelter dog welfare, as feeding twice per day prevents gastroenteric disorders (eg diarrhoea, gastric dilatation-volvulus) (Glickman *et al* 1997) and best controls food intake. Additionally, feeding dogs twice a day increases social contact between the dogs and people.

Principle of 'Good housing'

There was a difference in opinion on the weightings of measures within the criterion 'Comfort around resting', particularly for the welfare sub-measure, 'Type of bedding.' Some experts considered type of bedding irrelevant for animal welfare, assigning it a low weight (eg less than 5 on the 0–100 scale) and allocating greater weighting to the 'Presence of at least one bedding per dog' and 'Cleanliness/dryness of bedding.' Many experts commented that dogs have individual needs in relation to bedding type and, therefore, no single type of bedding should rate more highly in terms of animal welfare. Experts instead highlighted the importance of providing an adequate number of beds to avoid competition between dogs, agonistic reactions and stress that can promote the development of behavioural problems (Taylor & Mills 2007). Other experts simply divided the weight equally among the different measures composing the 'Comfort around resting' criterion. The availability of literature on this topic is scarce. As shelter dogs are often inactive and spend large periods of time resting or sleeping in the pen, the type of bedding should be considered in terms of their welfare needs (Normando *et al*

2009). Previous studies have reported shelter dogs' preference for fabric beds compared to plastic baskets, suggesting that polyethylene bedding is an optimal choice for dog welfare, as it provides safe and comfortable bedding (Normando *et al* 2014). Polyethylene bedding is recommended for geriatric laboratory beagles, puppies, and kennelled dogs, as the polyethylene is durable, waterproof, easily sanitised, warm and comfortable (Eisele 2001). Geriatric dogs comprise a large proportion of the shelter population in Italy (Arena *et al* 2019a), and are prone to displaying clinical problems, such as articular disorders (eg degenerative joint disease, lameness, arthritis), and skin alterations (eg loss of skin elasticity, wrinkling of skin, thickening of foot-pad, formation of callus) (Pati *et al* 2015; Arena *et al* 2019a). To improve shelter dog welfare, the choice of bedding materials should consider the age of the dogs and their physical conditions (Eisele 2001; Taylor & Mills 2007; Arena *et al* 2019a).

The results of Delphi 2 highlighted the importance of kennels as a bedding type for sheltered dogs. Kennels were considered the best option compared to baskets and other bedding materials (eg platforms and blankets). Although the use of kennels as a bedding type in shelters can make dogs feel secure and can be used in play behaviour (running around and jumping on top) and environmental enrichment (Hubrecht *et al* 1992), the real impact of kennels on dog welfare is currently unknown (Taylor & Mills 2007).

It may be beneficial to provide different types of bedding per pen to allow sheltered dogs the possibility of choice. In the bedding option, as with other environmental enrichments, individual preferences should be considered (Eisele 2001).

In the criterion, 'Thermal comfort', the welfare measure, 'Thermoregulation' received the highest weight. Most participants agreed that this measure provides a direct indication of the animal's thermal comfort (eg excessive panting in cases of warmth or shivering/huddling in cases of cold). This welfare measure is an animal-based measure and should be weighted more highly compared to the resource-based welfare measures within this criterion, such as 'Shelter from adverse weather conditions.' Experts emphasised that this measure requires knowledge of canine behaviour to discriminate between panting, shivering, and huddling as signs of thermal discomfort or as signs of fear or anxiety disorders. Individual anatomic characteristics, such as thick fur or long-haired coats, can make it difficult to detect behavioural signs of thermal discomfort, such as shivering, therefore possibly impairing the reliability of this measure.

Principle of 'Good health'

Considering the principle, 'Good health', experts weighted welfare sub-measures of the criterion, 'Absence of injuries' fairly equally. This may relate to all welfare sub-measures in this criterion being animal-based. Experts considered the 'Presence of wounds' more reliable than other sub-measures, as the presence of wounds is directly linked to management and the consequences of wounds (eg pain and possible infections) directly impact dog welfare (Cockram & Hughes 2011; Walsh 2016).

The criterion, 'Absence of disease' is also exclusively composed of animal-based measures ('Evidence of pain'; 'Signs of diarrhoea'; 'Coughing'). The expert panel agreed that 'Evidence of pain' should be weighted most highly. Pain can be caused by several pathological conditions and can manifest into behavioural changes (Overall 2013; Mills *et al* 2020). The welfare measure, 'Evidence of pain' reliably indicates poor health conditions and, as a result, highlights concerns regarding management of a shelter. Although many studies have tried to develop tools that recognise and assess pain in animals (Calvo *et al* 2014; Merola & Mills 2016a), accurate detection of animal suffering remains a challenge. In literature, the definition of acute and chronic pain is not univocal (Mellor *et al* 2000). Acute pain is usually defined as a temporary feeling that has evolved as a mechanism to avoid harmful stimuli and stimuli that can hinder the healing of damaged tissue. Chronic pain is longer lasting and can exacerbate tissue damage and negatively impact healing (Millan 1999; Meintjes 2012; Merola & Mills 2016b; Walsh 2016). This distinction between acute and chronic pain based on the duration is often arbitrary, with transitions between acute and chronic terminology occurring, in some instances, after a few days of pain, and in others, after several weeks. Definitions of pain based on duration do not identify the underlying casual agents. Instead, terminology such as 'adaptative' for acute pain and 'maladaptive' for chronic pain more accurately depict the individuals' capability to avoid noxious stimuli (Lascelles & Robertson 2010; Walsh 2016).

Within the criterion, 'Absence of pain induced by management procedures' the measure, 'Presence of operating procedures for post-surgical pain monitoring' received a high weighting compared to the other welfare measures of this criterion ('Presence of protocol for analgesia' and 'Presence of hospital pens'). This highlights that the experts consider pain control an important factor in ensuring high levels of shelter dog welfare.

Principle of 'Appropriate behaviour'

Although the principles were weighted almost equally, it is interesting that experts considered the 'Principle of appropriate behaviour' as most relevant to overall welfare. This is possibly due to dog behavioural responses being the first visible reaction to their living environment, providing important information on the capability of animals to cope with stimuli (Bracke & Hopster 2006). As behavioural assessments are non-invasive and non-intrusive, they are commonly used to measure animal welfare. Behaviour is the result of the animals' decision-making process (Dawkins 2004). When the ability to perform natural behaviours is restricted, stress-related behaviours can occur, such as increased aggression, auto-mutilations (eg acral lick dermatitis) and stereotyped behaviours (eg pacing, tail-chasing and circling) (Beerda *et al* 2000; Fraser *et al* 2013). Undesirable dog behaviours that do not meet the owners' expectations of appropriate behaviours can result in dogs being categorised as having behavioural problems. Dog behavioural problems are associated with a high risk of relinquishment or adoption failure (Shore 2005; Diesel *et al* 2010).

The experts emphasised the importance of using both canine behaviour and animal-based measures in the assessment of shelter dog welfare. Even if we assume that animal-based measures allow us to obtain direct information on the status of the animal (EFSA 2012), this also could be due to the presence of functionally analogous forms of human behaviour traits in dogs, possibly due to the co-evolution of dogs and humans (Topál *et al* 2009). Humans and dogs share similar mechanisms of communication (eg low fundamental frequency in the vocalisation indicates agonistic tendencies) and this aspect facilitates inter-species communication and human understanding of canine behaviours and vice versa (Hare & Tomasello 2005). Dogs have a close relationship with humans that differs from those between people and other domesticated animals. The human-dog bond is influenced both by canine and human characteristics (Zilcha-Mano *et al* 2011; McGreevy *et al* 2012; Payne *et al* 2015). Dogs seem to possess a great ability to interpret human signalling probably due to selective pressure on dogs to work with humans and to their pre-adaptation to domestication due to the social nature of the ancestors (Marshall-Pescini & Kaminski 2014).

'Presence of abnormal behaviours' (active-repetitive behaviours, eg pacing, circling, spinning, bouncing; other compulsive behaviours, eg auto-mutilation, licking/chewing furniture or bars) is also an animal-based measure and was considered the most relevant measure for assessing the criterion, 'Expression of other behaviours.' Abnormal behaviours can be considered as poor welfare indicators and can be related to stress and frustrating conditions (eg due to confinement in restricted and non-stimulating environments) (Hubrecht *et al* 1992; Hiby *et al* 2006; Mason & Rushen 2008). High incidences of abnormal behaviour are often associated with other poor welfare indicators (Mason & Rushen 2008). The performance of stereotypical behaviours in some cases may help animals to cope with their environment (Koolhaas *et al* 1999; Mason & Latham 2004; Pomerantz *et al* 2012; Denham *et al* 2014). Although animals displaying repetitive behaviours can show fewer symptoms of poor welfare (ie inadequate body condition or presence of diseases) than dogs without these behavioural manifestations, the presence of abnormal behaviour is still a valid welfare measure as it acts as a warning sign of distress and suffering. Abnormal behaviour is important but should not be used in isolation as an indication of welfare (Mason & Rushen 2008). Dogs are social animals, and it is important to promote the development of their social skills through appropriate management and activities. For this reason, the experts of Delphi 2 considered paired and group-housing pens as better housing options compared to single housing (Grigg *et al* 2017). Keeping dogs in pairs or groups provides a stimulating environment and encourages social contact, locomotion and exploratory behaviours (Petak 2013). Isolation of dogs is only considered acceptable for medical and safety reasons (Beerda *et al* 2000). Experts participating in Delphi 2 determined that pens with more than five dogs should be avoided, since a large number of dogs in one pen can be a stressor due to the reduction in space per dog and the lack of ability to avoid undesirable pen-mate interaction. Moreover, from a health and behaviour perspective, it is difficult to correctly manage a group of more than five dogs. Dogs should be grouped according to their size, sociality, temperament, age, sex and physiological status (Weipkema & Schouten 1990; Miller & Zawistowski 2015). Although Mertens and Unshlem (1996) did not see aggressive events among group-housed pen-mates, quarrels and agonistic encounters can occur in group-housing. Shelter staff should monitor dogs kept in groups to prevent negative consequences, such as aggression and biting events, or fear stress, and anxiety in some members of the group (Newbury *et al* 2010).

Dogs should also be provided with regular access to outdoor fenced areas and walks on the leash to ensure visual, auditory and olfactory stimulation, as well as social contact with people and conspecifics (Taylor & Mills 2007). Experts in Delphi 2 commented that regular walks and outdoor exercise reduces stress attributed to the shelter environment, therefore preventing outbreaks of behavioural problems (eg intra- and inter-specific aggression, anxiety-related disorders, fear). Regular walks and exercise can therefore increase shelter dog welfare, as well as increasing their adoptability (Luescher & Medlock 2009; Braun 2011; Menor-Campos *et al* 2011). The experts emphasised that restrictors of these activities mainly include the availability of personnel (eg volunteers or operators) and time to walk and supervise dogs exercising, particularly when the number of shelter dogs are large (Nardoia *et al* 2019). This could explain the lack of expert agreement on the variable, 'No or not regular access to outdoor fenced area and walking on leash.' The wording 'no or not regular' caused debate; some experts highlighted that the welfare impact of a lack of these activities differs depending upon whether the activities are never carried out, or whether the activities are carried out, but infrequently. It may therefore be advisable to split this measure into 'no activity routine' and 'infrequent activity routine.' The predictability of the living environment marked by routine meals, exercise and social activities can help shelter dogs cope with stress and control of the environment (Morgan & Tromborg 2007; Koolhaas *et al* 2011, 2013).

Good relationships with humans (eg shelter operators, volunteers) increases dog welfare and their chances of adoption (Coppola *et al* 2006; Menor-Campos *et al* 2011; Shih *et al* 2021). 'Reaction to people' obtained the highest weighting and was considered by experts the most important welfare measure in the criterion 'Good human-animal relationship.' Dogs with aggressive or fear reactions towards humans are prevalent in shelters, due to their low probability of adoption and high probability of relinquishment by owners (Shore 2005; Diesel *et al* 2010). The experts agreed that specific training programmes and behavioural rehabilitation should be customised towards dogs with aggressive or fear reactions towards humans, in order to improve their social skills, solve problematic behaviours, and increase the possibility of adoption, as supported by previous literature (Luescher & Medlock 2009; Normando *et al* 2009; Menor-Campos 2011; Protopopova *et al* 2018).

Experts highlighted the importance of the human-animal relationship by weighting the welfare criterion, 'Good human-animal relationship' highly. 'Good human-animal relationship' and 'Positive emotional state' received the highest weight amongst the criteria composing the principle, 'Appropriate behaviours.' Animals are sentient beings capable of experiencing sensations and emotions and it is now widely recognised that animal welfare is based not only on a good health status but also on good mental state (Broom 2011). A higher prevalence of positive emotions (eg pleasure, happiness) compared to negative emotions (eg fear, pain) leads to improved animal welfare (Boissy *et al* 2007). The experts participating in the Delphi 2 embraced this view by weighing positive adjectives more highly than negative adjectives in the evaluation of animal emotional state through qualitative behavioural assessments (QBA) (Arena *et al* 2017). The emotional state plays an important role that influences animal behaviour, communication, social bonding and cognitive functioning (Paul *et al* 2005). Currently, welfare research is geared towards the assessment of positive outcomes (eg play, affiliative behaviours, some vocalisations) and animal emotions (Wemelsfelder *et al* 2001; Boissy *et al* 2007; Zupan *et al* 2016; Arena *et al* 2017, 2019b). In light of the expert opinions in this study, and in the wider scientific community, the use of animal-based measures (eg physical condition, behavioural and/or physiological indicators), in addition to resource-based and environmental variables, has been recommended for welfare assessment. These measures assess welfare in a manner that more closely reflects an individual animal's perception of their environment (Eisele 2001; Whay *et al* 2003b; Boissy *et al* 2007; Kiddie & Collins 2014).

Animal welfare implications and conclusion

The lack of universally agreed guidelines on the use of Delphi techniques and a standardisation of methodology allowed us to design and adapt the Delphi technique to the objectives of the study. The Delphi techniques (Delphi 1 and Delphi 2) implemented in this study were both successful in achieving expert consensus on the relative weightings of welfare measures in an overall welfare score for the Shelter Quality Protocol. We suggest that continuing rounds until a consensus is agreed (Delphi 1) is a better Delphi technique than determining the number of rounds prior to implementation (Delphi 2), as this method balanced time and participant fatigue, whilst still achieving expert consensus in a few rounds (four). We found that both methods of determining a consensus threshold (SD in Delphi 1 and CV in Delphi 2) provided valid results.

The high weightings assigned by the experts to animal-based measures confirm that these are the most reliable measures of animal welfare condition. Through the Delphi technique, it was possible to obtain expert opinion on issues lacking scientific research, such as the best bedding or social housing option for shelter dogs. The Delphi technique therefore has the potential to assist in filling gaps or further consolidating scientific understanding of shelter dog welfare. The results of this study could be useful in creating guidelines for the correct management of dogs housed in long-term shelters.

Additionally, the results obtained in this study not only improve the understanding of shelter dog welfare, but also allowed the development of a numerical scoring system that can be integrated into the Shelter Quality protocol, providing an overall welfare score for shelter dogs. The final scoring system will allow shelters to be ranked according to dog welfare standards. Shelters will be able to use this scoring system to highlight focus areas (such as feeding or exercise) for improvement in order to increase their overall welfare score.

Declaration of interest

None.

References

Arena L, Berteselli GV, Lombardo F, Candeloro L, Dalla Villa P and De Massis F 2019a Application of a welfare assessment tool (Shelter Quality Protocol) in 64 Italian long-term dogs' shelters. Welfare hazard analysis. *Animal Welfare 28*: 353-363. https://doi.org/10.7120/09627286.28.3.353

Arena L, Wemelsfelder F, Messori S, Ferri N and Barnard S 2017 Application of Free Choice Profiling to assess the emotional state of dogs housed in shelter environments. *Applied Animal Behaviour Science 195*: 72-79. https://doi.org/10.1016/j.applanim.2017.06.005

Arena L, Wemelsfelder F, Messori S, Ferri N and Barnard S 2019b Development of a fixed list of descriptors for the qualitative behavioural assessment of shelter dogs. *BioRxiv*: 545020. https://doi.org/10.1101/545020

Atkinson NL and Gold RS 2001 Online research to guide knowledge management planning. *Health Education and Research 16(6)*: 747-763. https://doi.org/10.1093/her/16.6.747. https://doi.org/10.1093/her/16.6.747

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. *The Veterinary Record 178(1)*: 18. https://doi.org/10.1136/vr.103336

Beerda B, Schilder MBH, Van Hooff JARAM, De Vries HW and Mol JA 2000 Behavioural and hormonal indicators of enduring environmental stress in dogs. *Animal Welfare 9*: 49-62

Bennett RM, Broom DM, Henson SJ, Blaney SJP and Harper G 2004 Assessment of the impact of government animal welfare policy on farm animal welfare in the UK. *Animal Welfare 13(1)*: 1-11

Berteselli GV, Arena L, Candeloro L, Dalla Villa P and De Massis F 2019 Interobserver agreement and sensitivity to climatic conditions in sheltered dogs' welfare evaluation performed with welfare assessment protocol (Shelter Quality protocol). *Journal of Veterinary Behavior 29*: 45-52. https://doi.org/10.1016/j.jveb.2018.09.003

Boissy A, Manteuffel G, Jensen MB, Moe RO, Spruijt B, Keeling LJ, Winckler C, Forkman B, Dimitrov I, Langbein J, Bakken M, Veissier I and Aubert A 2007 Assessment of positive emotions in animals to improve their welfare. *Physiology & Behavior 92(3)*: 375-397. https://doi.org/10.1016/j.physbeh.2007.02.003

Bracke MB and Hopster H 2006 Assessing the importance of natural behavior for animal welfare. *Journal of Agricultural and Environmental Ethics 19(1)*: 77-89. https://doi.org/10.1007/s10806-005-4493-7

Bracke MB, Spruijt BM and Metz JH 1999 Overall animal welfare reviewed. Part 3: Welfare assessment based on needs and supported by expert opinion. *NJAS Wageningen Journal of Life Sciences 47(3)*: 307-322. https://doi.org/10.18174/njas.v47i3.468

Braun G 2011 Taking a shelter dog for walks as an important step in the resocialization process. *Journal of Veterinary Behavior 6(1)*: 100. https://doi.org/10.1016/j.jveb.2010.08.004

Broom DM 2001 Coping, stress and welfare. In: Broom DM (ed) *Coping with Challenge: Welfare in Animals including Humans* pp 1-9. Dahlem University Press: Berlin, Germany

Broom DM 2011 A history of animal welfare science. *Acta Biotheoretica 59*: 121-137. https://doi.org/10.1007/s10441-011-9123-3

Buckland EL, Corr SA, Abeyesinghe SM and Wathes CM 2014 Prioritisation of companion dog welfare issues using expert consensus. *Animal Welfare 23*: 39-49. https://doi.org/10.7120/09627286.23.1.039

Calvo G, Holden E, Reid J, Scott EM, Firth A, Bell A, Robertson S and Nolan AM 2014 Development of a behaviour-based measurement tool with defined intervention level for assessing acute pain in cats. *Journal of Small Animal Practice 55(12)*: 622-629. https://doi.org/10.1111/jsap.12280

Capdeville J and Veissier I 2001 A method of assessing welfare in loose housed dairy cows at farm level, focusing on animal observations. *Acta Agriculturae Scandinavica, Section A-Animal Science 51(S30)*: 62-68. https://doi.org/10.1080/090647001316923081

Cockram MS and Hughes BO 2011 Health and disease. In: Appleby MC, Hughes BO, Mench JA and Olsson AI (eds) *Animals Welfare, Second Edition* pp 120-137. CABI: Wallingford, UK. https://doi.org/10.1079/9781845936594.0120

Colditz IG and Hine BC 2016 Resilience in farm animals: biology, management, breeding and implications for animal welfare. *Animal Production Science 56(12)*: 1961-1983. https://doi.org/10.1071/AN15297

Coppola CL, Grandin T and Enns RM 2006 Human interaction and cortisol: can human contact reduce stress for shelter dogs? *Physiology & Behavior 87(3)*: 537-541. https://doi.org/10.1016/j.physbeh.2005.12.001

Dalla Villa P, Barnard S, Di Fede E, Podaliri M, Candeloro L, Di Nardo A, Siracura C and Serpell JA 2013 Behavioural and physiological responses of shelter dogs to long-term confinement. *Veterinaria Italiana 49(2)*: 231-241

Dawkins MS 2004 Using behaviour to assess animal welfare. *Animal Welfare 13(1)*: 3-7

Denham HDC, Bradshaw JWS and Rooney NJ 2014 Repetitive behaviour in kenneled domestic dog: stereotypical or not? *Physiology & Behavior 128*: 288-294. https://doi.org/10.1016/j.physbeh.2014.01.007

Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM and Wales PW 2014 Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology 67(4)*: 401-409. https://doi.org/10.1016/j.jclinepi.2013.12.002

Diesel G, Brodbelt D and Pfeiffer DU 2010 Characteristics of relinquished dogs and their owners at 14 rehoming centers in the United Kingdom. *Journal of Applied Animal Welfare Science 13(1)*: 15-30. https://doi.org/10.1080/10888700903369255

Donohoe HM and Needham RD 2009 Moving best practice forward: Delphi characteristics, advantages, potential problems, and solutions. *International Journal of Tourism Research 11(5)*: 415-437. https://doi.org/10.1002/jtr.709

EFSA Panel on Animal Health and Welfare 2012 Statement on the use of animal-based measures to assess the welfare of animals. *EFSA Journal 10(6)*: 2767-2795. https://doi.org/10.2903/j.efsa.2012.2767

Eisele PH 2001 A practical dog bed for environmental enrichment for geriatric beagles with applications for puppies and other small dogs. *Contemporary Topics in Laboratory Animal Science 40*: 36-38

English JM and Kernan GL 1976 The prediction of air travel and aircraft technology to the year 2000 using the Delphi method *Transportation Research 10(1)*: 1-8. https://doi.org/10.1016/0041-1647(76)90094-0

Evans C 1997 The use of consensus methods and expert panels in pharmacoeconomic studies. *Pharmacoeconomics 12*: 121-122. https://doi.org/10.2165/00019053-199712020-00003

Farm Animal Welfare Council 1992 *FAWC updates the Five Freedoms.* Veterinary Record *131*: 357

Fraser D, Duncan IJ, Edwards SA, Grandin T, Gregory NG, Guyonnet V, Hemsworth PH, Huertas SM, Huzzey JM, Mellor DJ, Mench JA, Spinka M and Whay HR 2013 General principles for the welfare of animals in production systems: the underlying science and its application. *The Veterinary Journal 198(1)*: 19-27. https://doi.org/10.1016/j.tvjl.2013.06.028

Glickman LT, Glickman NW, Schellenberg DB, Simpson K and Lantz GC 1997 Multiple risk factors for the gastric dilation-volvulus syndrome in dogs: a practitioner/owner case-control study. *Journal of the American Animal Hospital Association 33(3)*: 197-204. https://doi.org/10.5326/15473317-33-3-197

Greatorex J and Dexter T 2000 An accessible analytical approach for investigating what happens between the rounds of a Delphi study. *Journal of Advanced Nursing 32(4)*: 1016-1024. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01569.x

Grigg EK, Nibblett BM, Robinson JQ and Smits JE 2017 Evaluating pair versus solitary housing in kennelled domestic dogs (*Canis familiaris*) using behaviour and hair cortisol: a pilot study. *Veterinary Record Open 4(1)*: e000193. https://doi.org/10.1136/vetreco-2016-000193

Hare B and Tomasello M 2005 Human-like social skills in dogs? *Trends in Cognitive Sciences 9*: 405-454. https://doi.org/10.1016/j.tics.2005.07.003

Hasson F, Keeney S and McKenna H 2000 Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing 32*: 1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Hennessy MB 2013 Using hypothalamic-pituitary-adrenal measures for assessing and reducing the stress of dogs in shelters: A review. *Applied Animal Behaviour Science 149(1-4)*: 1-12. https://doi.org/10.1016/j.applanim.2013.09.004

Hennessy MB, Voith VL, Mazzei SJ, Buttram J, Miller DD and Linden F 2001 Behavior and cortisol levels of dogs in a public animal shelter, and an exploration of the ability of these measures to predict problem behavior after adoption. *Applied Animal Behaviour Science 73*: 217-233. https://doi.org/10.1016/S0168-1591(01)00139-3

Hiby EF, Rooney NJ and Bradshaw JWS 2006 Behavioural and physiological responses of dogs entering re-homing kennels. *Physiology & Behavior 89*: 385-391. https://doi.org/10.1016/S0168-1591(01)00139-3

Hsu CC and Sandford BA 2007 The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation 12*: 1-8

Hubrecht RC, Serpell JA and Poole TB 1992 Correlates of pen size and housing condition on behaviour of kennelled dogs. *Applied Animal Behaviour Science 34*: 365-383. https://doi.org/10.1016/S0168-1591(05)80096-6

Jairath N and Weinstein J 1994 The Delphi methodology (Part one): A useful administrative approach. *Canadian Journal of Nursing Administration 7(3)*: 29-42

Keeney S, Hasson F and McKenna H 2006 Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing 53(2)*: 205-212. https://doi.org/10.1111/j.1365-2648.2006.03716.x

Kerlinger FN 1973 *Foundations of Behavioral Research*. Holt, Rinehart, Winston, Inc: New York, NY, USA

Kiddie JL and Collins LM 2014 Development and validation of a quality of life assessment tool for use in kennelled dogs (*Canis familiaris*). *Applied Animal Behaviour Science 158*: 57-68. https://doi.org/10.1016/j.applanim.2014.05.008

Koolhaas JM, Bartolomucci A, Buwalda B, de Boer SF, Flügge G, Korte SM, Meerlo P, Murison R, Olivier B, Palanza P, Richter-Levin G, Sgoifo A, Steimer T, Stiedl O, van Dijk G, Wöhr M and Fuchs E 2011 Stress revisited: a critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews 35(5)*: 1291-1301. https://doi.org/10.1016/j.neubiorev.2011.02.003

Koolhaas JM, Coppens CM, de Boer SF, Buwalda B, Meerlo P and Timmermans PJ 2013 The resident-intruder paradigm: a standardized test for aggression, violence and social stress. *Journal of Visualized Experiments 77*: e4367. https://doi.org/10.3791/4367

Koolhaas JM, Korte SM, De Boer SF, Van Der Vegt BJ, Van Reenen CG, Hopster H, De Jong IC, Ruis MAW and Blokhuis HJ 1999 Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews 23*: 925-935. https://doi.org/10.1016/j.neubiorev.2011.02.003

Lascelles BD and Robertson SA 2010 DJD-associated pain in cats: what can we do to promote patient comfort? *Journal of Feline Medicine and Surgery 12*: 200-212. https://doi.org/10.1016/j.jfms.2010.01.003

Linstone HA and Turoff M 1975 Introduction. In: Linstone HA and Turoff M (eds) *The Delphi Method: Techniques and Applications* pp 3-12. Addison-Wesley Publishing Company: Reading, MA, USA

Luescher A and Medlock R 2009 The effects of training and environmental alterations on adoption success of shelter dogs. *Applied Animal Behaviour Science 117*: 63-68. https://doi.org/10.1016/j.applanim.2008.11.001

Mariti C, Gazzano A, Moore JL, Baragli P, Chelli L and Sighieri C 2012 Perception of dogs' stress by their owners. *Journal of Veterinary Behavior: Clinical Applications and Research 7(4)*: 213-219. https://doi.org/10.1016/j.jveb.2011.09.004

Marshall-Pescini S and Kaminski J 2014 The social dog: history and evolution. In: Kaminski J and Marshall-Pescini S (eds) *The Social Dog: Behavior and Cognition* pp 3-33. Academic Press: London, UK. https://doi.org/10.1016/B978-0-12-407818-5.00001-2

Mason G and Rushen J 2008 *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare*. CAB International: Wallingford, UK

Mason GJ and Latham NR 2004 Can't stop, won't stop: Is stereotypy a reliable animal. *Animal Welfare 13*: S57-S69

McGreevy PD, Starling M, Branson NJ, Cobb ML and Calnon D 2012 An overview of the dog-human dyad and ethograms within it. *Journal of Veterinary Behavior 7(2)*: 103-117. https://doi.org/10.1016/j.jveb.2011.06.001

Meintjes R 2012 An overview of the physiology of pain for the veterinarian. *The Veterinary Journal 193*: 344-348. https://doi.org/10.1016/j.tvjl.2012.03.001

Mellor D, Cook C and Stafford K 2000 Quantifying some responses to pain as a stressor. In: Moberg GP and Mench JA (eds) *The Biology of Animal Stress: Basic Principles and Implications for Welfare* pp 171-198. CAB International: Wallingford, UK. https://doi.org/10.1079/9780851993591.0171

Menor-Campos DJ, Molleda-Carbonell JM and López-Rodríguez R 2011 Effects of exercise and human contact on animal welfare in a dog shelter. *Veterinary Record 169(15)*. https://doi.org/10.1136/vr.d4757

Merola I and Mills DS 2016a Behavioural signs of pain cats: an expert consensus. *PLoS ONE 11(2)*. https://doi.org/10.1371/journal.pone.0150040. https://doi.org/10.1371/journal.pone.0150040

Merola I and Mills DS 2016b Systematic review of the behavioural assessment of pain in cats. *Journal of Feline Medicine and Surgery 18(2)*: 60-76. https://doi.org/10.1177/1098612X15578725

Mertens PA and Unshelm J 1996 Effects of group and individual housing on the behavior of kennelled dogs in animal shelters. *Anthrozoös 9(1)*: 40-51. https://doi.org/10.2752/089279396787001662

Miele M, Veissier I, Evans A and Botreau R 2011 Animal welfare: establishing a dialogue between science and society. *Animal Welfare 20(1)*: 103

Millan MJ 1999 The induction of pain: an integrative review. *Progress in Neurobiology 57*: 1-164. https://doi.org/10.1016/S0301-0082(98)00048-3

Miller L and Zawistowski S 2015 Housing, husbandry and behavior of dogs in animal shelters. In: Weiss E, Mohan-Gibbons H and Zawistowsky S (eds) *Animal Behavior for Shelter Veterinarians and Staff* pp 145-159. John Wiley & Sons: London, UK. https://doi.org/10.1002/9781119421313.ch7

Mills DS, Demontigny-Bédard I, Gruen M, Klinck MP, McPeake KJ, Barcelos AM, Hewison L, Van Haevermaet H, Denenberg S, Hauser H, Koch C, Ballantyne K, Wilson C, Mathkari CV, Pounder J, Garcia E, Darder P, Fatjó J and Levine E 2020 Pain and problem behavior in cats and dogs. *Animals 10(2)*: 318. https://doi.org/10.3390/ani10020318

Moberg GP 2000 Biological response to stress: implications for animal welfare. In: Moberg GP and Mench JA (eds) *The Biology of Animal Stress* pp 123-146. CABI Publishing: Wallingford, UK. https://doi.org/10.1079/9780851993591.0001

Morgan KN and Tromborg CT 2007 Sources of stress in captivity. *Applied Animal Behaviour Science 102(3-4)*: 262-302. https://doi.org/10.1016/j.applanim.2006.05.032

Mullen PM 2003 Delphi: myths and reality. *Journal of Health Organisation and Management 17(1)*: 37-52. https://doi.org/10.1108/14777260310469319

Nardoia M, Arena L, Berteselli G, Migliaccio P, Valerii L, Di Giustino L and Dalla Villa P 2019 Development of a questionnaire to evaluate the occupational stress in dog's shelter operators. *Veterinaria Italiana 55(1)*: 5-14. https://doi.org/10.12834/VetIt.1476.7954.2

Newbury S, Blinn MK, Bushby PA, Cox CB, Dinnage JD, Griffin B, Hurley KF, Isaza N, Jones W, Miller L, O'Quin J, Patronek GJ, Smith-Blackmore M and Spindel M 2010 *Guidelines for standards of care in animal shelters.* The Association of Shelter Veterinarians: Corning, NY, USA

Normando S, Contiero B, Marchesini G and Ricci R 2014 Effects of space allowance on the behaviour or long-term housed shelter dogs. *Behavioural Processes 103*: 306-314. https://doi.org/10.1016/j.beproc.2014.01.015

Normando S, Corain L, Salvadoretti M, Meers L and Valsecchi P 2009 Effects of an enhanced human interaction program on shelter dogs' behaviour analysed using a novel nonparametric test. *Applied Animal Behaviour Science 116*: 211-219. https://doi.org/10.1016/j.applanim.2008.10.005

Notari L 2009 Stress in veterinary behavioural medicine. In: Horwitz D and Mills D (eds) *BSAVA Manual of Canine and Feline Behavioural Medicine, Second Edition* pp 136-145. British Small Animal Veterinary Association: Gloucester, UK. https://doi.org/10.22233/9781905319879.13

Overall KL 2003 Medical differentials with potential behavioral manifestations. *Veterinary Clinical Small Animals 33*: 213-229. https://doi.org/10.1016/S0195-5616(02)00103-1

Part CE, Kiddie JL, Hayes WA, Mills DS, Neville RF, Morton DB and Collins LM 2014 Physiological, physical and behavioural changes in dogs (*Canis familiaris*) when kennelled: testing the validity of stress parameters. *Physiology & Behavior 133*: 260-271. https://doi.org/10.1016/j.physbeh.2014.05.018

Pati S, Panda SK, Acharya AP, Senapati S, Behera M and Behera SS 2015 Evaluation of geriatric changes in dogs. *Veterinary World 8(3)*: 273-278. https://doi.org/10.14202/vetworld.2015.273-278

Paul ES, Harding EJ and Mendl M 2005 Measuring emotional processes in animals: the utility of a cognitive approach. *Neuroscience & Biobehavioral Reviews 29*: 469-491. https://doi.org/10.1016/j.neubiorev.2005.01.002

Payne E, Bennett PC and McGreevy PD 2015 Current perspectives on attachment and bonding in the dog-human dyad. *Psychology Research and Behavior Management 8*: 71. https://doi.org/10.2147/PRBM.S74972

Petak I 2013 Communication patterns within a group of shelter dogs and implications for their welfare. *Journal of Applied Animal Welfare Science 16(2)*: 118-139. https://doi.org/10.1080/10888705.2013.741001

Pomerantz O, Paukner A and Terkel J 2012 Some stereotypic behaviors in rhesus macaques (*Macaca mulatta*) are correlated with both perseveration and the ability to cope with acute stressors. *Behavioural Brain Research 230(1)*: 274-280. https://doi.org/10.1016/j.bbr.2012.02.019

Powell C 2003 The Delphi technique: myths and realities. *Journal of Advanced Nursing 41(4)*: 376-382. https://doi.org/10.1046/j.1365-2648.2003.02537.x

Protopopova A 2016 Effects of sheltering on physiology, immune function, behavior, and the welfare of dogs. *Physiology & Behavior 159*: 95-103. https://doi.org/10.1016/j.physbeh.2016.03.020

Protopopova A, Hauser H, Goldman KJ and Wynne CD 2018 The effects of exercise and calm interactions on in-kennel behavior of shelter dogs. *Behavioural Processes 146*: 54-60. https://doi.org/10.1016/j.beproc.2017.11.013

Sharma DP, Nair PC and Balasubramanian R 2003 Analytical search of problems and prospects of power sector through Delphi study: case study of Kerala State, India. *Energy Policy 31(12)*: 1245-1255. https://doi.org/10.1016/S0301-4215(02)00185-4

Shih H-Y, Paterson MBA, Georgiou F and Phillips CJC 2021 Do canine behavioural assessments and characteristics predict the human-dog interaction when walking on a leash in a shelter setting? *Animals 11(1)*: 26. https://doi.org/10.3390/ani11010026

Shore ER 2005 Returning a recently adopted companion animal: adopters' reasons for and reactions to the failed adoption experience. *Journal of Applied Animal Welfare Science 8(3)*: 187-198. https://doi.org/10.1207/s15327604jaws0803_3

Taylor KD and Mills DS 2007 The effect of the kennel environment on canine welfare: a critical review of experimental studies. *Animal Welfare 16(4)*: 435-447

Te Velde H, Aarts N and Van Woerkum C 2002 Dealing with ambivalence: farmers' and consumers' perceptions of animal welfare in livestock breeding. *Journal of Agricultural and Environmental Ethics 15(2)*: 203-219. https://doi.org/10.1023/A:1015012403331

Topál J, Miklósi Á, Gácsi M, Dóka A, Pongrácz P, Kubinyi E, Virányi Z and Csányi V 2009 The dog as a model for understanding human social behavior. *Advances in the Study of Behavior 39*: 71-116. https://doi.org/10.1016/S0065-3454(09)39003-8

Vanhonacker F, Verbeke W, Van Poucke E and Tuyttens FA 2008 Do citizens and farmers interpret the concept of farm animal welfare differently? *Livestock Science 116(1)*: 126-136. https://doi.org/10.1016/j.livsci.2007.09.017

Walker AM and Selfe J 1996 The Delphi method: a useful tool for the allied health researcher. *British Journal of Therapy and Rehabilitation 3(12)*: 677-681. https://doi.org/10.12968/bjtr.1996.3.12.14731

Walsh K 2016 Chronic pain management in dogs and cats. *In Practice 38(4)*: 155-165. https://doi.org/10.1136/inp.i1489

Webster AJF, Main DCJ and Whay HR 2004 Welfare assessment: indices from clinical observation. *Animal Welfare 13(1)*: 93-98

Weipkema PR and Schouten WPG 1990 Mechanisms of coping in social situations. In: Zayan R and Dantzer R (eds) *Social Stress in Domestic Animals* pp 8-24. Kluwer Academic Publishers: Dordrecht, The Netherlands

Welfare Quality® 2009 *Welfare Quality assessment protocol for pigs (sows and piglets, growing and finishing pigs).* Welfare Quality® Consortium: Leystad, The Netherlands

Wells D, Graham L and Hepper PG 2002 The influence of length of time in a rescue shelter on the behaviour of kennelled dogs. *Animal Welfare 11*: 317-325. https://doi.org/10.1017/S1752756200007201

Wemelsfelder F, Hunter TEA, Mendl MT and Lawrence AB 2001 Assessing the 'whole animal': a free choice profiling approach. *Animal Behaviour 62*: 209-220. https://doi.org/10.1006/anbe.2001.1741

Whay HR, Main DCJ, Green LE and Webster AJF 2003b
Animal-based measures for the assessment of welfare state of
dairy cattle, pigs and laying hens: consensus of expert opinion.
Animal Welfare 12(2): 205-217

Whay HR, Main DCJ, Webster AJF and Green LE 2003a
Assessment of the welfare of dairy cattle using animal-based mea-
surements: Direct observations and investigation of farm records.
Veterinary Records 153: 197-202. https://doi.org/10.1136/vr.153.7.197

Williams PL and Webb C 1994 Delphi Technique: a method-
ological discussion. *Journal of Advanced Nursing 19(1)*: 180-186.
https://doi.org/10.1111/j.1365-2648.1994.tb01066.x

Yousuf MI 2007 Using experts' opinions through Delphi tech-
nique. *Practical Assessment, Research and Evaluation 12(4)*: 1-8

Zilcha-Mano S, Mikulincer M and Shaver PR 2011 An
attachment perspective on human-pet relationships: conceptu-
alisation and assessment of pet attachment orientations.
Journal of Research in Personality 45: 345-357.
https://doi.org/10.1016/j.jrp.2011.04.001

Zupan M, Buskas J, Altimiras J and Keeling LJ 2016 Assessing
positive emotional states in dogs using heart rate and heart rate
variability. *Applied Animal Behaviour Science 155*: 101-111.
https://doi.org/10.1016/j.phsbeh.2015.11.027